moderní statistika

HELMUT SWOBODA

S ÚVODEM NOSITELE NOBELOVY CENY

PROF. DR. RAGNARA FRISCHE

NAKLADATELSTVÍ

SVOBODA

PRAHA 1977

OBSAH

Úvod

nositele Nobelovy ceny profesora dr. Ragnara Frische

Dr. Helmut Swoboda vytvořil svou knihou o moderní statistice neobyčejné dílo. Vytkl si za úkol důvěrně seznámit se statistikou lidi, kteří o ní nemají vůbec žádnou představu. Neobyčejnost jeho činu je v tom, že se mu to v tak velké míře podařilo. Přátelsky vede čtenáře krok za krokem dál a dál. A je pochopitelné, že je nesnadné nedůvěřivě se uzavřít upřímným úmyslům laskavého průvodce.

Po chvíli však čtenář ke svému úžasu zjistí, že vnikl až do některých nejabstraktnějších a nejobtížnějších oblastí statistiky.

Nejpodivnější přitom však je, že se čtenář na svého lstivého učitele nemůže vůbec zlobit, neboť teď vidí vlastníma očima celou nádheru divuplné krajiny, do níž ho autor zavedl. A zvlášť vděčný je za to, že byl upozorněn na některá závažná zneužití statistiky.

Podívejme se jen na tento mistrovsky koncipovaný příklad: příběh muže, který zamýšlí koupit ve Zbohatlíkově pozemek. Tento muž dostane různé, zřejmě zcela si odporující údaje o průměrném ročním příjmu obyvatel Zbohatlíkova. Zprostředkovatel uvedl 82 320 tolarů a vysvětloval, že 20 % obyvatel prý má průměrný roční příjem 309 400 tolarů. Bankovní ředitel informoval, že více než polovina obyvatel má roční příjem přes 29 000 a nejčastější roční příjem je asi 18 000 tolarů. Místní učitel, velmi zběhlý ve statistice, tvrdil, že většina obyvatel má příjem nižší než 7500 tolarů. A nakonec statistický úřad na dotaz sdělil, že všechny tyto zdánlivě si tak odporující údaje jsou pravdivé. Vysvětlilo se to tím, že v tomto malém místě bydlel milionář a že matematické vlastnosti mediánu, aritmetického průměru, harmonického průměru, geometrického průměru a nejčetnější hodnoty se navzájem podstatně liší. Kdo tento případ promyslí, nikdy nezapomene na podstatu věci.

Autor přes zdánlivě tak snadný výklad nikdy neztrácí ze zřetele hlubší logické souvislosti. Jen se například podívejte, jak zdůrazňuje, že statistické ověřování hypotéz podle teorie Jerzy Neymana a E. S. Pearsona není ničím jiným než pokusem hypotézu zamítnout. Takové ověřování nikdy nemůže vést k závěru, že hypotéza je správná. Statistické ověřování hypotéz se skutečně může provádět mnoha způsoby, s větší nebo menší účinností. Při velmi důkladném postupu je dokonce možné představit si celou řadu dílčích ověření. Při každém z nich může být testovaná hypotéza zamítnuta — a v tom případě to znamená její konec. Jestliže ji však nezamítneme, můžeme pouze říci: „Provedeným testem jsme nemohli hypotézu zamítnout." Je možné, že se ji jednou podaří zamítnout některým jiným testem. Podobá se to dostihovému závodu s neomezeným trváním. Na každém skoku může kůň padnout, a tím by byl konec jeho závodění. Nepadne-li však, zbývá jen jedno — pokračovat v závodě.

Často se zlobím na nepřesný způsob vyjadřování, s nímž se můžeme setkat

i v přísně vědeckých pracích, že se totiž výraz „nezamítnuto" pokládá za významově stejný s „uznáno správným". Shoduji se s autorem, když tvrdě trvá na logickém rozdílu mezi těmito dvěma tvrzeními.
Jestliže se však provádějí stále přesnější ověřování, jejichž výsledkem je stále stejná odpověď: „při tomto zvláštním testu nebyla daná hypotéza zamítnuta", dojdeme nakonec k praktickému závěru, že v případě, kdy bychom byli donuceni jednat, jednali bychom tak, jako kdyby hypotéza byla správná. Avšak to je docela něco jiného než zjištění: „hypotéza je správná".

Je nápadné, že autor při svém vysvětlování velmi často používá ilustrací, obrázků mužů, žen a dětí. Dělá to proto, že ilustrace jsou pro většinu lidí velmi působivé. Ale lidé a jejich obliby jsou různé a někteří — třeba já — rozhodně dávají přednost abstraktnímu a systematickému výkladu.
Chtěl bych tedy pro ty čtenáře, kteří snad z této knihy a z doplňující literatury v ní uvedené již získali určité znalosti z oblasti statistiky, velmi stručně načrtnout logické základy teorie statistického ověřování hypotéz. Čtenáři, kteří se s prvními základy statistiky teprve seznamují, mohou tuto část klidně přeskočit, rozhodně se jí však nesmějí nechat zastrašit — rozumět jí není vůbec předpokladem pro pochopení této knihy!
Následuje tedy poněkud obtížnější část, určená vědecky orientovanému čtenáři.

U hypotézy v technickém a statistickém smyslu se soustřeďujeme na charakteristický způsob rozdělení pravděpodobnosti určitého jevu. Je zkrátka třeba určit speciální způsob tohoto rozdělení. Přitom nemusí jít o nějaký nový způsob, o němž se domníváme, že je správný. Někdy to může být i způsob, který pokládáme za nesprávný. Máme však přitom v úmyslu najít právě pomocí tohoto testu něco, co opravňuje daný způsob rozdělení pravděpodobnosti zamítnout.
Při statistickém ověřování hypotéz musíme vždy již něco vědět předem, nezačínáme tedy s čistým stolem. Volba ověřované hypotézy nesmí proto být zcela bezplánovitá. Ověřovaná hypotéza musí patřit k určité větší skupině hypotéz, k tzv. přípustným hypotézám, které označme Ω. Nejdůležitější je, aby tato přípustná skupina Ω byla určena.
Z větší skupiny vybereme podskupinu ω. To je ověřovaná hypotéza (někdy se označuje jako nulová hypotéza). Tato podskupina ω může být zčásti nebo plně specifikována. Zčásti bude specifikována, když způsob odpovídajícího rozdělení pravděpodobnosti obsahuje alespoň 1 parametr, který není číselně přesně stanoven. Význam takové hypotézy je prostě tento: „Je zde přesně definováno rozdělení pravděpodobnosti, které obsahuje tyto dosud číselně neurčené parametry."
Dodatkem k těmto číselně neurčeným parametrům může zčásti specifikovaná hypotéza, která se má ověřit, obsahovat jeden nebo více číselně určených parametrů. Testovaná hypotéza je plně specifikována, když všechny nejisté parametry rozdělení pravděpodobnosti jsou určeny číslem.
Ověřovanou hypotézu, popsanou uvedeným způsobem, budeme muset často testovat oproti jiné alternativní hypotéze uvnitř množiny přípustných hypotéz. Jakmile se dojde ke skutečné konstrukci určité testovací metody, uvidíme, jak

je nutné mít na zřeteli kromě testované hypotézy také některé jiné hypotézy, s nimiž se testovaná hypotéza porovná nebo, chcete-li, od nichž by se testovaná hypotéza měla odlišit. To dále ještě vyložím.
K provedení testu musíme použít statistického pozorování, na jehož podkladě se pokusíme určit, zda „testovaná hypotéza se zamítá" nebo zda „testovaná hypotéza se nezamítá".
Testovaná hypotéza má někdy určité charakteristické vlastnosti. Nebo obráceně, hypotéza bez zvláštních vlastností může oproti testované hypotéze být hypotézou se zvláštními vlastnostmi.
Proto musíme při konstruování metody testování hypotézy rozlišovat mezi dvěma druhy chyb, jichž se můžeme dopustit — chyby zamítnutí a chyby nezamítnutí.
Chyby zamítnutí (chyby prvního druhu) jsou chyby, které děláme, když testovanou hypotézu zamítneme v případě, v němž by tato hypotéza zamítnuta být neměla.
Chyby nezamítnutí (chyby druhého druhu) jsou chyby, jichž se dopouštíme, když testovanou hypotézu nezamítneme v případě, kdy by tato hypotéza musela být zamítnuta.
Stojíme před dvěma problémy. Na jedné straně bychom rádi provedli ověření takovým způsobem, aby vznikla jen nepatrná pravděpodobnost chybného zamítnutí. To znamená, že chceme být „fair" vůči hypotéze, kterou ověřujeme; neprovedeme tedy žádné nepřípustné zamítnutí. Na druhé straně bychom rádi zmenšili i možnost chyby nezamítnutí. To znamená, že nechceme svou slušnost přehánět a upustit od zamítnutí takové hypotézy v případech, v nichž by zamítnuta být měla. Musíme proto velmi uvážlivě kormidlovat mezi Scyllou a Charybdou. Podle druhu postupu, který chceme zvolit na základě ověření a v souladu s materiálními důsledky, které z našeho postupu mohou vyplynout, se někdy budeme nejvíce obávat toho, abychom se nedopustili chyby zamítnutí, jindy zase chyby nezamítnutí. Historicky to byl začátek.
Jak tedy najdeme nějaká kvantitativní měřítka, která nám umožní proplout mezi touto Scyllou a Charybdou?
Především použijeme pomocně něco, co označíme jako obor zamítnutí (kritický obor). Je to obor v prostoru pozorování, který je uspořádán tak, že testovaná hypotéza se zamítne, pokud vlastní pozorovaná hodnota spadá do tohoto oboru (nebo na jeho hranici). Leží-li však hodnota mimo obor zamítnutí, testovaná hypotéza se nezamítne. Prostor pozorování je prostě n-dimenzní prostor, v němž je obsaženo těchto n pozorovaných statistických údajů. Zvláštním případem je, když těchto n údajů jsou pozorováním konkrétně definované proměnné.
Je logické, že konstrukce oboru zamítnutí musí být zpracována dříve, než provedeme skutečná statistická pozorování. To je důležité. Druhý náš krok spočívá v tom, že stanovíme malou pravděpodobnost, např. 0,05, 0,02 nebo 0,01, tzv. pravděpodobnost zamítnutí. Je to pravděpodobnost toho, že pozorovaná hodnota spadá do kritického oboru s podmínkou, že ověřovaná hypotéza je správná. Je to zřejmě obvyklý způsob, jak udržet na nejnižším stupni pravděpodobnost, že uděláme chybu prvního druhu.
Je-li např. zpravodajským materiálem aritmetický průměr n nezávislých pozorování a testovaná hypotéza zní, že tento průměr má normální rozdělení s danou

střední hodnotou a danou směrodatnou odchylkou, je možno z tabulky normálního rozdělení vyčíst kritický obor, který odpovídá pravděpodobnosti zamítnutí 0,05, 0,02 nebo 0,01.

Nikoliv však jednoznačně! Pro každou udanou pravděpodobnost zamítnutí lze zde najít tři možné kritické obory úplně stejně, jako když vzpomeneme na odchylky od střední hodnoty jen směrem nahoru nebo jen směrem dolů anebo na odchylky bez ohledu na znaménka. Tyto tři možnosti konstrukce kritického oboru s proměnnou podle dané pravděpodobnosti zamítnutí jsou jen zcela zvláštním případem všeobecné neurčitosti kritického oboru, k níž dochází, všímáme-li si pouze obvyklým způsobem udané pravděpodobnosti α. Pozorujme nyní neurčitost kritického oboru poněkud blíže v n dimenzích (přičemž n je počet statistických údajů, které určují prostor pozorování). Předpokládejme, že je známa hustota pravděpodobnosti podle testované hypotézy v každém libovolném bodu prostoru pozorování. Když je udána pravděpodobnost zamítnutí (např. 0,05 nebo 0,02 či 0,01), zřejmě to nedostačuje k jednoznačnému určení kritického oboru. Všeobecně bude mnoho oborů toho druhu, kde celková pravděpodobnost oboru se rovná α.

Ukazuje se nutnost zavést jinou přípustnou hypotézu, se kterou musíme porovnat hypotézu původní, to však znamená jinou hypotézu, proti níž se musí testovat naše původní hypotéza.

To je jádrem moderní teorie statistického ověřování hypotéz. K definování toho, co rozumíme „nejsilnějším" (nejlepším) kritickým oborem, musíme určit jinou hypotézu, která nás nejvíce zajímá, pro srovnání s testovanou hypotézou.

Vede to k následujícímu pravidlu: Ze všech oborů v prostoru pozorování, které odpovídají pravděpodobnosti zamítnutí rovné udané hodnotě α podle rozdělení pravděpodobnosti testované hypotézy, vybereme ten obor (nebo ty obory), jenž ukazuje nejvyšší stupeň pravděpodobnosti podle rozdělení pravděpodobnosti jiné hypotézy, s níž testovaná hypotéza má být porovnána. Maximalizovat tuto pravděpodobnost je zřejmě totéž jako udržet na pokud možno nejnižším stupni pravděpodobnost připuštění chyby nezamítnutí.

Říkalo se, že ve „středověku" statistiky byl problém „nejsilnější" metody ověřování hypotéz diskutován se zřetelem na hypotézu *per se*. V moderní době se přihlíží ke skutečnosti, že „nejsilnější" metoda, tj. „nejsilnější" konstrukce kritického oboru, závisí na jiné přípustné hypotéze, s níž se musí porovnat hypotéza původní.

Je několik případů, kdy „nejsilnější" testovací metoda pro danou hypotézu je táž bez ohledu na to, kterou jinou přípustnou hypotézu s ní chceme porovnat. To se označuje jako případ, v němž je jeden stejnoměrně nejsilnější test. Takové případy jsou však vzácné, takže většinou si musíme zároveň všímat testované hypotézy a jiné přípustné hypotézy, s níž se p vní má porovnat.

Souhrnně řečeno: H je testovaná hypotéza, H^* je hypotéza, proti níž má H být testována. X budiž obor v prostoru pozorování a x bod v tomto prostoru.

Nejsilnější kritický obor pro H, když se má testovat H proti H^* s pravděpodobností zamítnutí, je určen:

(1) $P\{xcX|H^*\} \to \max!$ (c znamená „obsaženo v", $\to \max!$ znamená „má se maximalizovat")

(2) za vedlejší podmínky $P\{xcX|H\} = \alpha$

$P\{xcX|H^*\}$ v (1) označuje pravděpodobnost toho, že realizace bodu x v X je za předpokladu rozdělení pravděpodobnosti H^*. Zamítnout H znamená totéž jako dostat jeden bod x v X. To bychom přirozeně rádi viděli tak často, jak jen možno podle rozdělení pravděpodobnosti H^*. Potud pravidlo (1). Neradi bychom to však viděli často podle rozdělení pravděpodobnosti H. Podle tohoto rozdělení pravděpodobnosti bychom se s tím raději setkávali zřídka. Potud pravidlo (2).

Jednoduchý příklad

Hustota normálního rozdělení n nezávislých pozorování $x_1, x_2, \ldots x_n$ se stejnou střední hodnotou μ a stejnou směrodatnou odchylkou σ je určena:

$$(3) \quad \frac{1}{(2\pi)^{\frac{n}{2}} \sigma^n} \exp. \left(- \frac{1}{2\sigma^2} \sum_{i=1}^{n} (x_i - \mu)^2 \right)$$

(4) Ω má být libovolné rozdělení pravděpodobnosti tvaru (3) s konečným μ a σ. Chceme testovat hypotézu $\mu = \mu_0$, přičemž μ_0 je předem udané číslo. A tuto zkoušku chceme testovat proti hypotéze $\mu > \mu_0$, přičemž σ budiž známá konstanta. Nejdříve si všimneme jednoduššího problému testu $\mu = \mu_0$ proti $\mu = \mu_1$, přičemž μ_1 je jiné dané číslo $\mu_1 > \mu_0$. Obě tyto hypotézy jsou plně určeny. Problém (1) až (2) může být proto řešen přímo s použitím Lagrangeova multiplikátoru.

Výsledkem je, že

$$(5) \quad \bar{x} \geqq \mu_0 + \xi_\alpha \frac{\sigma}{\sqrt{n}}$$

je nejsilnější kritický obor pro hypotézu $\mu = \mu_0$, testováno proti hypotéze $\mu = \mu_1$, přičemž $\mu_1 > \mu_0$. V (5) je ξ_α horní mezí normálního rozdělení podle pravděpodobnosti zamítnutí α. Jinými slovy: vztah mezi ξ_α a α je dán

$$(6) \quad \frac{1}{\sqrt{2\pi}} \int_{\xi=\xi_\alpha}^{+\infty} e^{-\frac{\xi^2}{2}} d\xi = \alpha \; (\xi_\alpha \text{ pozitivní}).$$

Tento vztah mezi ξ_α a α lze vyčíst z tabulky normálního rozdělení. V (5) je $\bar{x}$ prostě nevážený aritmetický průměr pozorování. Pozoruhodné je, že μ_1 se v (5) nevyskytuje. Podle toho je stejný kritický obor nejsilnější, když hypotéza $\mu = \mu_0$ se testuje proti nějaké jiné hypotéze spočívající v tom, že se řekne: střední hodnota budiž větší než μ_0. V tomto případě máme proto stejnoměrně nejsilnější test pro $\mu = \mu_0$, totiž ten, který je určen kritickým oborem (5). Není však žádný stejnoměrně nejsilnější test proti hypotéze $\mu \neq \mu_0$.

Funkce síly testu

Předpokládejme, že nějakým způsobem byl určen kritický obor X pro hypotézu H. Předpokládejme X jako konstantu. Tu se může objevit zajímavá otázka: Jak ostře oddělující bude tento test pro H s daným X, když se bude testovat H proti libovolně vybrané hypotéze H^{**} v Ω? To je dáno funkcí „síly testu" („silofunkcí").

(7) $P\{x \subset X | H^{**}\}$,

přičemž H^{**} se může obměňovat se všemi H^{**} v Ω. V případě, že různé hypotézy H^{**} jsou charakterizovány hodnotou jediného parametru Θ, může se psát (7) jako

(8) $P\{x \subset X | \Theta\}$.

Ideální tvar „silofunkce" (8) je ten, který v (8) je blízko 1 pro všechny hodnoty Θ s výjimkou zvláštní hodnoty Θ, jež charakterizuje ověřovanou hypotézu. Pro tuto zvláštní hodnotu Θ má (8) tak rychle, jak jen možno, poklesnout k pravděpodobnosti zamítnutí α.

„Silofunkce" (7) — nebo její zvláštní případ (8) — se nazývá „nezkresleně", pokud jde o pravděpodobnost, nebo „maximálně pravděpodobný odhad" — když její minimum leží při $H^{**} = H$.

Chceme-li v našem předcházejícím příkladu ověřovat hypotézu $\mu = \mu_0$ proti hypotéze $\mu \neq \mu_0$, nemáme — jak jsem se zmínil — žádný stejnoměrně nejostřeji odlišující test. Jestliže však přidáme podmínku, že „silofunkce" má být nezkreslená, tzn. když mezi všemi testy s nezkreslenou „silofunkcí" hledáme ten, který splňuje (1) a (2), zjistíme, že je jeden stejnoměrně nejostřeji odlišující test, totiž ten, který je charakterizován

(9) $\left|\bar{x} - \mu_0\right| \geqq \frac{\sigma}{\sqrt{n}} \xi^{**}_{\alpha}$, přičemž ξ^{**}_{α} je určeno vzorcem

(10) $$\frac{1}{\sqrt{2\pi}}\left|\int_{\xi=-\infty}^{-\xi^{**}_{\alpha}} e^{-\frac{\xi^2}{2}} d\xi + \int_{\xi=+\xi^{**}_{\alpha}}^{+\infty} e^{-\frac{\xi^2}{2}} d\xi\right| = \alpha$$

Uvedené výsledky jsou přirozeně jen některé krajně jednoduché příklady. Další případy s důkazy jsou uvedeny v mých norských přednáškách o ověřování hypotéz (Memorandum z 9. srpna 1952, Universita Oslo, Sociálně ekonomický ústav).

Venkovské sídlo v Røyse, Norsko, svatodušní neděle 1971.

Ragnar Frisch

Autor tohoto úvodu, profesor dr. Ragnar Frisch, se narodil 3. března 1895 v Oslo; studoval na tamější univerzitě hospodářské vědy a filozofii. V roce 1931 byl povolán na univerzitu v Oslo v hodnosti profesora hospodářských věd. Od roku 1932 je ředitelem Sociálně ekonomického ústavu univerzity.

Profesor Frisch patří ke spoluzakladatelům Ekonometrické společnosti, existující od roku 1931, jejíž časopis „Econometrica" dlouhý čas vydával (od r. 1935 do r. 1955). Je členem četných vědeckých akademií a společností v různých zemích a čestným doktorem mnoha univerzit. K jeho hlavním dílům patří „Maxima a minima. Teorie a hospodářské využití." Na základě svých prací, které vytyčovaly směr v oblasti konjunktury a růstu, dostal v roce 1969 (společně s holandským národohospodářem Janem Tinbergenem) Nobelovu cenu za ekonomii, která byla v tomto roce udělena poprvé. Cena mu byla udělena za práce k vývoji modelů pro rozbor hospodářských procesů.

1 Co je statistika?

1.1 Od hřbitova čísel k pomůcce při rozhodování

„Jsou tři druhy lží: lži, odsouzeníhodné lži a statistiky." Toto rčení se vyskytuje v různých variantách, jako např. jsou lži, lži z nouze a statistiky, a představuje již asi 100 let nejčastěji citovaný výrok o statistice. Všechny verze uvedeného rčení, ať už jsou vyhroceny jakkoliv, mají jedno společné: naznačují, že *statistika je zvlášť rafinovanou formou lži.*

Je kuriózní, že mezi vědci nikdy nedošlo ke shodě, komu tento zlomyslný výrok připsat. Nejčastěji se autorství připisuje Benjaminu Disraelimu, jindy se jako autor uvádí protivník Disraeliho lord Palmerston. Je pravděpodobné, že jde o zlomyslnou jízlivost některého člena parlamentu, který asi byl statistickými údaji zahnán do úzkých, a když se mu obrana nezdařila jeho vlastními statistikami, bránil se tímto způsobem. Velmi podivné však je, že poznámka, která byla před 100 lety přijata v britském parlamentě s úsměškem, se stále ještě cituje po celém světě. Ještě podivnější je, že se většinou cituje nekriticky, bez nebo téměř bez znalosti toho, co statistika vlastně je. Jestliže se stále opakuje, že statistiky lžou, začne se tomu nakonec věřit tím spíše, že čas od času se vyskytne možnost přesvědčit se, že ta nebo jiná statistika *skutečně* „lže".

Tento obecně rozšířený názor je však podivuhodným způsobem spojen se zcela opačným postojem, totiž že *statistiky se zároveň považují za vrchol nevyvratitelného,* neboť mají magické kouzlo matematické přesnosti — a co by mohlo být přesvědčivější než číselný údaj? Proto platí: kdo potřebuje nezvratný důkaz, vyzbrojí se statistickými údaji a rozbory a v rozhodujícím okamžiku je hodí na stůl: zde — *statistika dokazuje.*

Většinou pak proti sobě stojí dvě statistiky, ze kterých se odvozují zcela protichůdná tvrzení a úplně zmatený laik je jen posilován ve svém přesvědčení, že statistiky skutečně lžou. Je-li však v přemýšlivé náladě, pak se ptá, jak je možné, že dochází k tak těsnému a zdánlivě neodlučitelnému propojení matematické přesnosti se zřejmou lží. Statistika se v důsledku toho nakonec stává pro laika téměř tajemnou vědou.

To vše by nebylo tak zlé, kdyby šlo o vědu nebo pracovní metodu, s níž se normální občan obyčejně dostane jen zřídkakdy do styku. Se statistikou je to však jinak. Setkáváme se s ní nejen ve zprávách, informacích a poznámkách všeho druhu, ale tvoří zároveň *základnu pro plánování, organizaci a celé moderní soužití.* Ačkoliv se ve své moderní podobě zabývá i malými vzorky, poskytuje v zásadě návod k zacházení se soubory a četnostmi, s rozsáhlými číselnými údaji o věcech a lidech. Ve světě plánování a systematické přípravy budoucnosti je základem nesčetných propočtů a nesčetných rozhodnutí. Asi není správné tvrzení, že ze života má *více* ten, kdo se zabývá statistikou. Jisté však je, že takový člověk *ví* více o zákonitostech života a *že se i v budoucnu o nich více dozví.*

Kde však hledat kořeny tohoto zvláštního nerozhodného postoje ke statistice? Mohou za to především *dvě skutečnosti:* za prvé nedostatečná znalost cílů, metod a možností statistiky a za druhé, že za statistiku se pokládá i to, co je ve skutečnosti pseudostatistikou.

K prvnímu důvodu jednou Horace Levinson velmi správně poznamenal: „Statistika se těší pochybnému vyznamenání tím, že je nejvíce nepochopeným vědním oborem. Neznamená to však, že je nejméně známá. Nepochopení nějaké věci totiž předpokládá, že se o ní něco ví nebo přinejmenším se myslí, že se ví... Se statistikou je to však tak, že panuje všeobecné mínění, že z každého, kdo se ve škole naučil trochu počítat, lze bez obtíží udělat statistika prostě tak, jako když se udělí titul."

První příčina však přímo vyvolává příčinu druhou. Protože všude vládne mylný názor, že k vypracování statistiky postačují jen obstojné znalosti počtů, stává se statistika velmi lehce rejdištěm amatérů. Výsledek jejich práce se pak posílá do světa jako „statistika", kterou šikovný novinář případně opatří tučným titulkem, jenž se třeba k obsahu ani nehodí, a čtenář pak bere toto pochybné dílo jako vrchol moudrosti nebo jako další důkaz, že statistikové marní svůj čas pošetilostmi nebo že pomocí čísel fabrikují lži.

Ve středověku a v raném novověku byl všeobecný postoj vůči medicíně velmi podobný tomu, jaký se dnes zaujímá ke statistice, a to ze zcela obdobného důvodu. Pro laika bylo prostě příliš obtížné rozpoznat rozdíl mezi lazebníkem a lékařem, mezi mastičkářem a učencem.

U statistiky k tomu přistupuje ještě jedna obtížná okolnost. Lékař a pacient jsou zpravidla v osobním důvěrném styku, který se rozkolísá, jakmile vznikne podezření, že lékař je ve službách jiné, „nepřátelské" osoby, jako je tomu např. u odvodního lékaře za války nebo u posudkového lékaře nemocenské pojišťovny. K soukromému statistikovi se zpravidla na konzultaci nechodí, a tak důvěra ke statistikovi, který je v „cizích" službách, kolísá od bezmyšlenkovité víry, pokud nedochází k ohrožení vlastního zájmu, až po bezmocný vztek, když se operuje číselným materiálem, který by mohl vlastním zájmům škodit. Pak statistika nesporně „lže" a je-li zájmová skupina, která se cítí ohrožena, dosti majetná a aktivní, rychle angažuje vlastního statistika, aby mohla jinými čísly prokázat *vlastní* pravdu.

V dalším textu budeme mít ještě dost příležitostí ukázat, proč se zdá, že statistiky tak častou lžou. Nyní jen heslovitě některé hlavní příčiny: nepozorné čtení přesně formulovaných podmínek šetření, neúmyslně nebo úmyslně mylná interpretace jednotlivými instancemi, a *last not least* (poslední, ne však méně důležitá — angl.) stará moudrost, že každá věc má přinejmenším dvě stránky.

Tato kniha sleduje současně *dva* navzájem zcela slučitelné cíle: Poskytnout *kurs první pomoci proti špatným statistikám*, tj. ukázat, před kterými chybami, klamnými triky, omyly a úskoky se má uživatel statistik chránit (popř. také jejich zpracovatel, jde-li o statistika amatéra). Dále má kniha poskytnout *úvod do myšlenkových základů a rutinních postupů vědy a metody*, jejíž význam stále roste už i proto, že moderní statistika, jak brzy uvidíme, nemá téměř nic společného s těmi „hřbitovy čísel", které před 100 a snad ještě i před 50 lety byly pro tuto vědu charakteristické.

„Statistika je dnes výkonnější než dříve," řekl lapidárně Erwin Kreyszig před několika lety, a jako důvod dodal: „Dříve se musela omezit na prostý empirický popis, dnes může posloužit při posuzování situací, protože byla vhodně rozvinuta její matematická základna." Již na tomto místě lze uvést to, čím se budeme dále zabývat častěji: Moderní statistika klade ve své teoretické i praktické práci hlavní důraz na analýzu výběrových šetření, a ne na hromadná vyčerpávající šetření.

Proto je statistika dnes více než kdykoliv jindy ve všech životních situacích a povoláních — od domácnosti až po světovou politiku — především *pomocí při rozhodování*. Vím něco, ale ne vše — jak se zachovat pokud možno nejracionálněji? Wallis a Roberts uvedli svou učebnici „*Methoden der Statistik*" (Statistické metody) právem větou: „*Statistika je soubor metod, které nám umožňují činit rozumná rozhodnutí v případě nejistoty.*" Je však nepochybné, že *moderní statistika tvoří základ teorie rozhodování.*

Zároveň však nutno dodat, že starší způsoby statistické práce nebyly ještě plně překonány a zatlačeny novými metodami a teoriemi, ale spíše přeformulovány a obohaceny. Statistický rozbor umožňuje např. mnohem rozsáhlejší a zčásti dokonce přesnější odhady s menšími náklady než dřívější způsoby vyčerpávajících hromadných šetření.

Ještě poznámku k *dvojímu významu pojmu* „statistika". Nejen v němčině a v češtině se tímto slovem označuje jak jednotlivé statisticky vyjádřené šetření, tak i věda, která se těmito pracemi zabývá v nejširším smyslu. Je proto nutno rozlišovat mezi *„jednou (dílčí) statistikou"* a „statistikou", přičemž se samozřejmě ihned ukáže, co jsme právě naznačili: v průběhu staletí, a především v průběhu našeho, dvacátého století se význam statistiky podstatně změnil.

1.2 Od nauky o státu k analytické statistice

Před více než 100 lety napsal Gustav Rümelin, že existuje nejméně 63 definic pojmu „statistika". Od té doby jich přibylo daleko více, neboť v Rümelinově době ještě nenastala ona dalekosáhlá přeměna, která úplně změnila strukturu statistiky. Budeme ještě mít příležitost ukázat některé podrobnosti z dějin statistiky. Zde postačí jen stručný náčrt tří etap, na které lze rozdělit její historický vývoj.

Nejstarší statistikou je „*popis státu*", spočívající v zobrazení daného zeměpisného, hospodářského a politického stavu. „Status" je stav, ale také stát, neboť stát sám je stav, totiž „status rei publicae", stav společenství. Jedno z prvních státovědných děl vyšlo v r. 1562 v Benátkách: Francesco Sansovina, „*Del governo et amministratione di diversi regni*" (O vládě a správě v různých královstvích). Přesně o 100 let později uveřejnil Veit Ludwig von Seckendorff svou knihu „*Teutscher Fürsten-Staat*" (Německý knížecí stát), státovědnou příručku, která byla opakovaně vydávána až do 18. století. Ve stejné době měl historik práv a lékař Hermann

DEL GOVERNO
ET AMMINISTRATIONE
DI DIVERSI REGNI, ET REPVBLICHE,
Cosi antiche, come moderne,
DI M. FRANCESCO SANSOVINO
Libri XXII.

Cioè di Germania. Francia. Spagna. Inghilterra. Polonia. Portogallo. Napoli. Turchia. Persia. Tunisi. Fessa. Roma antica. Roma moderna. Athene. Sparta. Venetia. Genova. Lucca. Suizzeri. Norimberga. Ragugi. Vtopia.

Ne' quali si contengono diversi ordini, leggi, Magistrati, vsanze, costumi, & altre cose notabili, appartenenti alla historia, vtili ad ogni buomo di Stato & ciuile, & buone cosi à tempi di pace, come di guerra.

CON PRIVILEGIO.

In Vinegia, Presso Altobello Salicato, M D LXXXIII.
Alla Libraria della Fortezza.
238467-B

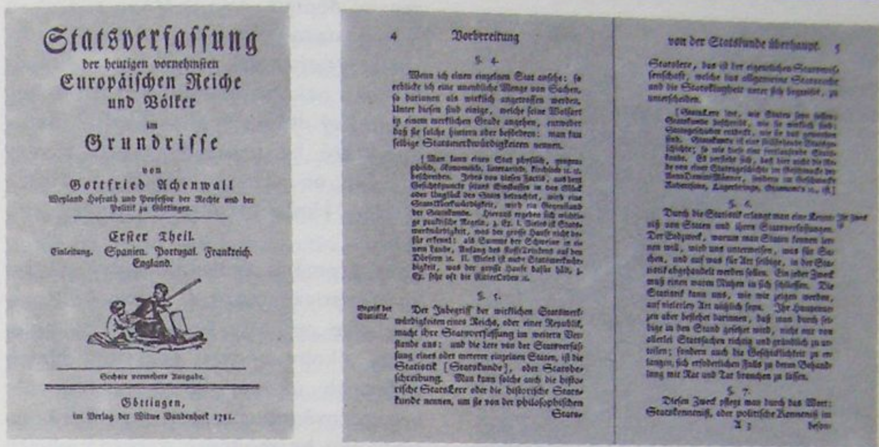
Nejstarší statistika ještě ve smyslu učení o státě je sbírka popisů států „Del governo... di diversi regni" (1562) od Franceska Sansovina, ve které je dokonce popsán ideální stát „Utopie" Thomase Moora.

Conring v Lipsku první přednášky o popisné státovědě.
První velký teoretik statistiky v německé jazykové oblasti, Gottfried Achenwall, se již v polovině 18. století výslovně odvolává na etymologický původ slova „status" jako „stav" a jako „stát". Statistika měla ukazovat „skutečné pamětihodnosti měšťanské společnosti" — podle našeho způsobu vyjadřování důležité zvláštnosti státu. Poslední zbytky této nejstarší podoby statistické vědy ještě dnes nalézáme na prvních stránkách statistických ročenek řady států, kde se uvádí řada geografických údajů, jako např. délka hranic, počet ostrovů a jejich rozloha, nejvyšší horské vrcholy a místa s nejnižší nadmořskou výškou, délka toků řek, klimatické údaje apod.

Zcela jiný okruh statistiky vznikl mezitím v Anglii, a to takzvaná *„politická aritmetika"*, která vycházela z údajů o narozeních a úmrtích a na tomto základě se pokoušela srovnávat a pozorovat číselný vývoj obyvatelstva za delší časové úseky. Graunt a Petty zpracovali v druhé polovině 17. století základní díla, v Německu dosáhl tento směr v tvrdém boji proti Achenwallově — Schlözerově škole prvního rozhodujícího úspěchu velkolepým výkonem

Teutscher Fürsten-Stat/
Oder:
Gründliche vnd kurtze Beschreibung/
Veit Ludwig von Seckendorff/ auff Obern-Zenn/ der Zeit Fürstl. Sächsischen Hoffrath zu Gotha.
M. DC. LVI
Gedruckt zu Hanaw/ bey Johann Aubry/
In Verlegung Thomæ Matthiæ Götzens/ zu Franckfurth am Mäyn.

„Teutscher Fürsten-Staat" (Německý knížecí stát) od Veita Ludwiga von Seckendorffa představuje předchůdce oněch popisů státu, jimž dali později Achenwall a Schlözer označení „statistika".



Statsverfassung
der heutigen vornehmsten
Europäischen Reiche
und Völker
im
Grundrisse
von
Gottfried Achenwall
Weyland Hofrath und Professor der Rechte und der Politik zu Göttingen.
Erster Theil.
Einleitung. Spanien. Portugal. Frankreich. England.
Göttingen,
im Verlag der Witwe Vandenhoek 1781.

Gottfried Achenwall definoval statistiku jako „vědu o státě", jako popis a zobrazení jednotlivých států, přibližně tak, jak ji dnes poskytují např. díla, jako je „The Statesman's Year Book". Moderní statistika má naproti tomu s Achenwallovou vědou o státě dnes společné jen jméno.

pruského polního kazatele a pozdějšího vrchního konzistorního rady Johanna Petera Süssmilcha.
Süssmilch — „meteor, zářící a osamocený", jak ho později označil teolog a morální statistik von Oettingen — napsal své hlavní dílo „*Betrachtungen über die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, dem Tode und der Fortpflanzung desselben erwiesen" (1741)* (Úvahy o božském pořádku v proměnách lidského rodu, dovozených z jeho narození, smrti a rozmnožování), knihu o statistice obyvatelstva, která v prokazovaných zákonitostech doufá nalézt moudrou vůli boží (obr. str. 22).
V následujících desetiletích se tato politickoaritmetická statistika změnila na světskou statistiku. Matematici a teoretici pravděpodobnosti nacházeli v běhu života lidské společnosti stále více zákonitostí. Co začalo jako velebení boha, stalo se početním příkladem ostrovtipných matematiků (a astronomů). Vzpomenutý vliv základních teoretických koncepcí pravděpodobnosti na statistiku se začal projevovat stále výrazněji. Belgičan Adolphe Quételet vypočítal z různorodých lidských jedinců *„homme moyen"* (průměrného člověka), ideální typ, o který se příroda snaží a který je v podstatě nereálný. Tím vytvořil nejen matematicko-mystickou bájnou bytost, nýbrž i důležitou základnu pro celou budoucí statistiku, totiž koncept *normálního rozdělení, normální křivky, střední hodnoty a rozptylu.*
Z četných matematiků a všestranných vědců, kteří v 18. a 19. století vytvořili cenné a pro další rozvoj statistiky fundamentální práce, se zpravidla uvádějí tři Bernoulliové (Jacob, Daniel a Nicolas), Halley, Lagrange, Euler, dále zejména Laplace, de Moivre a Gauss a v řadě s nimi i geniální outsider, anglický duchovní Thomas Bayes.
Téměř celá statistická činnost v 19. a na začátku 20. století je charakterizována tím, že se statistika *zabývá hromadnými*

Die göttliche
Ordnung
in den Veränderungen
des menſchlichen
Geſchlechts,
aus
der Geburt, Tod, und
Fortpflanzung
deſſelben
erwieſen von
Johann Peter Süßmilch,
Prediger beym hochlöblichen Kalckſteiniſchen
Regiment.
Nebſt einer Vorrede
Herrn Chriſtian Wolffens.
Berlin,
Zu finden bey J. C. Spener.
1741.

J. P. Süsmilch uvedl svým dílem „Göttlicher Ordnung" (Úvahy o božském pořádku) (1741) poprvé do střední Evropy statistiku obyvatelstva. Na podkladě kostelních registrů v Kuronsku prokázal Süsmilch pravidelnost narození a úmrtí.

jevy. Magické heslo znělo: *vyčerpávající šetření*, to znamená přesně zachytit veškeré obyvatelstvo pomocí pečlivě připraveného sčítání lidu, pokud možno vyčerpávající záznamy o demografických a hospodářských jevech. Heslem bylo: čísla, stále více a stále úplnější. To byla statistika, jak si ji většinou představuje laik ještě dnes.

Na přelomu století a zejména pak ve třicátých letech našeho století přichází nová proměna ve vývoji statistiky. Zrodila se *moderní statistika, analytická statistika, induktivní statistika*. K této charakteristice bude nutné později mnohé doplnit. Zde je třeba zdůraznit jen podstatný rozdíl oproti dřívější statistice: jestliže dříve byla snaha namáhavou a nekonečnou prací zjistit každý jednotlivý detail, nyní se hledají metody, které by umožnily tvořit závěry o celku na základě výběru a dílčích šetření. Heslo moderní statistiky zní: *výběr*.

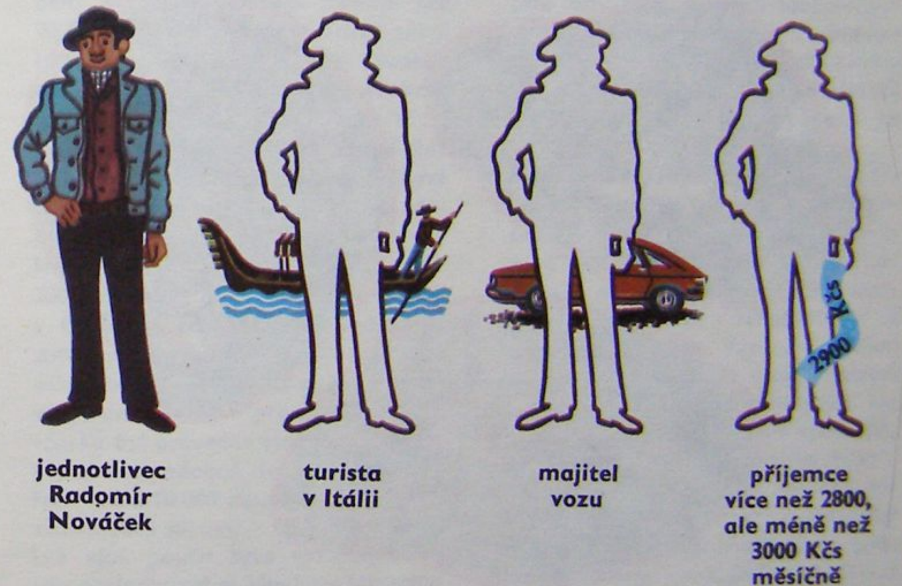
Tato proměna způsobila *úplný převrat* ve statistické teorii a praxi. *Matematika* více než kdy jindy ovládla pole, a to nikoliv pomocí čtyř základních početních úkonů, které postačovaly pro státovědnou statistiku a ještě i pro statistiku hromadných vyčerpávajících šetření. Matematická statistika se vyvinula v *samostatný vědní obor vyšší matematiky*, s vlastními postupy, jako jsou analýza rozptylu, korelační počet a ověřování hypotéz s četnými pracovními metodami.

Základy k tomuto vývoji položili na začátku 20. století Lexis a Bortkiewicz v Německu a Čebyšev, Čuprov, Ljapunov a Markov v Rusku. Brzy potom se tato disciplína stala doménou *Angloameričanů* a v menší míře také *Skandinávců*. Po druhé světové válce začínala ve střední Evropě statistika prakticky úplně znovu, i když Oskar Anderson a někteří jiní se snažili již ve třicátých letech o uplatnění nových metod. Ještě v roce 1965 konstatoval Erwin Kreyszig: „Matematická statistika nemá v německé jazykové oblasti stále ještě dostatečný význam a použití."

Převažující postavení Angloameričanů je především zásluhou velké osobnosti sira Ronalda A. Fischera, který sám nebo spolu s jinými vytvořil většinu dnes obvyklých pracovních metod statistické analýzy. I když jsou dnes on a „biometrická škola" v mnohých podrobnostech překonáni, jeho historický

význam pro statistiku je a zůstane ohromný. K nejvýznamnějším statistikům v anglo-americké oblasti patří dále G. U. Yale, Karl Pearson a Američan polského původu Jerzy Neyman, který společně s Pearsonovým synem E. S. Pearsonem byl otcem teorie výběru. Dnes je vědecky vzdělaný statistik současně také vysoce kvalifikovaným matematikem — i když ne nutně ve všech oborech matematiky. To však nemůže být pro laika důvodem, aby v tom spatřoval potvrzení své nechuti nebo svého odporu ke statistice. Ve všech vědních oborech mohou jen odborníci dosáhnout vysoce kvalifikovaných výsledků. Inteligentní současníci se však přesto snaží porozumět do určité míry nejvýznamnějším vymoženostem a procesům moderního života, i když netvoří součást jejich vlastního pracovního zaměření. Představy o světě dosažené moderní přírodovědou nebo poznatky sociologie jsou důležité i pro toho, kdo není přírodovědcem nebo sociologem.

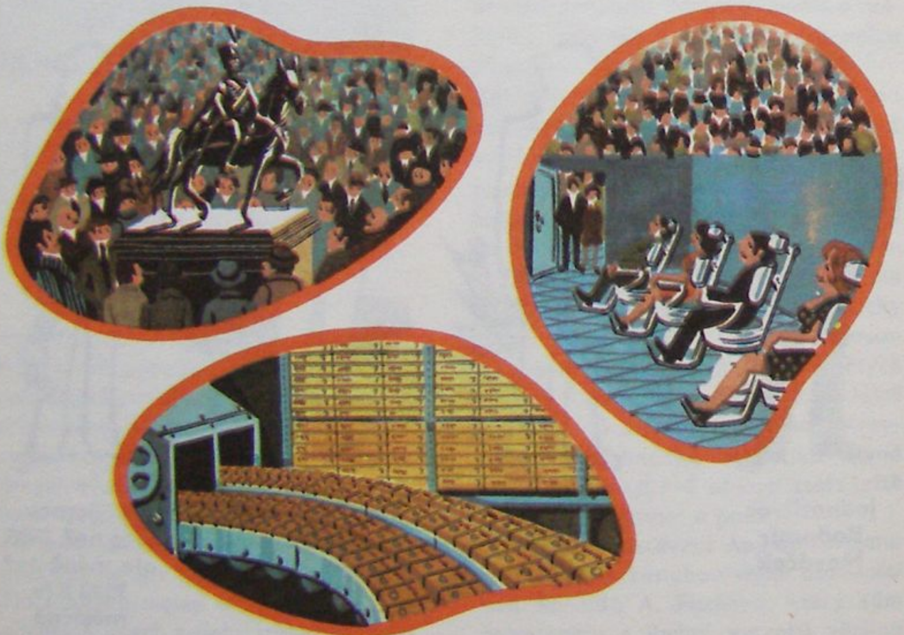
Stejně je tomu i se statistikou: v době, kdy roste význam statistických metod a informací, je znalost nejdůležitějších zásad statistiky potřebná pro každého, koho povolání nebo i běžný život přivedou do styku se „statistikami", a to je dnes každý, jehož zájem jde i jen tak daleko, že pravidelně čte svůj denní tisk.



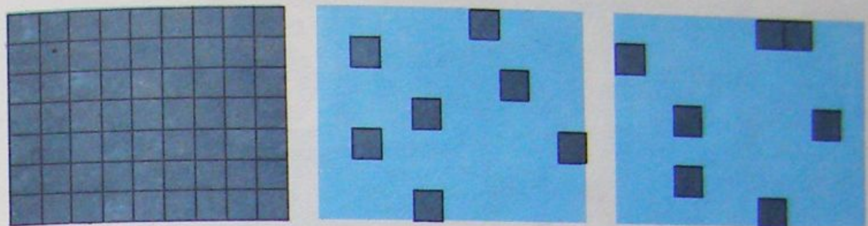
Jak může statistika jako učení o hromadných jevech vypovídat něco rozumného o jednotlivci? Může a dělá to tím způsobem, že nezkoumá jednotlivce v individuální mnohosti jeho činnosti, nýbrž jen jako anonymního nositele některé činnosti vedle mnoha jiných, např. jako turistu v Itálii, jako majitele vozu, jako příjemce mzdy určité příjmové skupiny atd.

1.3 Základní soubor — rozdělení

Hlavním obsahem této knihy je uvést a vysvětlit nejdůležitější statistické pojmy a metody. Nemělo by proto příliš mnoho smyslu, kdybychom je hned na začátku uvedli vyčerpávajícími definicemi, neboť se mohou stát srozumitelnými teprve ve vzájemných souvislostech. Základní myšlenku, která se jako červená nit vine dalším výkladem, chceme však nastínit hned na samém počátku.
Statistika se zabývá i ve své moderní podobě výběrové analýzy zásadně *hromadnými jevy*. Tím se nevylučuje, že ve výběru nejsou přesně zkoumány *jednotlivé věci nebo osoby*, avšak nikoliv proto, aby se zjistila jejich individuálnost, nýbrž proto, aby se zjistila *existence nebo neexistence nějakého znaku*, o němž se domníváme, že se vyskytuje i jinde, znaku, který je rozložen v základním souboru.
S následujícími dvěma slovy se budeme setkávat často. Jedno zní „*základní soubor*“, druhé „*rozdělení*“. Podívejme se nejdříve krátce na *základní soubor*. Je to většinou myšlenková konstrukce, která se nevyskytuje v přirozeném stavu, ale je formulována teprve v průběhu statistické činnosti. Takovým základním souborem může být „obyvatelstvo světa“ nebo „obyvatelstvo

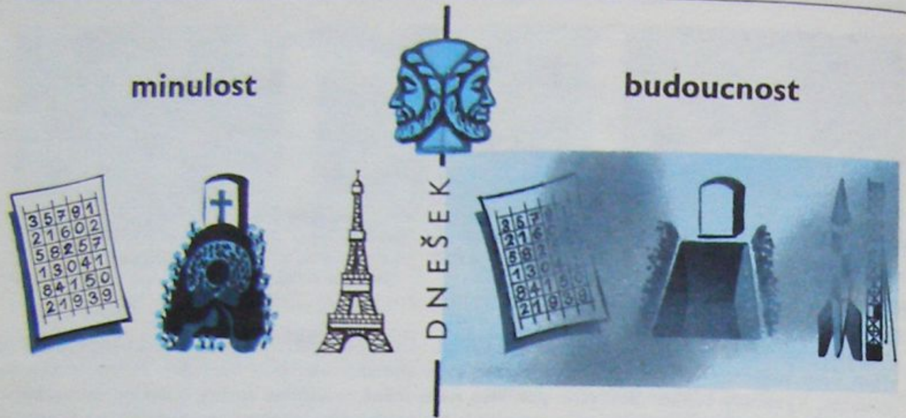


„Statistická masa“, „základní soubor“ může být tvořen podle libosti — musí však zahrnovat všechny jednotlivé prvky souboru: např. všechny ženaté muže v Brně, pacienty městské zubní kliniky v minulém roce nebo výrobu cigaret v závodě Y v měsíci dubnu.



Je-li nemožné nebo příliš časově náročné, příliš drahé nebo neúčelné zachytit základní soubor ve všech jeho jednotlivých složkách, používá se výběrového souboru a libovolně se vyjme několik jednotek. Výběrový soubor poskytne pak více nebo méně spolehlivé závěry o základním souboru, a to podle rozsahu základního souboru, rozsahu výběrového souboru a rozdělení zkoumaných znaků.

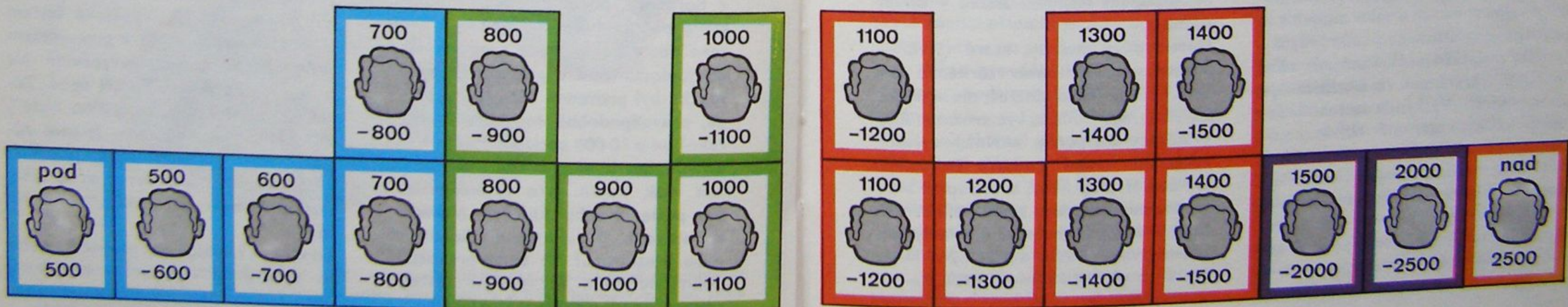
ČSSR k 1. 1. 1972“ nebo také „ženatí muži v Plzni“ anebo „samostatní zemědělci s ročním příjmem menším než 20 000 DM“ či „pacienti městské zubní kliniky v Bordeaux“ nebo „vozidla z roku 1970“ anebo „cigarety vyrobené v měsíci dubnu závodem X f my Y“.
Vymezení základního souboru není vždy docela bez problémů. Často je lákavé pokušení rozšířit jej více, než je zdůvodněno statisticky šetřeným materiálem. Když zkoumáme např. výběr 300 pacientů městské zubní kliniky v Bordeaux podle příjmu, rodinného stavu, počtu zubů, které ještě mají, nebo také podle barvy vlasů anebo tělesné váhy, dostaneme v případě, že výběr vzorku byl proveden statisticky správně, pravděpodobně dobrou představu řekněme o 10 000 pacientech, kteří navštívili tuto kliniku v uplynulém roce. Lze však použít tyto výsledky také pro pacienty zubní kliniky v Marseille nebo Lyonu, či dokonce v Edinburghu nebo Stuttgartu? To bude pravděpodobně nemožné bez dalšího šetření, zejména jde-li o zahraniční kliniku.
Teprve když jsme pomocí odpovídajícího výběrového šetření poznali do jisté míry např. základní soubor „pacienti zubní kliniky Stuttgart“, budeme moci napříště — stále však ještě s náležitou obezřetností — aplikovat některé závěry šetření v Bordeaux také na Stuttgart.
Zásadně však platí, že *výběrový soubor (vzorek) vypovídá jen o tom základním souboru, z něhož byl odvozen:* dotazníková akce v Mnichově poskytuje informace o názorech a poměrech v Mnichově, nikoliv však v Hamburku a už vůbec ne v Paříži. Výběrové šetření o mzdách vyplacených v papírenském průmyslu Hessenska nevypovídá nic o mzdách horníků v Porúří apod. *Základní soubor*, „soubor vyššího řádu“, jak jej nazývá Anderson, *je onen základní soubor, v němž každý jednotlivý díl má stejnou naději dostat se do výběrového souboru* — „souboru nižšího řádu“.
Co však je v základních souborech nebo ve výběrových souborech *rozděleno?* *Znak*, který je předmětem šetření nebo,

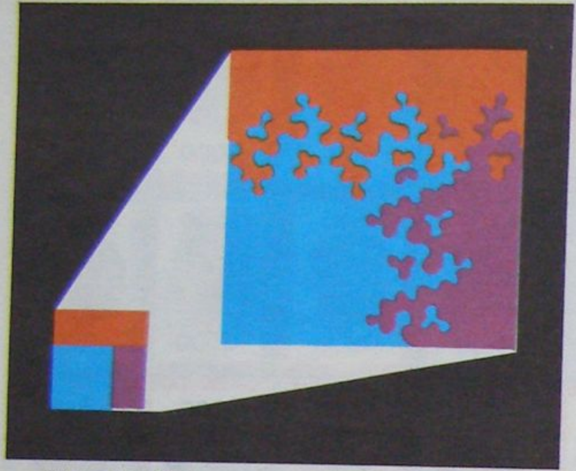


Statistické údaje jsou vždy „události", a proto skutečnosti z minulosti. Jejich projekce do budoucna je vždy zatížena nejistotou i při použití nejlepších postupů. Včerejší a dnešní čísla dovolují jen dohady a odhady čísel zítřejších, neumožňují však jejich přesný výpočet. Ze současné úmrtnosti se může pouze přibližně usuzovat o budoucí, z technických konstrukcí minulosti jen přibližně o konstrukcích v budoucnosti.

Téměř všechna statistická šetření se zásadně zakládají na pozorování nebo měření znaků v individuálních vyjádřeních: podle toho se zjišťují četnosti jednotlivých znaků. 25 osob dotázaných na svůj příjem se stane bezprostředně potom úplně neosobními „nositeli znaku" — nositeli, kteří se rozlišují jen více nebo méně četnými vyjádřeními znaku: dva z 25 vydělávají mezi 1500 a 2500 Kčs, pět z 25 vydělává méně než 800 Kčs — kdo jsou, jak se jmenují, zda mají příbuzné nebo dluhy, jsou nebo nejsou spokojeni, to je pro statistické zkoumání znaku „výše příjmu" prozatím bezvýznamné, což samozřejmě nevylučuje, aby v průběhu dalšího zkoumání souvislostí mezi příjmem a věkem, příjmem a školním vzděláním, příjmem a velikostí rodiny atd. se uvedené znaky zjišťovaly.

Znak: příjem





výběrový soubor neznámý základní soubor

Přesně známé je složení výběrového souboru. „Promítne-li" se však tento dílčí výsledek do neznámého základního souboru, přesnost se nutně ztrácí — zůstává jen odhad, ovšem v přesně propočitatelných mezích pravděpodobnosti a spolehlivosti.

přesněji řečeno, *četnost*, s níž se vyskytují určité charakteristiky znaku. Tak je třeba „rozdělen" důchod mezi příjemce důchodu. Někteří dostávají méně než 500 DM měsíčně, jiní mezi 500 a 550, další mezi 551 a 600 atd. až nad 1500 DM, kdy počet příjemců vyšších důchodů výrazně klesá (str. 26—27). *„Rozděleny"* jsou také četnosti, např. sled barev v ruletě. Velmi charakteristické rozdělení vykazují, jak ještě uvidíme, četná biologická měření, jako třeba tělesná výška nebo objem hrudníku: velmi vysokých a velmi malých lidí je málo, naproti tomu „průměrných" je velký počet. To je takzvané *„normální rozdělení"*, které má ve statistice významnou roli. Vyskytuje se také *„rozdělení" řídkých událostí*. Nikdo např. nemůže předvídat, kdo v příštím slosování bude mít 6 správných čísel, avšak docela dobře se dá předpovědět, kolik výherců přibližně bude.

Základní myšlenka, která je dnes základem značné části statistické činnosti, je tedy tato: má se propočítat (správněji: početním výkonem odhadnout), *jaké rozdělení určitého znaku je možno očekávat v základním souboru*, jestliže ve výběrovém vzorku jsem našel určité rozdělení.

Od průzkumu veřejného mínění ke kontrole výroby a od hospodářské prognózy k laboratornímu výzkumu se prostírá onen bezmála nekonečný prostor, v němž se na základě plánovaných experimentů, vzorků a dílčích poznatků usuzuje z *rozdělení* jednoho znaku *vzorku* na rozdělení stejného znaku v celém souboru.

Matematické metody, které vytvářejí početní souvislosti mezi vzorkem a celkem, jsou poněkud obtížné, ale nejdůležitější z nich mohou být srozumitelně vyjádřeny alespoň v hlavních rysech. Počtářské umění moderní statistiky, jak čtenář brzy uvidí, tvoří úplně svérázná matematika, spíše rafinovaná hra s pravděpodobnostmi a odhady než přesnost na několik desetinných míst, která se většinou s představou matematiky spojuje.

1.4 Pravděpodobnost a přesnost

Základem celé moderní statistiky je *počet pravděpodobnosti*, a to nikoliv jen z prostého a jasného důvodu, že zájem o údaje z minulosti vzniká především na základě zájmu o budoucnost; dílčí výsledky výběrových šetření mají především poskytnout pomůcku pro budoucí jednání.

Jde tedy vždy o odvození *závěrů z dílčích poznatků*, závěrů, které nesmějí nikdy být nesprávně chápány jako nesporné předpovědi, nýbrž jako předpovědi, které jsou vždy obklopeny pojmy jako „pravděpodobnost", „obor spolehlivosti", „maximální věrohodnost", „pravděpodobnost jistoty", „očekávaná hodnota", „rozptyl", „hypotéza" a „odhadovaná hodnota". Jednoznačně poznatelná může být nanejvýš minulost, nikdy budoucnost. „Poznání" (znalost) daného vzorku (výběrového souboru) však opravňuje k výpovědi o základním souboru, z něhož byl výběr proveden jen v přesně určených mezích.

K tomu, aby se ze vzorků a podobných dílčích výsledků mohly odvodit závěry, jejichž vypovídací schopnost lze popsat, je nezbytně nutné, aby výběr ze základního souboru byl proveden *nahodile*, a nikoliv na podkladě záměrné volby.

Jen v tomto případě je náhoda ponechána jaksi sama sobě, náhoda, kterou sice počet pravděpodobnosti dovede vymezit a podchytit, která je však falšována a křivena, jakmile se ji snažíme obejít subjektivně provedenou volbou (obr. na str. 31).

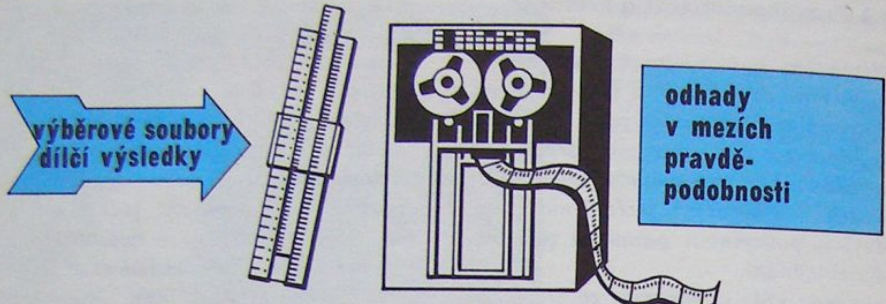
Z tohoto důvodu tvoří úvod do studia statistiky většinou studium počtu pravděpodobnosti a obráceně. Je velice obtížné vymezit přesnou hranici mezi počtem pravděpodobnosti a moderní statistikou. Neustále se objevují formulace, že statistika je koneckonců jen aplikovaný počet pravděpodobnosti. Je možné, že tato formulace není zcela správná, ale nikdo nepopře velmi úzkou souvislost mezi nimi.

Leonard Savage, jeden z nejvýznamnějších statistiků naší doby — a také, jak později ještě uvidíme, jedna z nejspornějších osobností — napsal k tomu ve svém díle „*The Foundations of Statistics*" (Základy statistiky): „Všeobecně se plně souhlasí s tím, že statistika se nějak zakládá na pravděpodobnosti. Co však se má rozumět slovem »pravděpodobnost« a jak je spojeno se statistikou, na to se názory tak různí a je tak malá možnost dorozumět se, jak sotva kdy od stavby babylónské věže."

V našem výkladu se ovšem nemůžeme podrobně zabývat těmito interními rozpory mezi velekněži vědecké statistiky. Základní teoretická vysvětlení pravděpodobnosti jsme omezili na minimum, které je nezbytně nutné k porozumění statistické činnosti. Nicméně považujeme za vhodné poukázat již zde na toto úzké sepětí. *Propočet pravděpodobnosti, kvantitativní výpovědi o hypotézách, odhady a domněnky*, to jsou dnes hlavní úkoly statistiky. Matematicky dokonalé a přesné jsou jen *metody*, *výsledkem* jsou však odhady a pravděpodobnosti.

Zdánlivá přesnost statistiky klame jen laika. Odborník ví, že statistika je *spíše umění odhadu a „nauka o odhadu" než počtářská technika*. Ernst Wagemann, prezident Říšského statistického úřadu ve 20. letech, napsal: „Počtářské umění je jen služkou své mnohem vzdělanější paní, statistického odhadu."

Britský statistik M. Y. Moroney končí svou knihu „*Facts from Figures*" (Skutečnost z čísel) dokonce těmito slovy: „Ve škole se nám stále nalévá do hlavy,

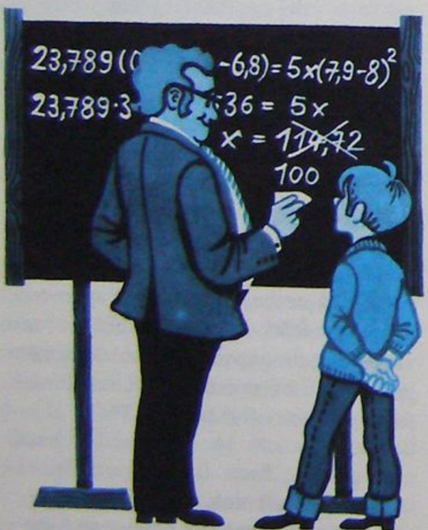


Ani použitím všemožných pomůcek nelze dosáhnout toho, aby z výsledků výběrových souborů bylo možno učinit nesporné závěry o celkovém souboru. Je však zcela možné přesně zjistit kvalitu odhadu, např. tímto způsobem: „S 95% jistotou je tento odhad plus nebo minus dvě procenta správný."

že aritmetika je exaktní věda, že všechny úkoly v učebnicích mají svá správná řešení. A pokud slyšíme něco o odhadech, pak vždy s poznámkou, že přitom jde o hrubou přibližnost a že lze samozřejmě nalézt přesné řešení... To je ale žalostná příprava pro budoucí život. Snad až na pokladníka v bance, který počítá peníze jiných, je jinak úplná přesnost aritmetiky k ničemu... Učitelé by prokázali svým žákům velmi cennou službu, kdyby je učili kritickému přístupu k aritmetice. Měli by je učit umění škrtnout 114,72 a místo toho napsat 100..."

Pro státovědnou statistiku 18. a 19. století měla pečlivá a přesná zjištění a úzkostlivě exaktní početní operace podstatný význam. Dnešní matematická statistika je naproti tomu převážně uměním určit, za *jak* nepřesné lze pokládat takové vypočtené údaje, jako je *pravděpodobný* výsledek, hypotéza či tvrzení.

Počátky tohoto poznání sahají daleko do minulosti. Quételet již před více než 100 lety varoval před přehnaným přesností. Laici, statističtí amatéři — a zvláště v minulých letech bohužel také statistikové z povolání — vždy znovu podléhali pokušení předstírat přesnost výsledků, která je popravdě prostě nedosažitelná. Oskar Anderson, jeden z prvních, kdo se v německé jazykové oblasti zasazovali o moderní statistiku, napsal ještě před necelými 40 lety na svou obranu toto: „Pro mne platí tyto dvě nezvratné zásady: 1. nemá praktický smysl chtít odvažovat fůru sena na chemicky přesných váhách; 2. není nic platné odhadnout vzdálenost mezi dvěma městy zhruba v tisíci krocích a pak k výsledku připočítat tloušťku městských hradeb v milimetrech." (Viz obr. na str. 32.)



„Učitelé by se měli naučit umění škrtnout 114,72 a místo toho napsat 100," říká anglický statistik Moroney a vyjadřuje tím, že v denním životě se mnohem častěji setkáváme s odhady a přibližnými hodnotami než s úplně přesnými údaji.

Uvedená druhá „zásada" není ani dnes všude uznávána. Úcta před přesností aritmetiky, kterou už Moroney tak odsoudil, stále ještě silně působí. A pokud se skutečně někdy uvádí hrubý odhad, pak první, komu padne do ruky, s ním nakládá jako se svátostí. Tak můžeme číst, že 1609 km vysoko ve vesmíru nebo také 1609 km daleko od pobřeží se stalo to nebo ono: *odhad „1000 námořních mil"* byl přepočten s úzkostlivou přesností. Z téhož důvodu se katastrofy stávají často přesně 32 km od Tokia nebo San Franciska — *to zase, když se něco stane ve vzdálenosti asi 20 mil.*

Statistika opovrhuje tímto druhem počtářského umění — a právem. Není pochyb o tom, že statistická výpověď může být uvedena přesně až na desetinná místa, ale skutečně dobré statistice nikdy nebude chybět odkaz na to, jak velká je pravděpodobnost pro přibližnou správnost údajů. Jestliže statistik nechce říci: „asi 1600", řekne: „s 95% pravděpodobností ne méně než 1583 a ne více než 1631".

Zásadně musí každý jednotlivec mít stejnou šanci být pojat do výběrového souboru; jinak nelze použít postupů počtu pravděpodobnosti. Nemohu se tedy dotazovat jen lidí s klobouky nebo lidí s fotoaparáty, chci-li se dovědět něco o souboru návštěvníků náměstí sv. Marka. A nesmím se také přirozeně dotazovat ve stejnou denní dobu. Předně: návštěvníci náměstí sv. Marka vybraní pro výběrový soubor jsou jen návštěvníci náměstí sv. Marka, a nikoliv „Benátčané" nebo „turisté z Itálie" anebo „návštěvníci italských pamětihodností" či „hoteloví hosté v Benátkách".

Bylo by nesprávné jen *statistice připisovat k tíži takové zdůrazňování pravděpodobnosti a poukazování na nepřesnosti.* Zde se jen otevřeně vyslovuje to, co nechceme přiznat v denním životě, v běžném hovoru a často i ve vědě — totiž že i naše „ano" a „ne", dokonce naše „docela jistě" a „určitě ne" jsou zatížena často nejistotou a snad i zcela nepatrnou pochybností. Pro zjednodušení a také proto, abychom se vyhnuli nenapravitelnému zmatku, nevšímáme si zpravidla náhod, překvapení a vzniklých nepravděpodobností. Matematická statistika to však činí viditelným, *vymezuje hranice pravděpodobnosti* a za to platí *zpochybněním přesnosti*.

Znovu však se musíme ptát spolu s Moroneyem: *k čemu je toto úsilí po přesnosti?* Pro správu železnic může být důležitá znalost počtu cestujících, kteří pojedou v příštím červenci z určitého města do Jugoslávie. K tomu ale plně postačuje znát počet cestujících v minulém roce zaokrouhleně na 1000 osob. Několik stovek více nebo méně už nehraje roli.

Ani vyčerpávající šetření není totiž bez vad. Nezřídka má větší vady než dobře a svědomitě provedené výběrové šetření. Jak problematickou se např. může stát při bližším pohledu i taková na první pohled prostá otázka, jako je počet cestujících v minulém létě z NSR do Itálie. Léto — co je to léto? Definujme: od začátku června do konce září. Co dál? Máme obejít všechny malé i velké cestovní kanceláře a sečíst záznamy cest do Itálie a prosit celníky na jižní hranici, aby se zeptali každého automobilisty na cíl jeho dovolené? Pak chybějí ještě jednotlivci cestující železnicí a ti, kteří jedou přes Francii. Ale cestující v osobních autech s paušální úhradou organizované cesty jsou již počítáni dvakrát. A co s těmi, kteří sice zaplatili let do Říma, ale odtud snad poletí dál do Káhiry nebo Madridu?

„Není nic platné, když vzdálenost mezi dvěma městy odhadneme a potom k výsledku připočteme tloušťku městských hradeb v milimetrech," řekl právem Oskar Anderson. Jestliže se k nějakému odhadu připočítá přesná míra, není tím celkový výsledek přesnější, nýbrž zůstává nadále odhadem.

Mají se mezi cestující do Itálie počítat také ti, kdo stráví dovolenou u Vrbského jezera v Korutanech a na tři dny si zajedou do Benátek?

Použije-li se údajů o přenocování a o přihláškách v cizině, budou čísla zase jiná, avšak nemusejí být lepší. Kolik majitelů soukromých pokojů raději své hosty nehlásí! Toho, kdo cestuje s obytným přívěsem, nelze skoro vůbec podchytit. Dotazníkovou akcí pomocí výběrového souboru několika tisíc občanů NSR se pravděpodobně dosáhne s mnohem menšími náklady správnějšího výsledku šetření než všemi uvedenými pokusy o vyčerpávající šetření.

Ještě většími zdroji chyb jsou údaje o tom, kolik DM nebo dolarů vydal v průměru turista ve Španělsku, Rakousku nebo Řecku. Všechny devizy totiž nenajdou cestu do Národní banky a dolary nemusejí vždy pocházet od Američanů a marky od Němců. Jestliže tedy čteme, že západoněmecký občan vydá v Itálii na osobu a den průměrně 3567 lir, je toto číslo především pouze odhadem a správně by mělo být vyjádřeno asi takto: „s 95% jistotou od 3500 do 3630 lir"; za druhé je toto číslo průměrem, který vzniká mechanickým rozdělením všech v Itálii vydaných DM na všechny německé cestující. Nicméně průměry mají ve statistice svůj velký — a často velmi špatně pochopený — význam. Věnujme jim proto pozornost hned na začátku.

Teorie pravděpodobnosti a statistika zacházejí s pojmy jako „jistota" a „určitost" pečlivěji než hovorová řeč. Pro budoucnost se nedají dělat žádné „absolutně jisté" předpovědi. Vědecký výhled do budoucnosti zná pojem „s nejvyšší pravděpodobností", nikoliv však „jistě".

2 Řád a zlořád průměrů

2.1 Aritmetický průměr

Průměr má ve statistice velmi významnou úlohu — a v představách, které o statistice existují, je jeho význam ještě daleko větší. Zatímco laik považuje různé „průměry" za velmi důležité, statistik zná všechny jejich záludnosti a nedostatky. Neodborník si říká: konečně mám koncentrovaný, informací nabitý údaj, který dovede vyjádřit najednou vše, co je jinak skryto v nekonečných a matoucích tabulkách. A tak se průměr až příliš lehce povyšuje na opravdu magické číslo.

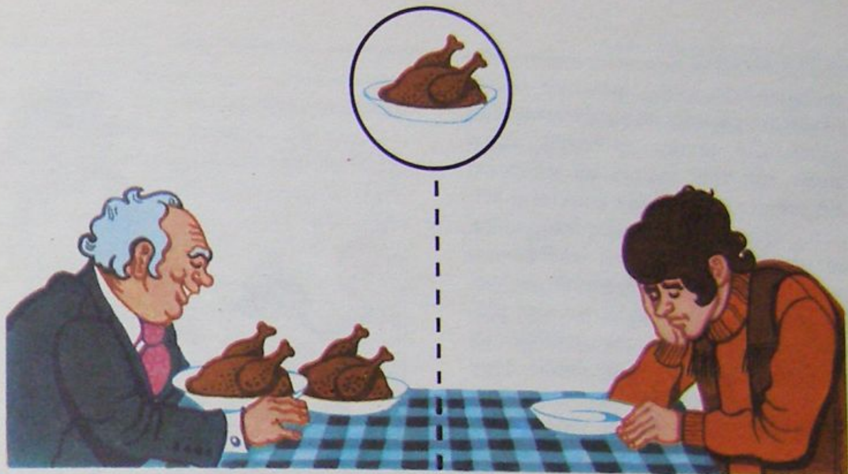
Průměr jako výsledek hodnocení nepřehledného množství čísel je pro statistika velmi důležitý vzhledem k tomu, že vnáší řád a přehled do tisíců nebo dokonce miliónů jednotlivých informací, a takové přehledné uspořádání je jedním z hlavních úkolů statistiky. Pro odborníka je však průměr jen prostředkem pro zjednodušené zobrazení.

Podívejme se však nejdříve podrobněji na onen průměr, který má zpravidla na mysli neodborník, když slyší toto slovo. K tomu je třeba sečíst větší nebo menší počet jednotlivých výsledků měření nebo jiných údajů a tento součet pak dělit počtem sčítanců. Toto „dělení" je etymologicky již opravdu „průměrem", neboť je rozdělováním a jako početní výsledek poskytuje rozdělení, tj. průměr: jestliže deset dělím pěti, tedy deset jednotek na pět stejných dílů, sestává každý díl ze dvou jednotek. Má-li např. 5 000 rodin dohromady měsíční příjem 10 miliónů DM, vydělávají „v průměru" 2000 DM. Nebo, jak praví jedno žertovné úsloví: „Sním-li dvě husy a ty žádnou, pak jsme v průměru snědli každý jednu."

Přesto má tento průměr magický účinek, kvůli němuž laik zapomene, že jde jen o jakýsi druh *orientační pomůcky v uspořádání veličin*, v žádném případě o prognózu pro jednotlivý případ. Očekávaná délka života třicetiletého muže, občana ČSSR, může v průměru činit ještě dalších 40,9 roku, čímž však ještě není řečeno, že všichni dnes třicetiletí muži (nebo i jen jejich značná část) zemřou krátce před dosažením 71. roku. Průměr obecně a aritmetický průměr zvláště slouží k tomu, aby bylo možno sjednotit značně rozdílné údaje. Třicetiletý jednotlivec se může zítra stát obětí dopravní nehody, stejně tak jako se v plné duševní svěžesti dožít roku 2020.

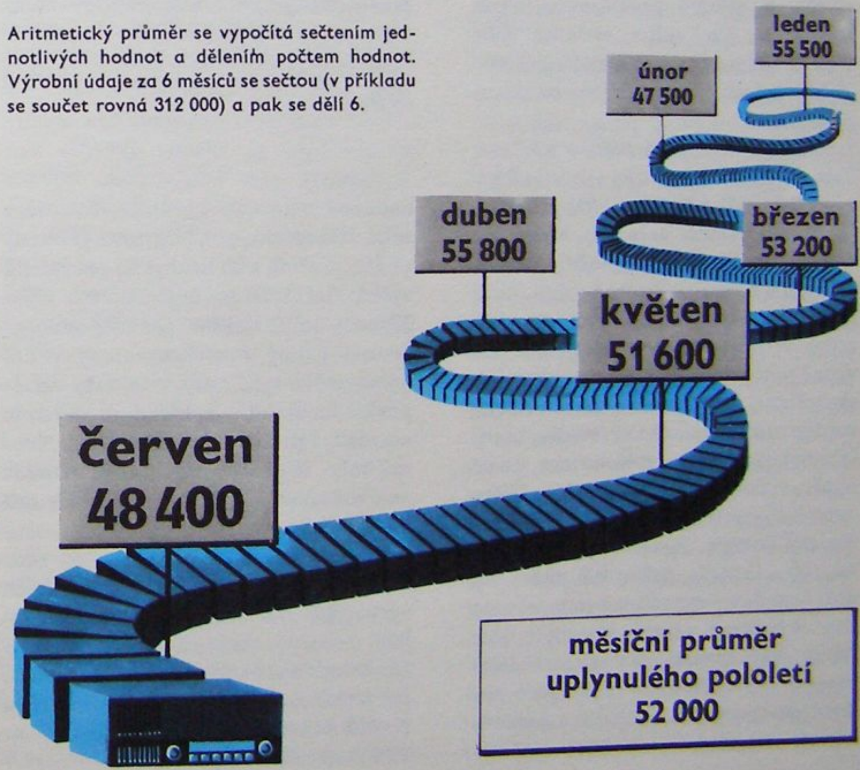
Jako součást „*obyvatelstva*", ve dvojím smyslu slova, jako občan i součást statistického základního souboru, je jeho osobní osud zahrnut do „*průměrné délky lidského života*", ať už proběhne jakkoliv. Průměrnou délku života lze stanovit pro „velké číslo", jímž je souhrnný počet všech třicetiletých, velmi přesně, protože se zjišťuje z ještě většího čísla a získaných zkušeností; pro jednotlivce má však relativně malý význam.

Individuální význam bude nepatrný, zejména tehdy, když „*rozptyl*" (tento pojem bude později podrobně vysvětlen — viz str. 47) je relativně velký, to znamená, když např. přibližně stejný počet



„Průměr" předstírá často rovnoměrnost nebo normu, která neexistuje. Když v „průměru" každý sní husu, je zcela možné, že někteří lidé snědí dvě, resp. i více, a jiní žádnou.

Aritmetický průměr se vypočítá sečtením jednotlivých hodnot a dělením počtem hodnot. Výrobní údaje za 6 měsíců se sečtou (v příkladu se součet rovná 312 000) a pak se dělí 6.



nynějších třicátníků nakonec zemře v šedesáti, v sedmdesáti nebo osmdesáti letech. Čím je naproti tomu rozptyl menší, tím větší vypovídací schopnost má průměr i pro každý konkrétní případ. Kdyby např. statistika dokazovala, že převážná většina lidí umírá mezi sedmdesátým a sedmdesátým druhým rokem věku, pak by bylo rozumné plánovat svůj život tak, jako kdybychom se nedožili více než sedmdesáti dvou let. Avšak to statistika vůbec netvrdí.

V nejjednodušším případě vypadá zjišťování „průměru" takto: Firma X vyrobila v lednu 55 500 rozhlasových přijímačů (nebo také zubních párátek), v únoru 47 500, v březnu 53 200, v dubnu 55 800, v květnu 51 600 a v červnu 48 400. K zjištění průměru je nutno těchto šest čísel sečíst, výsledek dělit šesti, a tak dostaneme měsíční průměr 52 000 jednotek vyrobených za první pololetí.

Prozkoumáme-li ještě jednou uvedená čísla výroby, zjistíme, že jen v květnu se výroba s 51 600 velmi blíží průměru (52 000), zatímco v lednu, březnu a dubnu bylo vyrobeno značně více a v únoru a červnu podstatně méně. To je zvláštnost všech matematických průměrů včetně průměru aritmetického: vypočítané číslo většinou neodpovídá skutečnosti. Tento fakt vynesl statistice mnoho neoprávněného výsměchu, který je mnohdy přímo vyvoláván tím, že se např. vypočítává průměr 2,2 dítěte nebo 0,36 oběti nehod. Anglický humoristický časopis „Punch" o tom jednou napsal: „Protože počet 2,2 dítěte na dospělou ženu vypadá v jistém ohledu značně podivně, navrhla kterási komise Jejího Veličenstva, aby se středním vrstvám poskytla finanční podpora pro dosažení zaokrouhlenějších a pohlednějších průměrů."

Aritmetický průměr nemá většinou žádný odraz ve skutečnosti. Když průměrná rodina má 2,2 dítěte, naštěstí to ještě neznamená takovou pochmurnou grotesku, jakou vidíme na obrázku.

Podobný aritmetický průměr však může také leccos skrývat. Zjistíme-li např. ve školní třídě s 30 hochy, že průměrná výška činí 1,67 m a průměrná váha 58,3 kg, je především pravděpodobné, že ani jediný z těchto mladých lidí nemá právě tyto míry a že tedy neodpovídá Quételetovu ideálnímu „homme moyen", „průměrnému člověku". Kromě toho třída se může navíc skládat ze dvou extrémních skupin: z více než tuctu těžkých vah a stejného počtu malých útlých hošíků. „Průměr" tuto skutečnost úplně skryje. V počátcích výroby pánské i dámské konfekce podlehli mnozí výrobci podobným klamným závěrům a vyráběli průměrné velikosti, po nichž se stěží nacházela poptávka prostě proto, že „ideální postavy" se vyskytují velmi zřídka.

Jiný příklad: hledáme-li ideální bydliště s příjemnou průměrnou roční teplotou a peníze nehrají žádnou roli, pak bychom v užší volbě mohli vybírat třeba mezi Pekingem, Milánem nebo Quitem, neboť všechna tato tři místa mají „*roční průměr*" kolem 12°. V Pekingu je však v zimě chladnější než ve Stockholmu a v létě takové vedro jako v horkých měsících v Rio de Janeiru. „Průměr" Pekingu vyplývá ze součtu a dělení extrémů. V Miláně se výkyvy udržují v hranicích obvyklých ve střední Evropě a roční průměr je tedy výsledkem mírně studených zim a mírně teplých letních období. Konečně v Quitu, hlavním městě Ekvádoru, se teplota vzduchu po celý rok téměř nemění. Příčinou je vysoká nadmořská výška, 3000 m nad mořem, a blízkost rovníku.

Můžeme si také představit velkou akciovou společnost, třeba „Akciovou společnost pro statistiku a výzkum čísel", jejíž akciový kapitál je rozdělen na 500 000 akcií. Akcie této společnosti vlastní 50 000 osob, a průměrně má tedy každý akcionář deset akcií. Při bližším zkoumání se však docela dobře může ukázat, že jediný majoritní akcionář má 320 000 podílů, další akcionář 30 102, sto osob vlastní po 1000 akcií a zbývajících 49 898 akcionářů má jen po jediném podílu ve své hubené peněžence. Jestliže však chci tuto skutečnost skrýt a přitom nemluvit o prostém průměru, mohu také ještě říci, že vlastnictví akcií je „neobyčejně široce rozptýleno" — přes 99 % akcionářů má méně než 500 akcií! (Může se také říci „méně než 50" nebo „méně než 800" — podle toho, která formulace se jeví účelnější.) Ve statistice General Motors Corporation bylo jednou uvedeno, že prý 78 % akcionářů vlastní méně než 50 akcií a jen 8 % více než 100 — v těchto 8 % by však snadno mohl být jen jediný akcionář s absolutní většinou všech akcií.

2.2 Modus a medián

Pošleme-li „průměrného akcionáře" naší akciové společnosti pro statistiku do říše fantazie, musíme místo něj hledat reálnějšího reprezentanta. Mohli bychom se např. zeptat, který je to ten „typický akcionář", vyskytující se nejčastěji. V našem případě lze tuto otázku zodpovědět snadno a jasně. Protože 49 898 z celkového počtu 50 000 akcionářů vlastní jen jediný podíl, je tento typ nejmenšího akcionáře nepochybně nejčastější a z tohoto hlediska „*typický*".

Statistika skutečně uznává tzv. „nejčetnější hodnotu" (nebo „modus") jako důležitou formu průměru, je ovšem jasné, že v příkladu, který jsme uvedli, nás zjištěná nejčetnější hodnota nemůže plně uspokojit, neboť majoritní akcionář se svými milióny se úplně ztratil. „Nejčetnější hodnota" neříká totiž vůbec nic o krajnostech, neprozrazuje, v kterém směru a jak jsou veliké. Naproti tomu prozradí někdy chybnou statistickou konstrukci: vyskytnou-li se např. v jednom rozdělení dvě různé „nejčetnější" (tedy přibližně stejně četné) hodnoty, byly zde bezpochyby omylem smíšeny dva různé celky. Takové „*bimodální*" neboli dvouvrcholové rozdělení se vyskytuje např. u tělesné výšky dospělého obyvatelstva jako celku — jedna nejčetnější hodnota pro ženy a jedna pro muže.

Zkusme to však nejdříve s dalším průměrem. K tomu si vybereme abiturientskou třídu určitého ročníku — řekněme 1965 — a dotážeme se jí na příjmové poměry. O problematice sběru informací budeme ještě podrobněji mluvit později; na tomto místě zatím naivně předpokládáme, že poskytnuté údaje jsou správné.

Počet dotázaných	vydělává vydělávají	DM ročně
1		100 000
1		60 000
2	po	40 000
3	po	25 000
5	po	20 000
1		18 000
12	po	13 000

Pomocí listu papíru a tužky nebo malé kancelářské počítačky rychle zjistíme aritmetický průměr — součet ročních příjmů činí 589 000 DM, průměr je 23 560 DM, tedy zase částka, kterou nikdo z dotazovaných ve skutečnosti nevydělává.

Nejčetnější hodnota je od ní velmi vzdálena; představuje ji na konci tabulky oněch dvanáct chudáků s 13 000 markami. Co teď?

Po krátkém uvažování se nabízí následující řešení: za průměrný příjem by bylo nejvhodnější považovat ten „prostřední" — prostřední podle pořadí v tabulce uspořádané podle výše příjmů. Jinými slovy: mezi uvedenými 25 osobami vyhledáme tu, která vydělává více než dvanáct nejchudších, avšak méně než dvanáct nejbohatších. To je, jak zjistíme z tabulky, muž s 18 000 DM.

Statistika již dávno objevila a používá tuto „*prostřední hodnotu*", „*centrální hodnotu*" nebo „*medián*". Má mimo jiné tu výhodu, že ji lze použít také v tzv. „*topologických stupnicích*" nebo řadách, u nichž výpočet aritmetického průměru nemá žádný smysl. Je možné zjistit „prostředního" žáka — před ním je 15 lepších, za ním 15 horších, zatímco aritmetický průměr ze součtu známek na vysvědčení by spíše mátl. Mimoto lze mediánu použít i v „*otevřených stupnicích*". Kdybychom třeba řekli, že dva z našich abiturientů vydělávali „přes 50 000 DM", pak by nebylo možno vypočítat aritmetický průměr, medián však zůstane stejný.

Z těchto důvodů medián nalézáme např. v rakouské statistice věku uzavírání sňatků, protože věk novomanželů sahá od patnáctého až k devadesátému (a někdy i vyššímu) roku věku a prostřední hodnota je velmi názorná:

Rodinný stav před uzavřením sňatku

Rok	Věk ženicha				Věk nevěsty			
	svobodný	ovdovělý	rozvedený	celkem	svobodná	ovdovělá	rozvedená	celkem
1937	28,3	50,2	38,6	28,9	25,5	42,6	34,0	25,8
1959	25,8	56,7	39,3	27,0	22,3	46,2	35,6	23,3
1969	24,7	58,3	36,5	25,6	21,6	48,8	30,9	22,2

Průměry mohou klamat, když se dva extrémy zdánlivě vyrovnávají. „Průměrem" z vyzáblých dlouhánů a malých tlouštíků je dobře vyvinutá a úměrná střední postava. V počátcích rozvoje textilní konfekce se někdy zhotovovaly oděvy pro „normální" velikosti, které se ve skutečnosti sotva vyskytují: „průměry" umožnily vznik falešné představy.

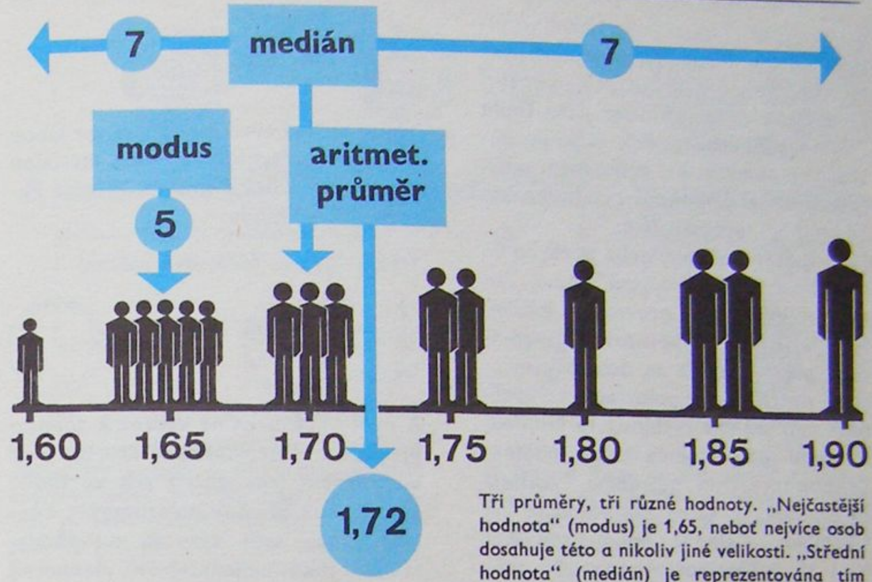
Jestliže se tedy uvádí, že průměrný sňatkový věk svobodných nevěst činil v roce 1969 21,6 roku, vidíme z toho, že polovině všech až dosud svobodných nevěst bylo méně a polovině více než 21,6 roku; zda v tom byly nebo nebyly i šedesátnice, to medián zamlčuje. Z těchto důvodů jej lze účelně použít tam, kde chybí omezení nahoru nebo dolů, protože se třeba shrnou všechny nízké a všechny vysoké hodnoty pod označení typu „více než...", „méně než...".

Přesto má i tato „prostřední hodnota" své nevýhody. Z příkladu našich akcionářů je vidět, že bylo nutno uvést jednoho z 49 898 nejmenších akcionářů jako „prostředního", ačkoliv všichni mají stejně málo. Stejně nevhodný se ukazuje v případě rodin s jedním dítětem. Dokonce ani v naší abiturientské třídě nepředstavuje nejlepší řešení. Co by se stalo, kdyby onen muž s 18 000 DM, náš „prostřední příjemce", vydělával místo toho jen 14 000 nebo naopak 19 000 DM? Medián by jej věrně provázel, zatímco aritmetický průměr by na tuto změnu reagoval jen nepatrně. Jestliže se však prostřední hodnota — medián — může tak značně změnit nahodilým výsledkem, musíme se na něj dívat s určitými pochybnostmi.

K tomu je však třeba přičinit dvě poznámky. Především: takové nebezpečí sotva hrozí v rozsáhlé oblasti popisné statistiky — klasické statistiky hromadných jevů; jen u malých výběrových souborů (vzorků) je nutno být nanejvýš opatrný. Čím rozsáhlejší je však číselný materiál získaný šetřením, tím menší

je nebezpečí, že průměr klame. Kdybychom měli v tabulce např. 2000 výdělečných osob místo 25, pak by jistě, i za předpokladu zcela podobného strukturálního rozdělení, nebyla mezera mezi 14 000 a 19 000, prostřední hodnota by mohla být zjištěna přesně a byla by uchráněna vlivu nahodilých výkyvů.
Za druhé, v povaze všech průměrů však je, že něco ponechávají nevyjádřeno. Vždyť průměry vytváříme s úmyslem zbavit se číselné změti jednotlivých údajů. Proto nemůžeme od průměrů požadovat, aby na první pohled vyjadřovaly podstatné a zároveň odpovídaly na detailní otázky. Každý průměr zakrývá a uhlazuje krajnosti a je jimi současně ovlivňován. Vzniknou-li pochybnosti, lze jej použít jen s doplňujícími údaji, např. o rozptylu. V případě pochybností je nutno se vždy ptát, o který z četných průměrů se vlastně jedná. (Nebo správněji: o kterou z mnoha středních hodnot. To proto, že statistika pojmu „průměr"

Průměrná roční teplota je jedním z „průměrů", které velmi lehce mýlí. Quito, Peking a Milán mají téměř stejný roční průměr, avšak zatímco v Quitu je skutečně po celý rok přibližně stejně teplo, je v Miláně v zimě citelně chladno a v létě velmi teplo, v Pekingu je třeskutá zima a žhavé léto. (Starý vtip praví: Stojí-li někdo jednou nohou na plotně a druhou na ledu, pak statistik říká, že v průměru je mu příjemně teplo.)



Tři průměry, tři různé hodnoty. „Nejčastější hodnota" (modus) je 1,65, neboť nejvíce osob dosahuje této a nikoliv jiné velikosti. „Střední hodnota" (medián) je reprezentována tím mužem, který má vedle sebe na jedné straně 7 větších a na druhé straně 7 menších. „Aritmetický průměr", prakticky nejdůležitější střední hodnota, je jen početní veličinou vzniklou ze součtu všech tělesných velikostí a dělený 15. Činí 1,723 m. Měřítko je pro názornost silně zkresleno, což je v tomto případě oprávněné. Později poznáme statistická zobrazení, v nichž taková zkreslení nebezpečně klamou.

moc ráda nepoužívá, a to tím méně, že v teorii množin se tímto pojmem vyjadřuje něco zcela jiného.)
Přehled dosud uvedených středních hodnot vypadá takto:
1. *aritmetický průměr* (jednotlivé hodnoty se sečtou a dělí se jejich počtem),
2. nejčetnější hodnota neboli *modus* (udává, který jednotlivý výsledek je zastoupen nejčastěji),
3. prostřední hodnota neboli *medián*, označovaný také jako centrální hodnota (má nad i pod sebou stejný počet jednotlivých měrných hodnot).
Při sudém počtu pozorování nebo měření není pochopitelně žádná reálná prostřední hodnota; definuje se pak jako aritmetický průměr z nejvyššího čísla dolní poloviny a nejnižšího čísla horní poloviny údajů uspořádaných podle velikosti.
Máme tedy tři střední hodnoty, a přesto nemůžeme žádnou z nich použít, chceme-li vypočítat něco tak všedního, jako je *průměrná rychlost jízdy* auta. Předpokládejme, že jedeme třicet km daleko a prvních deset km jedeme rychlostí 60 km/hod., dalších deset rychlostí 80 km/hod. a posledních deset rychlostí 100 km/hod. Jakou průměrnou rychlostí jsme jeli?
Aritmetický průměr odpovídá zdánlivě jednoznačně: 80 km/hod., protože $60 + 80 + 100 = 240 : 3 = 80$. Chceme-li u těchto tří čísel určit prostřední hodnotu — medián, uvedený výpočet se jen potvrzuje; nejčetnější hodnotu nemáme. Tedy „průměrně 80 km v hodině".

Z opatrnosti, kterou jsme získali již při zacházení se středními hodnotami, jsme se při odjezdu a v cíli podívali na hodinky a zjistili jsme, že doba jízdy činila 23 minut 30 sekund. Pak však, po důkladném zamyšlení a opětovném počítání, vychází průměrná rychlost sotva 77 km/hod., přesněji 76,6.
Jdou špatně hodinky, nebo je chyba ve statistice? Ani jedno, ani druhé. Jen zprůměrování bylo provedeno špatně. Jestliže začneme posuzovat extrémnější případ, ukáže se daleko jasněji, co se vlastně stalo: polovinu trati (patnáct km) jedeme rychlostí 15 km/hod. K tomu potřebujeme celou hodinu. Druhou polovinu pojedeme rychlostí 75 km/hod., k tomu potřebujeme jen 12 minut. Celková doba jízdy činí 60 + 12 = 72 minut. Aritmetický průměr nás ošálí výsledkem 15 + 75 = = 90 : 2 = 45 km/hod., při kterém bychom k projetí tratě potřebovali 40 minut. K zjištění správného výsledku doby jízdy je nutno pro oba stejné úseky trati vypočítat průměr: 60 minut + 12 minut = 72 minut : 2 = 36 minut pro každý z obou patnáctikilometrových úseků trati, což činí 25 km/hod.

2.3 Harmonický a geometrický průměr

S rychlostmi však můžeme pracovat přímo, jestliže použijeme tzv. *harmonického průměru*. Ukážeme si to na našem prvním příkladu. Nejdříve sečteme reciproké rychlosti, tedy jejich převrácené hodnoty: 1/60 + 1/80 + 1/100 = 47/1200. Toto číslo dosadíme do jmenovatele zlomku, do jehož čitatele jsme dosadili počet měření (3). Výsledek vypadá takto:

$$\frac{3}{\frac{47}{1200}} = \frac{3600}{47} = 76{,}6.$$

Tímto způsobem vypočítáme správnou průměrnou rychlost. Stejnou metodou provedeme ještě kontrolu našeho extrémního příkladu:

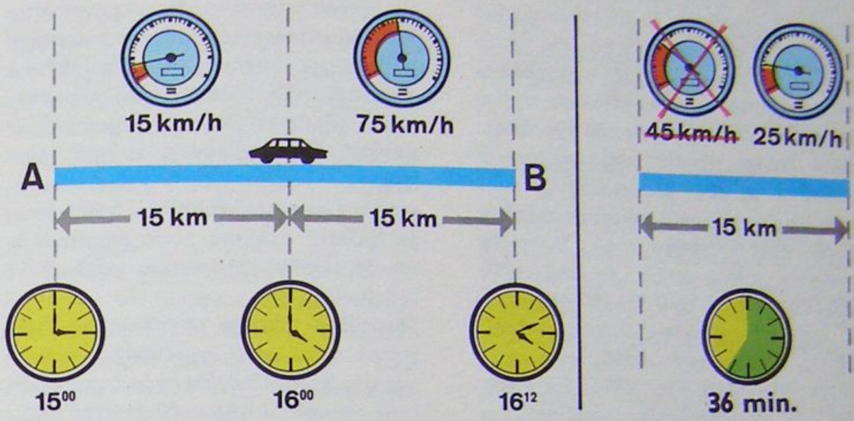
1/15 + 1/75 = 6/75; dvě měření

$$\frac{2}{\frac{6}{75}} = \frac{150}{6} = 25.$$

A nyní hned uveďme vzorec k tomuto propočtu. Matematické vzorce byly postrachem mnohého z nás ve škole. Jsou-li však užívány v rozumných dávkách, mají svůj význam a výhody, zvláště jako nejjednodušší všeobecné znázornění způsobu propočtu, kterým jsme se právě zabývali. Vzorec k výpočtu harmonického průměru vypadá takto:

$$\boxed{X_h = \frac{n}{\sum\left(\frac{1}{x}\right)}}$$

Vzorec vypadá hůře, než jaký je ve skutečnosti. Velké řecké písmeno sigma (Σ) je znaménkem součtu; všechno, co je uvedeno za ním, se má sečíst, v tomto případě tedy všechna $\frac{1}{x}$, tj. převrácené hodnoty výsledků měření (1/60, 1/80, 1/100). Počet měření je *n*. Symbol $\overline{X}_h$ znamená „harmonický průměr", někdy také $\bar{x}_H$ nebo zkráceně *H*.
Poznámka k $\overline{X}_h$: používat jen pro stejnoměrně rozdělené jízdní dráhy! Deset km à 60 km/hod. a šedesát km à 80 km/hod. dává tedy 7 měření po desíti kilometrech.



Volba nejvýstižnější střední hodnoty se většinou přenechává statistikovi, přesto však musí být dodržena určitá pravidla. Např. z rychlosti jízdy se nedá vypočítat správný aritmetický průměr: kdo ujede polovinu trati rychlostí 15 km/hod. a druhou polovinu 75 km/hod., nedosáhl „průměru" 45 km/hod., nýbrž jen 25 km/hod. Proto se musí vypočítat průměr celkové doby jízdy: 30 km ujetých za 72 minut dává v „průměru" 36 minut pro každou polovinu trati po 15 km. Elegantněji se počítá s pomocí „harmonického průměru".

A nyní k aritmetickému průměru:

$$\boxed{\overline{X}_a = \frac{\Sigma x}{n}}$$

V čitateli je tedy součet výsledků jednotlivých měření, ve jmenovateli počet měření. Vzpomeňme na firmu, která vyrobila v měsíci lednu až červnu 55,5 — 47,5 — 53,2 — 55,8 — 51,6 a 48,4 tisíce jednotek, což dává aritmetický průměr 52 (součet čísel výroby dělený šesti). $\bar{x}_a$ je aritmetický průměr: vystupuje také jako *A*, jako $\bar{a}$, a zejména jako $\bar{x}$ (vyslov x s pruhem), často také μ (vyslov mí).
Mezi $\overline{X}$, $\bar{x}$ a μ je však ten rozdíl, že řecká písmena μ a σ se zpravidla používají při popisu *základních souborů*, zatímco $\bar{x}$ a *s* se používají pro charakteristiku *výběrových souborů*.
To, co jsme vyjádřili jako

$$\bar{x}_a = \frac{\Sigma x}{n} \text{ nebo také } \bar{x} = \frac{\Sigma x}{n}$$

se někdy formuluje takto:

$$\bar{a} = \frac{1}{n}\sum_{i=1}^{n} a_i \text{ nebo } \mu = \frac{\sum_{i=1}^{N} a_i}{N}.$$

Princip je samozřejmě stejný. Čtenář však musí počítat s tím, že někdy narazí na jiné (i když smyslem stejné) vzorce, než jakých zde bylo použito. (*N* místo *n* je v jednom vzorci proto, že při μ jde o základní soubor a míry základního souboru jsou většinou označovány velkými písmeny — u výběrových souborů malými písmeny. Pokud jde o a_i, označuje se tím „cokoliv z mnoha *a*", např. kterákoliv z možných hodnot a_1, a_2, a_3, a_4 a a_5. Součet všech měření se tedy píše Σa.)
Vzorce existují samozřejmě i pro tzv. *polohou typické střední hodnoty*, tedy

medián (prostřední hodnota) a modus (nejčetnější hodnota). Ty nás zde již víc nezajímají. Ze sestavených a uspořádaných tabulek je lze většinou vyčíst „prostým okem", popřípadě vypočítat. Ještě jednu střední hodnotu nutno uvést, a to tzv. *geometrický průměr*, kterému se většinou přiděluje zvláštní úkol. Předpokládejme např., že rychle rostoucí město mělo v roce 1960 200 000 a v roce 1970 300 000 obyvatel. Z mezidobí nemáme žádné šetření. Přesto bychom rádi věděli, kolik obyvatel mělo město v roce 1965 (za předpokladu stejnoměrného organického růstu, tedy bez náhlého přírůstku třeba v roce 1964 nebo 1967).

Aritmetický průměr říká, že 250 000, což na první pohled vypadá zcela věrohodně. Uvažujeme-li však dále, vzniknou určité pochybnosti, zda přírůstek 100 000 obyvatel připadá stejnoměrně na obě pětiletí, nebo zda jde spíše o „populační explozi v malém" s typicky lavinovitým efektem, „*exponenciálním růstem*", s jeho vždy rychlejším zdvojnásobením. V daném případě se uchýlíme ke geometrickému průměru. Jeho vzorec zní:

$$\boxed{\bar{x}_g = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n}}$$

Někdy se označuje též $\bar{X}_g$ nebo G. Stanoví se tedy n-tá odmocnina ze součinu n výsledků měření. U našich dvou měření to máme jednoduché: $n = 2$, a druhou odmocninu ze součinu 200 000 a 300 000, tedy $\sqrt{60 \text{ miliard}} \approx 245\,000$ ($\approx$ znamená „přibližně rovné").

Ale jak se stanoví čtvrtá, devátá, dvacátá odmocnina? To není třeba. Místo nich se použije *logaritmů*.

Nemůžeme a také nechceme zde opakovat kurs počítání s logaritmy, musíme však zdůraznit, že pro statistiku jsou logaritmy elementárním způsobem počtů s podobným významem, jaký má pro statistika logaritmické pravítko, které představuje jeho nejjednodušší nástroj. Obojí vyhovuje způsobu statistické práce tím více, že provádí automaticky ona zaokrouhlení, která jsou, jak později uvidíme, zcela nevyhnutelná. Nyní však zůstaneme ještě chvíli u průměrů.

Především musíme poukázat na podstatný *rozdíl mezi vypočítanými a polohou typickými středními hodnotami*, které jsme naznačili již na počátku této kapitoly: matematické průměry (aritmetický, geometrický, harmonický) se vyskytují ve skutečném životě velmi zřídka, a to jen za pomoci fantazie; nemůže přece existovat „průměrná rodina" s 1,7 nebo 2,2 dítěte.

Podobně je tomu s dělením „na jednoho obyvatele". Jestliže rovnoměrně rozdělíme spotřebu doutníků, pak kouří i kojenci; rozdělíme-li výrobu uměleckých tisků, připadá z toho celá hromada na každou vesnickou chaloupku. Podobné průměry nelze brát doslova, ale jen takové jaké jsou: jako užitečné

Při přepočtu „na hlavu obyvatele" vznikají někdy velmi klamné průměry. Potom jsou dokonce i kojenci mimo jiné také spotřebiteli doutníků a whisky.

početní pomůcky a srovnávací čísla. Chceme-li však získat průměrné číslo vyskytující se také ve skutečnosti, musíme se uchýlit k polohou typickým průměrům, jako je prostřední nebo nejčetnější hodnota. Ale ani v tomto případě nemusíme být ušetřeni překvapení, nebudeme-li postupovat pečlivě a obezřetně.

Tak např. může medián, prostřední hodnota, při nerovnoměrném rozdělení statistického souboru zdůraznit nějaký extrém. Dejme tomu, že z uvedených jedenácti čísel uspořádaných podle velikosti v řadě 10, 10, 9, 9, 8, 3, 3, 3, 3, 2, 2 je mediánem číslo 3. Jestliže tuto řadu jen nepatrně změníme na 10, 10, 9, 9, 8, 8, 3, 3, 3, 2, 2, je prostřední hodnotou číslo 8. (Aritmetický průměr by byl v prvním případě $62/11 = 5{,}62$ a v druhém případě $67/11 = 6{,}07$. Nepatrnost změny by se v něm dobře projevila, ale 6 není obsaženo v žádné z obou řad.)

Nejčastější hodnota (modus) může být také výsledkem pouhé náhody a také nemusí nic vypovídat. Utvoříme-li opět řadu z jedenácti čísel, která vypadá např. takto: 10, 10, 8, 7, 6, 5, 4, 3, 2, 1, 0, pak „nejčetnější hodnota" je 10 (jen ta se vyskytuje dvakrát), ale aritmetický průměr je $56/11 = 5{,}1$. Máme-li však stovky nebo tisíce měřených hodnot, stěží dojde k takovému zkreslení; avšak statistika nepracuje vždy jen „s velkými čísly", ale také často s výběrovými soubory, které se tvoří jen z několika málo údajů. Medián a modus mají většinou pouze speciální úkoly, zatímco při možnosti volby se častěji použije aritmetického průměru.

Někdy se statistika označuje jako „učení o průměrných hodnotách". I když je to jedna z příliš vyhrocených formulací, nemůže být pochyb o tom, že střední hodnoty a jejich správná volba mají ve statistice úlohu, kterou lze jen stěží přecenit.

2.4 Statistika ve Zbohatlíkově

Pro ilustraci získaných poznatků o průměrech nechť poslouží historka, která se odehrává ve vybájené zemi, kde se počítá na tolary, avšak jinak jsou tamní poměry velmi podobné západoněmeckým.

V této zemi přišel jednou jeden muž do realitní kanceláře a řekl: „Chtěl bych pozemek na venkově, s lesem, loukami, ne příliš daleko od města, v pěkné krajině, za kterou by se člověk nemusel stydět. Samozřejmě že cenově výhodný." Zprostředkovatel chvíli usilovně přemýšlel, pak se rozzářil a pravil: „Zbohatlíkov na Malém Nadutci. To je přesně to, co hledáte. Měl bych tam ještě 20 hektarů nádherných pozemků."

Zájemce ohrnul nos: „Nedávno jsem tím hnízdem projížděl, vypadá to tam dost chudě." Zprostředkovatel zavrtěl hlavou: „To se musíte mýlit. Zde mám nejnovější údaje: průměrný příjem ve Zbohatlíkově činí 82 320 tolarů, to je průměr, který tak lehce nenajdete ani ve vilových předměstích."

Kupec slíbil, že celou věc ještě promyslí, a navštívil svého starého známého, ředitele banky, aby se zeptal, jestli snad neví něco o Zbohatlíkově. Ten prohrabal své záznamy a pravil: „Mám tu jakousi statistiku, kdepak je, aha tady je výtah. Nuže, roční příjem více než poloviny obyvatel je 29 000 tolarů a více. To je zatím všechno, co mám, ale nechám to zjistit podrobněji."

Kupec uvažoval: průměrný příjem činí 82 320 tolarů a více než polovina obyvatel má přes 29 000 tolarů; i když to není bezpodmínečně v rozporu, je to přesto podivné. Pak se rychle rozhodl, sedl do svého mercedesu a odjel směrem na Zbohatlíkov. Pět kilometrů před cílem se zastavil na okresním úřadě a vyžádal si informaci. Zde uslyšel: „Dosti chudé místo, průměrný příjem je kolem 29 000 tolarů. Chcete-li se dovědět něco bližšího, promluvte si s učitelem Počtářem. Ten je ve Zbohatlíkově doma a zcela nedávno se podrobně zabýval místní statistikou. Vidí sice všechno černě, je pesimista, ale je pečlivý a chytrá hlava."

Zájemce poděkoval a pak nejprve zavolal svého přítele, ředitele banky, zda nezískal další podklady. „Velmi zajímavý detail," řekl ředitel, „nejsilněji zastoupená příjmová kategorie je od 12 000 do 24 000 tolarů; nejčetnější příjem činí poměrně přesně 18 000 tolarů."

Soptě hněvem sedl kupec do svého vozu a odejel k učiteli Počtáři do Zbohatlíkova, který dělal stále stejně chudobný dojem. Na dotaz o příjmové struktuře v místě se učitel hořce zasmál a pravil: „Neutěšená, což je řečeno velmi mírně. Dvě třetiny rodin mají méně než 30 000 tolarů. Pokud jde o příjem na hlavu, nemá většina lidí ani 7500 tolarů ročně a 88 % má méně než 25 000. Jen náš pan vesnický milionář si žije skvěle v radovánkách na svém statku, obklopen několika tucty chudých otroků."

Na druhý den, když se zájemce o koupi poněkud uklidnil, zavolal konečně jednoho přáteli doporučeného úředníka Statistického úřadu a požádal jej zdvořile o autentický číselný materiál,

Ve Zbohatlíkově bydlí vedle převážně chudších lidí také milionář, a to v přepychové vile. Jak vysoký je průměrný příjem „obyvatel Zbohatlíkova"? Aritmetický průměr, geometrický průměr, nejčetnější hodnota, medián poskytují čtyři značně odlišné hodnoty, které jsou však všechny početně zcela „správné".

stručně mu vylíčiv vše, co se mu přihodilo. Úředník slíbil, že materiál obstará. Dva dny nato sděloval: „Okresní úřad vám udal správné číslo, 29 000 tolarů je medián. Také váš přítel z banky vás informoval správně, neboť nejčetnější hodnota je skutečně 18 000 tolarů. Pokud jde o učitele Počtáře, neuvedl jste sice přesné znění jeho vysvětlení, ale znám ho jako nejvýš svědomitého statistika-amatéra a nedovedu si představit, že by vám uvedl nesprávná čísla."

„Takže vychází, že vlastně lhal zprostředkovatel se svými 82 320 tolary?" zeptal se vyděšeně zájemce. Úředník však pravil mírně: „Aritmetický průměr není v tomto případě nejvhodnější, ale činí přes 82 320 tolarů. Situaci však lépe vystihuje geometrický průměr, který se pohybuje kolem 32 730 tolarů." Po těchto slovech zájemce bez okolků zavěsil, neboť teď věděl nad veškerou pochybnost, co už dávno tušil: statistikové jsou notoričtí lháři, pravdu překrucují. Po jeho zkušenostech mu toto mínění nemůžeme mít za zlé, vždyť i jinak docela chytří lidé nereagují někdy příliš rozumně, dostanou-li se do styku se statistickými údaji.

Byl to například onen britský ministr války, který se zeptal svého hospodářského statistika: „Co teď stojí vlna?" Ten začal vysvětlovat, že jsou velké výkyvy podle kvality a země původu, u koupě za hotové a koupě termínové a..., načež ho lord přerušil a pravil: „Dal jsem vám jednoduchou otázku a chci jednoduchou odpověď." Nuže, naše jednoduchá otázka zněla: „Jak vysoký je průměrný příjem ve Zbohatlíkově?" Chce-li někdo na to mít jednoduchou odpověď, musí riskovat, že nebude jednoznačná.

Chcete-li však případ Zbohatlíkov propočítávat sami ještě dál a přezkoumat uvedené údaje, uvádíme *„prvotní záznam"* příjmů 25 rodin a (v závorce) počet členů každé rodiny:

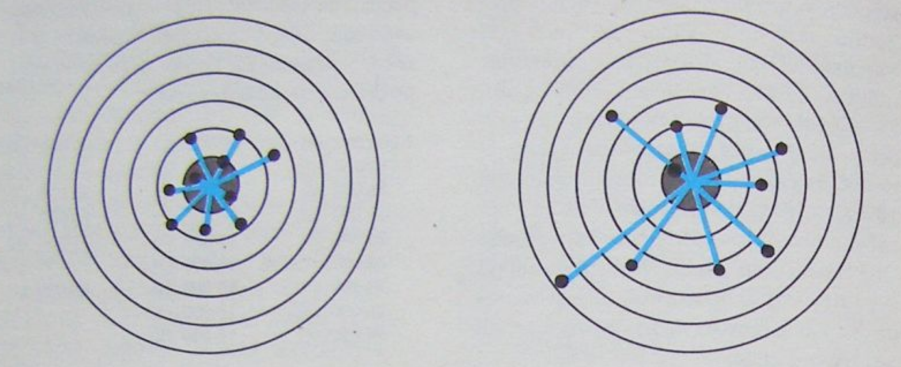
1 200 000 (3)	60 000 (1)	45 000 (2)
150 000 (5)	51 000 (3)	42 000 (2)
86 000 (4)	49 000 (4)	38 000 (4)
37 000 (3)	20 000 (7)	14 000 (1)
35 000 (5)	18 000 (3)	13 000 (4)
32 000 (3)	18 000 (8)	11 000 (1)
29 000 (3)	18 000 (4)	10 000 (2)
26 000 (4)	16 000 (3)	
24 000 (4)	16 000 (2)	

Hrubé zkreslení aritmetického průměru vzniklo tím, že rodina milionáře svou převahou potlačila ostatních 24 rodin, má větší příjmy než všichni ostatní dohromady. Jak vznikly další údaje, které tak popletly našeho zájemce o pozemky, lze poměrně snadno vypočítat za pomoci toho, co jsme se naučili v této kapitole. Poučení však zní: *také ve statistice mnoho záleží na správném kladení otázek.* A nejen to, patří sem i porozumění pro situaci. Neexistuje bezvadný průměr! Gerhard Mackenroth k tomu říká: „Neexistuje ani jedna správná střední hodnota, statistik však musí podle svého materiálu a problému mít možnost rozhodnout, kterou střední hodnotu zvolí."

2.5 Směrodatná odchylka a rozptyl

Ze všeho, co jsme se dosud dověděli o průměrech, bylo stále zřetelnější jedno: *střední hodnoty* (a především tak často používaný aritmetický průměr) *potřebují* k správnému zhodnocení svého významu a vypovídací schopnosti *ještě jeden rozměr*.

Takovým nejjednodušším doplněním je, že se uvede, *jak jsou extrémy od sebe vzdáleny*: „Nejnižší příjem je 13 000



Deset zásahů dobrého střelce leží blízko středu terče, „rozptylují" se jen málo. Zásahy špatného střelce leží většinou daleko od středu terče, mají velký rozptyl.

marek, nejvyšší 100 000, aritmetický průměr činí 23 560." Toto rozpětí „mezi 13 000 a 100 000" však neříká nic o tom, jak jsou jednotlivé hodnoty rozděleny: leží skoro všechny kolem 23 000 a jen dvě extrémní hodnoty nahoře a dole, anebo mají skoro všichni příjem 13 000 marek a jen dvě osoby 100 000 a pak je tedy aritmetický průměr stejně nesmyslný, jako byl v příkladu o dvou husách: jeden neměl žádnou a ten druhý snědl obě?

Statistika zná sice také „rozpětí", vymyslela však lepší míru, která ve spojení s průměrem umožňuje dobrý přehled o rozdělení jednotlivých hodnot shrnutých v průměru. Je to míra rozptylu — tzv. *směrodatná odchylka*. Vysvětluje se obvykle na příkladu střeleckého terče: je-li na terč vypáleno 50 výstřelů, „rozptýlí se" kolem středu. Leží-li zásahy těsně u středu a střílel-li tedy dobrý střelec, jsou vzdálenosti ke středu skoro všechny malé. Střílel-li však špatný střelec, budou zásahy od sebe značně vzdáleny.

Přeneseno na statistiku: jsou-li měřené údaje nebo sledované hodnoty jen velmi málo vzdáleny od aritmetického průměru, je rozptyl velmi malý. Když se však průměrují odlehlá čísla, je rozptyl veliký — tak jako je tomu na obrázku zásahů na levém a pravém terči, mezi nimiž uprostřed leží černý střed.

Výpočet „směrodatné odchylky" znázorníme nejdříve na docela jednoduchém příkladu. Aritmetický průměr čísel 7, 8 a 9 je 8, aritmetický průměr čísel 1, 10 a 13 je však také 8. Tuto zcela odlišnou situaci je možno vyjádřit naší mírou rozptylu. Začneme u první skupiny čísel (7, 8, 9).

Nejprve porovnáme každé číslo s vypočítaným aritmetickým průměrem a zjistíme rozdíl — přeneseno na střelecký terč: u každého zásahu změříme, jak je vzdálen od středu. Tak dostáváme:

7 je od 8 vzdálen —1
8 je od 8 vzdálen 0
9 je od 8 vzdálen 1
V druhém případě:
1 je od 8 vzdálen —7
10 je od 8 vzdálen 2
13 je od 8 vzdálen 5

Tím jsme určili záporné a kladné odchylky, které ovšem nemůžeme vzájemně sečítat. Na kladné je převedeme tak, že je umocníme na druhou, a tím zároveň dosáhneme toho, že extrémy ležící daleko od sebe vyniknou zvlášť zřetelně a připomínají nám: opatrně, zde dochází k velkému rozptylu.
Počítáme tedy dále:

První skupina

Odchylka	Druhá mocnina odchylky
— 1	1
0	0
1	1
	2

Druhá skupina

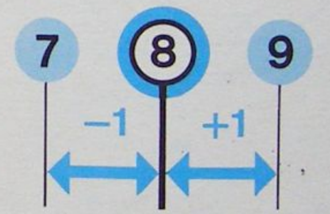
Odchylka	Druhá mocnina odchylky
— 7	49
2	4
5	25
	78

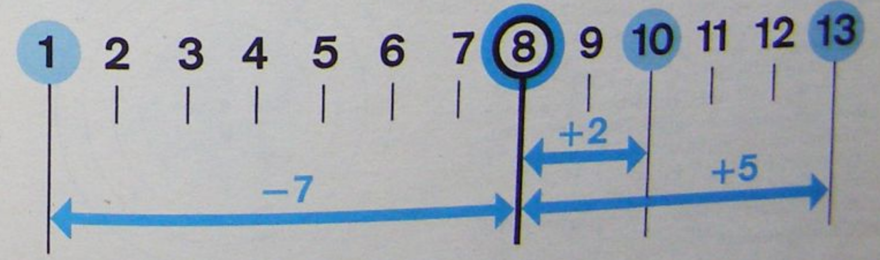
Rozdíl mezi 78 a 2 již skutečně bije do očí. Ale tato čísla ještě klamou: jsou tím větší, čím více jednotlivých měření sečítáme. Aby se přiměřeným způsobem vyjádřil počet měření, dělí se získaná čísla počtem měření — v našem případě jsme měli 3 hodnoty — dělíme tedy třemi. Tím získáme jen první míru rozptylu, tzv. „*rozptyl*" *nebo* „*varianci*" (var x nebo σ^2).

V době, která byla ještě ovlivňována Quételetovými pojmy „průměrného člověka" jako ideálního typu, označil americký astronom Simon Newcombe tuto míru rozptylu jako „evil", tj. „zlý duch". Dnes používáme raději neutrálního pojmu „*rozptyl*" nebo „*variance*", které v našem případě vypočítáme takto:

$$\text{var } x_1 = \frac{2}{3} = 0{,}67; \quad \text{var } x_2 = \frac{78}{3} = 26$$

Rozptyl se vine celou rozlehlou oblastí statistiky jako příslovečná červená nit. Znovu a znovu jsou umocňovány na druhou odchylky, vzdálenosti — od tohoto elementárního výchozího bodu až do vzdálených výšin analýzy rozptylu.

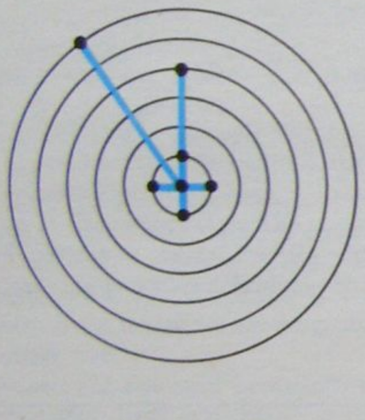


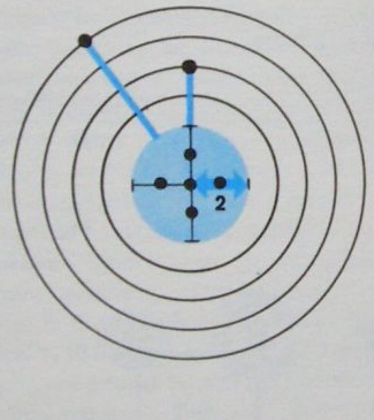


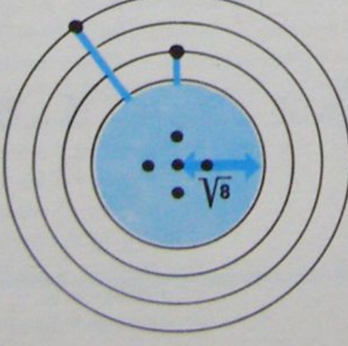
Aritmetický průměr neříká nic o rozptylu. Průměr ze 7, 8 a 9 činí 8 stejně jako průměr z 1, 10 a 13.

Ještě jednou upozorňujeme na důležitost této početní manipulace: za prvé, vylučujeme záporné hodnoty a za druhé, a to zejména, zdůrazňujeme větší odchylky. Máme-li např. absolutní odchylky 0, 0, 0 a 4, pomyslíme nejdříve na „průměrnou" vzdálenost 1. Jestliže 0, 0, 0, 4 umocníme a pak 16/4 odmocníme, dostaneme již 2, tedy míru podstatně charakterističtější, vzpomeneme-li si, že jen jeden ze série čtyř výstřelů byl vzdálen od cíle 4 jednotky. Přesně totéž provedeme s naším příkladem: k získání původního uspořádání veličin a ke konstatování „v průměru jsou zásahy tak a tak daleko od středu", provedeme odmocnění a tím dostaneme *směrodatnou odchylku* (s nebo σ, neboť se většinou označuje malým řeckým písmenem „sigma"). Měli jsme rozptyl $\sigma_1^2 = \frac{2}{3} = 0{,}67$ a $\sigma_2^2 = 26$. Nyní tedy dostaneme $\sigma_1 = \sqrt{0{,}67} = 0{,}82$; $\sigma_2 = \sqrt{26} = 5{,}1$.

Teď je zcela zřejmé, jak odlišný smysl má stejný průměr $\mu = 8$ v obou rozděleních. Směrodatná odchylka činí v prvním případě méně než 1, v druhém případě naproti tomu více než 5. Obě tato čísla je třeba číst v souvislosti se střední hodnotou 8; pak vyjadřují







Jak lze stanovit rozptyl s poměrně slušnou vypovídací schopností pro těchto 7 zásahů v terči? Jeden zásah je ve středu, čtyři ve vzdálenosti 1, jeden je vzdálen 4 a poslední 6 jednotek. Průměr $(0 + 1 + 1 + 1 + 1 + 4 + 6 = 14/7 = 2)$, znázorněný vpravo nahoře, neukazuje na příliš velkou odchylku. Vzdálenosti se tedy umocní, čímž se extrémy zvýrazní: $0 + 1 + 1 + 1 + 1 + 16 + 36 = 56/7 = 8$ — a tento výsledek se odmocní. Míra variance („směrodatná odchylka") činí v daném příkladě $\sqrt{8}$.

v prvním případě toto: většina čísel se odchyluje od střední hodnoty o méně než 1 v obou směrech, leží tedy mezi 7 a 9. Ve druhém případě to znamená: většina čísel se odchyluje od střední hodnoty 8 o více než 5 v obou směrech, leží tedy mezi 3 a 13.

Kvůli jednoduchosti jsme zvolili jen tři čísla. Není však obtížné představit si, že při padesáti, stu nebo desetitisíci čísel poskytuje charakteristika pomocí směrodatné odchylky dobrý obrázek, aniž se nejdříve musí zkoumat nebo alespoň přehlížet celé sloupce čísel. Statistik provedl výpočet a nyní už člověku znalému věci postačí jen dva údaje o průměrné hodnotě a o rozptylu, aby poznal, jaká je situace.

V našem druhém případě to znamená: jestliže průměr je 8 a směrodatná odchylka činí více než 5, je tento průměr zcela necharakteristický; s největší pravděpodobností byly seskupeny úplně nestejnorodé množiny údajů, jako např. v našem dřívějším příkladě s průměrnou velikostí žáků, když se zjišťoval průměr velkých a silných a malých a hubených žáků.

Je-li směrodatná odchylka svou absolutní hodnotou závislá na aritmetickém průměru, je možno vyjádřit ji také tak, že se vykazuje jako poměr mezi oběma mírami. Dosadí-li se totiž směrodatná odchylka do zlomku jako čitatel a průměr jako jmenovatel a násobí-li se 100, dostáváme tzv. *variační koeficient* (V); v našem případě

$$V_1 = \frac{0{,}82}{8} \cdot 100 = 10{,}2\ \%$$

$$V_2 = \frac{5{,}1}{8} \cdot 100 = 63{,}5\ \%$$

Podle velmi hrubého pravidla prozrazuje variační koeficient vyšší než 50 % silnou „nesourodost statistického souboru", a to v takové míře, že použití aritmetického průměru je už stěží oprávněné. Přesně to jsme zjistili u druhé skupiny našich čísel: 1, 10 a 13 dávají jen s použitím „násilí" průměrnou hodnotu $\mu = 8$.

Průměry jsou tu dále také k tomu, aby zjednodušeně zobrazily rozdílné údaje, měření a pozorování, avšak nějaké minimum podobnosti mezi těmito hodnotami je přece jen žádoucí. Jestliže se zjišťuje tělesná výška 20 žen kmene Pygmejů z pralesů Konga a 20 válečníků kmene Masajů ze stepí Tanzánie a na tomto základě se dospěje k závěru, že „Afričané" mají průměrnou výšku 1,65 metru, pak směrodatná odchylka a variační koeficient ihned prozradí, že „obyvatelstvo" bylo vytvořeno z nesourodých prvků (a vzpomeňme si: měli bychom zase dvě typické hodnoty — dominanty, jak tomu bylo u tělesné výšky obyvatelstva na str. 36).

Stejným způsobem prozradí velmi brzy míra rozptylu, jak je to s průměrnými teplotami v Pekingu, Miláně a Quitu, o nichž jsme mluvili na začátku této kapitoly. Stačilo by, aby místo ročního průměru bylo za základ výpočtu použito měsíčních průměrů uvedených měst a prozkoumalo se, jak se „rozptylují" od ročního průměru. Pak bychom s určitostí dostali v Pekingu velmi vysoký variační koeficient, v Quitu naproti tomu jen velmi malý rozptyl; tam je totiž soubor (tj. souhrn měsíců) velmi sourodý — průměrné teploty v květnu se příliš neliší od teplot v červenci, říjnu nebo únoru. A mohlo by se jít tak daleko — ačkoliv je to jen neužitečná počtářská hra — že by se vypočítal rozptyl oproti ročnímu průměru z denních 365 průměrů (které zase mohou být průměrem tří měření za den!).

Výsledek by byl však v podstatě stejný: nepatrná směrodatná odchylka pro Quito, vyšší pro Milán, velmi vysoká pro Peking.

Nyní bychom ještě mohli zkusit vypočítat směrodatnou odchylku v příjmech občanů Zbohatlíkova, ale to by nemělo — jak už víme — moc smyslu. Již jen krátká úvaha ukazuje, že příjem milionáře se částkou vyšší než 1,1 miliónu odchyluje od průměrné hodnoty $\bar{x} = 82\,300$; to po umocnění na druhou dává skoro 1200 miliard. Rozptyl činí asi 50 miliard (1200 miliard děleno 25 představuje již samo o sobě tuto částku), směrodatná odchylka je tedy asi 230 000 tolarů. Protože však aritmetický průměr činí pouze 82 000, je směrodatná odchylka přibližně třikrát tak velká, tzn. že variační koeficient činí téměř 300 %. Pošetilost takové průměrné hodnoty je z toho jistě zcela zřejmá.

Lepší vyhlídky naproti tomu má provedení potřebných početních úkonů s našimi 25 členy abiturientského ročníku 1965, ačkoliv i v tomto případě víme, že jeden zvlášť vysoký příjem způsobuje, že rozdělení je dosti nerovnoměrné. Rozpětí, které se ostatně ve statistických propočtech sotva kdy objevuje, sahá od 13 000 do 100 000 při 25 měřeních.

Pokusme se o štěstí:

Příjem (v tisících)	Četnost	Odchylka od průměru ($\bar{x} = 23{,}5$)	Druhá mocnina odchylky (cca)
100	1	76,5	5850 · 1 = 5850
60	1	36,5	1330 · 1 = 1330
40	2	16,5	270 · 2 = 540
25	3	1,5	2 · 3 = 6
20	5	— 3,5	12 · 5 = 60
18	1	— 5,5	30 · 1 = 30
13	12	—10,5	110 · 12 = 1320
	25		9136

Rozptyl (var x) činí $\frac{9136}{25} = 36{,}6$ a tím je směrodatná odchylka poněkud vyšší než 6 (6,04); a konečně variační koeficient $V = 100 \cdot \frac{6}{23{,}5} = 25\,\%$. Tato vysoká hodnota variačního koeficientu jasně říká, že rozptyl je veliký. Kdyby odpadl špičkový příjemce, byla by situace rázem jiná.

Odpadnutí extrémní hodnoty vede vždy k podstatnému zmenšení rozptylu a směrodatné odchylky.

Ještě něco je třeba dodat. V dosavadním výkladu jsme už častěji označili průměrnou hodnotu $\bar{x}$, směrodatnou odchylku σ místo s. Je to poněkud nedůsledné, nikoliv však zcela neoprávněné. Při kontrole jakosti se totiž mezi jiným mluví zcela všeobecně o *„pravidlu dvou sigma“*, i když přitom vůbec nejde o rozdělení souboru. Řecké μ se pro střední hodnotu naproti tomu většinou velmi přesně píše jen tehdy, když se jedná o aritmetický průměr základního souboru, zatímco pro výběrový soubor se vždy používá $\bar{x}$.

V případě směrodatné odchylky výběrového souboru je situace bohužel poněkud složitější. Znaku σ pro základní soubor odpovídá sice ve výběrovém souboru znak s, ale toto s se většinou nevypočítává stejným způsobem jako

naše σ, nýbrž tak, že se celkový rozptyl nedělí celkovým počtem měření, ale počtem měření zmenšeným o 1 a teprve potom se odmocňuje. To znamená, že místo

$$\sigma^2 = \frac{\Sigma (x - \mu)^2}{n}$$

bychom psali:

$$s^2 = \frac{\Sigma (x - \bar{x})^2}{n - 1}$$

Mnohem později se dozvíme, proč tomu tak je. Prozatím budeme naše výpočty rozptylu provádět nadále podle jednoduchého schématu stanoveného pro σ, i když někdy budeme mluvit o s. Důvod rozdílného propočtu je v podstatě ten, že s výběrového souboru se tím poněkud zvětšuje a poskytuje tak správnější měřítko skutečného rozptylu základního souboru. Prozatím jsme však ještě nedošli k teorii výběrových souborů.

2.6 Třídění

A nyní se musíme k něčemu přiznat: ve skutečnosti to není tak, že pět z našich abiturientů vydělává přesně 20 000 DM ročně, další tři 25 000 a ostatní rovněž tak pěkně zaokrouhlené částky, které nám ulehčovaly propočet. Dopustili jsme se tedy určitého zjednodušení, za které se dodatečně omlouváme a které můžeme hned zdůvodnit tím, že nás přivedlo k důležitým pojmům, jako jsou prvotní záznam a rozdělení četnosti.

Prvotní záznam je sestava všech zjištěných údajů neuspořádaně zaznamenaných nebo — pro přehlednost — seřazených podle velikosti. Uvedených 25 osob při zjišťování svých příjmových průměrů uvedlo následující údaje (seřazené od nejnižšího k nejvyššímu příjmu):

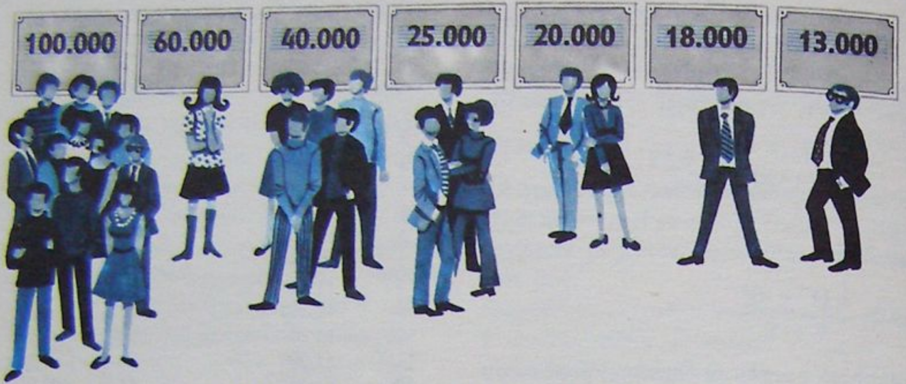
Osoba	DM	Osoba	DM
1	11 850	14	19 100
2	12 200	15	19 300
3	12 342	16	20 000
4	12 700	17	20 300
5	12 750	18	22 000
6	12 917	19	24 300
7	13 000	20	24 950
8	13 000	21	26 000
9	13 380	22	38 500
10	13 500	23	41 000
11	14 000	24	61 500
12	14 439	25	105 000
13	18 000		

Tato čísla jsme v předchozím textu *seřadili do charakteristických skupin:* osoby 1 až 12 jsme zařadili do skupiny s příjmem „průměrně 13 000 DM“, osobu 13 s příjmem 18 000 DM, osoby 14 až 18 „průměrně 20 000“, osoby 19—21 „průměrně 25 000“, osoby 22 až 23 „průměrně 40 000“. U obou špičkových příjemců jsme částky zaokrouhlili na 60 000 a na 100 000.

Byl tento postup nesprávný?

Zásadně nikoliv. Musíme si však být vědomi toho, že jsme sem zavedli prvek libovůle. Vždyť jsme stejně dobře mohli osoby 13 až 21 zařadit jako skupinu s „průměrným příjmem 21 000“ nebo obráceně: osoby 1 až 6 zařadit jako skupinu s příjmem „méně než 13 000“. V tomto případě jsme postupovali podle volného uvážení, které je oprávněné, pokud odpovídá danému číselnému materiálu.

Zařazování do skupin je totiž jedním z nejvýznamnějších statistických postupů, protože slouží k tomu, aby se nepřehledné množství jednotlivých údajů přehledně rozčlenilo. *Členění základního souboru do skupin (dílčích souborů)* podle určitých společných znaků je hlavním úkolem popisné statistiky.



Meze tříd ve statistice musí být přesně označeny. Zde stojí muž s 20 000 DM bezmocně mezi dvěma skupinami a nepatří k žádné. Kdyby vlevo znamenalo „do dvaceti tisíc", bylo by jeho místo vlevo, kdyby vpravo znamenalo „dvacet tisíc a více", musel by se zařadit vpravo.

Chceme-li uvedených 25 příjmů zařadit do „tříd", musíme nejdříve stanovit *stejnoměrné meze tříd*. Zkusme to třeba s intervalem 20 000:

Příjmy od	Osoby
0 — 20 000 DM	16
20 001 — 40 000 DM	6
40 001 — 60 000 DM	1
60 001 — 80 000 DM	1
80 001 — 100 000 DM	0
100 001 — 120 000 DM	1

Kdybychom však meze posunuli jen o jednu jedinou marku, nastala by dosti podstatná změna vzhledem k tomu, že první dvě třídy by pak vypadaly následovně:

Příjmy od	Osoby
0 — 19 999 DM	15
20 000 — 39 999 DM	7

Zůstaňme však u původně zvolené tabulky a určeme pro další výpočty *střed třídy*, to znamená:

Interval	Střed třídy	Obsazení (počet osob)
0 — 20 000	10 000	16
20 001 — 40 000	30 000	6
40 001 — 60 000	50 000	1
60 001 — 80 000	70 000	1
80 001 — 100 000	90 000	0
100 001 — 120 000	110 000	1

Tento postup spočívá ovšem už na předpokladu, že uvnitř intervalu je rozdělení přibližně symetrické, což však v našem případě není, protože příjmy uvedené v prvním intervalu (0—20 000) jsou vyšší než 10 000. V zájmu jednoduchosti však již na tom nebudeme nic měnit.

Další početní postup zjištění průměrného výdělku a směrodatné odchylky probíhá následujícím, již známým způsobem:

Střed třídy	Počet	Střed třídy krát počet
10 000	16	160 000
30 000	6	180 000
50 000	1	50 000
70 000	1	70 000
90 000	0	0
110 000	1	110 000
6 středů tříd	25	570 000

$$\text{Aritmetický průměr } \bar{x}: \frac{570\,000}{25} = 22\,800$$

Tento „průměrný příjem" se sice podobá tomu, který jsme již zjistili se zaokrouhlenými čísly (činil 23 560), ale je přece jen o něco nižší. Postupujeme dále v propočtu, abychom zjistili směrodatnou odchylku a variační koeficient: bychom však výslovně dokázali oprávněnost svého postupu. Při výpočtu rozptylu jsme druhou mocninu odchylky vždy „*vážili*" počtem osob příslušné třídy, tedy např. šestnáctkrát 164, ale pouze jedenkrát 1380. Tímto způsobem jsme získali přesný vážený aritmetický průměr.

Předpokládejme nyní toto: našich 25 abiturientů rozložíme do dvou velikých kategorií: a) zaměstnanci, b) samostatně činní. Prvních dvacet osob z našeho prvotního seznamu jsou zaměstnanci, zbývajících pět je samostatně činných. Jestliže vypočítáme průměrný příjem zaměstnanců, dostaneme $\bar{x}_1 = 16\,200$ a u zbývajících $\bar{x}_2 = 54\,400$. Následující argumentace by byla dosti nebezpečná: jestliže průměrný výdělek zaměstnanců činí 16 200 a samostatně činných 54 400, je aritmetický průměr těchto dvou druhů příjmů $(16\,200 + 54\,400) : 2 = 35\,300$. To by však byl

Střed třídy (v tisících)	Počet	Odchylka od průměru ($\bar{x} = 22{,}8$)	Druhá mocnina odchylky
10	16	—12,8	16 · 164 = 2624
30	6	+ 7,2	6 · 52 = 312
50	1	+37,2	1 · 1380 = 1380
70	1	+57,2	1 · 3260 = 3260
90	0	+77,2	0 · 5950 = 0
110	1	+97,2	1 · 9450 = 9450
	25		17026

$$\text{Rozptyl var } x = \frac{17\,026}{25} = 68{,}1; \quad V = \sqrt{68{,}1} = 8{,}25; \quad \sigma = 100 \,.\, \frac{8{,}25}{22{,}8} = 36\,\%$$

I v tomto případě jsou čísla podobná, ale nejsou stejná jako ta, která jsme získali při původním předpokladu. Každá nová početní metoda (jiné intervaly, jiné středy tříd) by vedla k jiným, *nicméně do jisté míry podobným hodnotám*.

Ještě jsme dlužni další upozornění: opět jsme postupovali zcela správně, aniž chybný, nevážený průměr, s nímž se podrobněji seznámíme později na podkladě příkladu ze začátků statistických indexů (viz str. 113 an.). Střední hodnoty se nemohou sčítat ani opakovaně průměrovat, pokud nejsou k dispozici takové dodatečné informace, které umožňují vážení.

Mnichovské průmyslové závody podle velikostních tříd

Závody s ... zaměstnanců	Závody		Zaměstnanci		Obrat (1000 DM)	
	abs.	%	abs.	%	abs.	%
1—9	1107	56,2	3 483	1,8	14 626	1,4
10—19	205	10,4	2 903	1,5	25 936	2,5
20—49	272	13,8	8 594	4,6	42 685	4,0
50—99	144	7,3	10 189	5,4	72 167	6,8
100—199	111	5,6	15 594	8,3	67 527	6,4
200—499	75	3,8	23 603	12,5	180 663	17,1
500—999	24	1,2	17 419	9,2	136 924	13,0
1000 a více	34	1,7	107 021	56,7	514 283	48,8
Celkem	1972	100	188 806	100	1 054 811	100

Jestliže však vím, že se při propočtu použilo za $\bar{x}_1$ 20 (nebo docela obecně m) jednotlivých hodnot a za $\bar{x}_2$ 5 (nebo n) jednotlivých hodnot, je možno vypočítat vážený průměr obou skupin.

$$\bar{x} = \frac{20\,\bar{x}_1 + 5\,\bar{x}_2}{25}$$

$$= \frac{20 \,.\, 16\,200 + 5 \,.\, 54\,400}{25}$$

$$= \frac{324\,000 + 272\,000}{25} = 23\,840$$

nebo obecně: $\bar{x} = \dfrac{m\bar{x}_1 + n\bar{x}_2}{m + n}$

Těchto 23 840 marek jako průměrný příjem („vážený aritmetický průměr") je mimochodem nejpřesnější z dosud stanovených hodnot, protože byl vypočítán přímo z prvotního záznamu, bez zkreslení vzniklého zaokrouhlováním a třídními intervaly. V souvislosti s tím vzniká otázka, proč se tímto způsobem nepostupuje hned od počátku. Odpověď je jasná: jestliže se jedná — jako v našem jednoduchém případě — o základní soubor v počtu 25 osob (nebo obecně nositelů znaků), nepřijde žádný rozumný člověk na nápad, aby použil statistické metody tohoto druhu. Je však také zřejmé, že je podstatně snazší a přitom dostatečně přesné v případě 250 nebo 2500 či dokonce 25 000 jednotlivých údajů nepracovat s prvotním záznamem, ale použít rozdělení četností v omezeném počtu účelně členěných tříd, které mohou být dále členěny, což potom vypadá tak, jak ukazuje tabulka v záhlaví.

V tomto případě se ukázalo také užitečné zvolit různě veliké třídní intervaly: kdyby se kategorie „500 až 999 zaměstnanců" rozdělila na desítkové intervaly, byla by tabulka nekonečná a nepřehledná, a proto by nevyhovovala hlavnímu účelu celé deskriptivní statistiky, tj. zpracovat a přehledně připravit velký počet údajů. Ve všech takových případech, kdy skupiny v tabulce jsou nahoru nebo dolů otevřené, hrozí nebezpečí úmyslného nebo neúmyslného potlačení informací. Vraťme se ještě jednou k našim abiturientům a rozdělme si je takto:

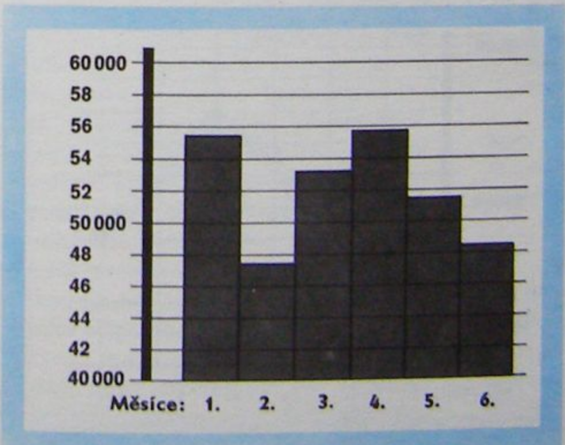
Příjem		
pod	15 000 DM	12
15 001 — 20 000 DM		4
20 001 — 25 000 DM		4
25 001 — 30 000 DM		1
přes	30 000 DM	4

Tato sestava je zcela neuspokojivá. Chybí totiž jakýkoliv náznak, jak hluboko pod hranicí 15 000 DM leží dvanáct nejnižších příjmů (teoreticky by to mohlo být i jen několik set marek) a stejně tak není jasné, zda 4 pánové ze skupiny „přes 30 000 marek" jsou těsně u této hranice anebo zda tři z nich nemají miliónové příjmy.

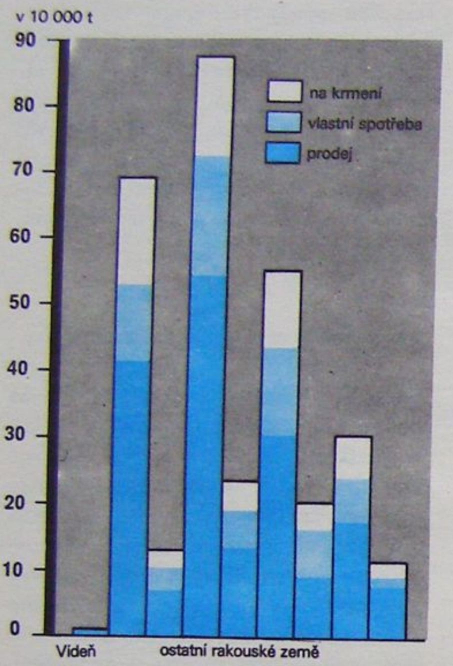
Výše uvedená podniková statistika uniká tomuto nebezpečí především tím, že se uvádí celkový počet zaměstnanců, takže z údajů může být přibližně odhadnuta kategorie „1000 a více". I bez statistických znalostí je hned vidět, že uvedených 34 podniků vzpomenuté kategorie musí mít nejméně 34 000 zaměstnanců; k rozdělení tedy zbývá ještě 73 021. To znamená: s absolutní jistotou nemůže mít jediný podnik více než 74 021 zaměstnanců (pouze za groteskního předpokladu, že zbývajících 33 podniků má přesně po 1000 zaměstnanců), s největší pravděpodobností nebude mít ani jeden podnik byť i jen 50 000 zaměstnanců, ačkoliv by bylo docela možné, že je v této kategorii zahrnuto 33 podniků s počtem 1000—2000 zaměstnanců vedle jednoho mamutího podniku.

2.7 Histogram a polygon četnosti

Četnosti, rozdělení četností — tyto pojmy se táhnou jako červená nit celou statistikou, i když nejsou vždy výslovně uvedeny. Lze je vyjádřit v tabulkách, mohou však být také vyjádřeny graficky. Výhodami i nevýhodami statistických zobrazení se budeme později zabývat ještě podrobněji. Na tomto místě považujeme za nutné představit dvě elementární možnosti zobrazení: histogram a polygon četnosti. Obojí zní hůř, než jaké to ve skutečnosti je. Slovník vycházející z medicíny definuje *histogram* jako *„drobnohledné zobrazení*



Histogram výrobních údajů z našeho dřívějšího příkladu.

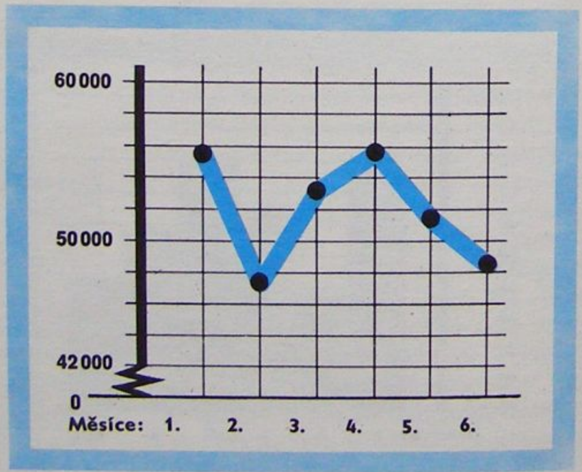


Sloupkový diagram: Použití mléka (v 10 000 t) vyrobeného v rakouských spolkových zemích.

tkaniny obvazu"; pro naše potřeby by se to dalo nejlépe přeložit slovy „*strukturální obraz*". V nejjednodušším případě dané zobrazení vypadá tak, že různě velké bloky (dvojrozměrně zobrazené, tedy vlastně plochy) jsou zobrazeny vedle sebe tak, aby vypovídaly o vzájemné velikosti nebo četnosti. Histogram půlroční výroby podniku na rozhlasové přijímače, která byla uvedena na počátku tohoto oddílu, ukazuje obrázek na str. 57.

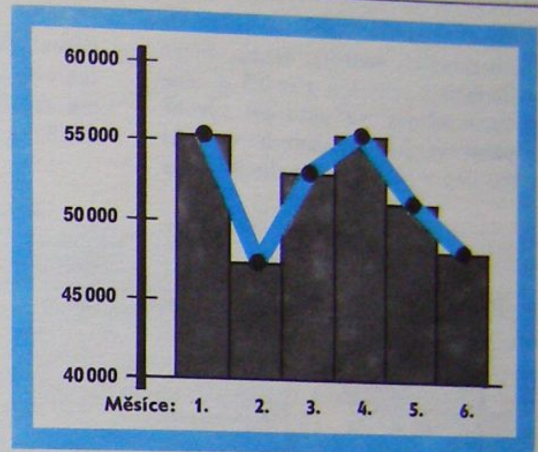
Nebo histogram spotřeby mléka v Rakousku — abychom uvedli příklad z úřední statistiky a současně ukázali vnitřní členění uvedených sloupků a ploch. (Jde jen o použití mléka vyrobeného v dané oblasti, tím se vysvětluje ono minimální „použití" v miliónové Vídni, kde je ovšem jen velmi málo krav.)

Polygon četnosti, který mnohdy ani není polygonem četnosti, nýbrž jakýmkoliv jiným polygonem, se po jazykovém rozboru jeví jako „mnohoúhelník", úzce spřízněný s pětiúhlým pentagonem a



Polygon (četnosti) stejných výrobních údajů.

Histogram — polygon: zde stejné vyjádření různými optickými prostředky. Záměnu obou metod znázornění nelze však vždy účelně provést.



trigonometrií, která vychází z trojúhelníku. Při použití tohoto způsobu vyjádření by vypadaly četnosti podnikové výroby tak, jako je tomu na obr. na str. 58.

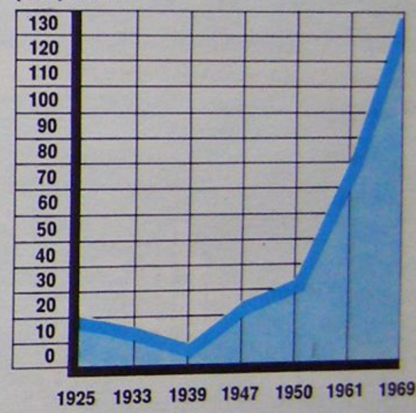
Souvislost mezi histogramem a polygonem je úplně zřejmá, jestliže promítneme obě vyobrazení přes sebe: uprostřed horního konce jednotlivých sloupků histogramu jako by byl fixován bod, a spojíme-li tyto body, vznikne přímka, někdy také křivka.

Zobrazení úsečkami, resp. křivkami se s oblibou používá u časových řad, tedy grafického zobrazení hodnot, které byly měřeny v delších časových obdobích. Časový prvek se přitom vždy nanáší na osu x (abscisa). Také k tomu uvádíme příklad z oblasti komunální statistiky (viz obr. vpravo).

Mnohá rozdělení četností mají v grafickém znázornění velmi charakteristickou formu. Jednou z nejdůležitějších a také nejproslulejších křivek rozdělení četnosti je tzv. *normální* nebo *zvonovitá křivka*, které chceme pro její značnou důležitost věnovat větší část následujícího oddílu.

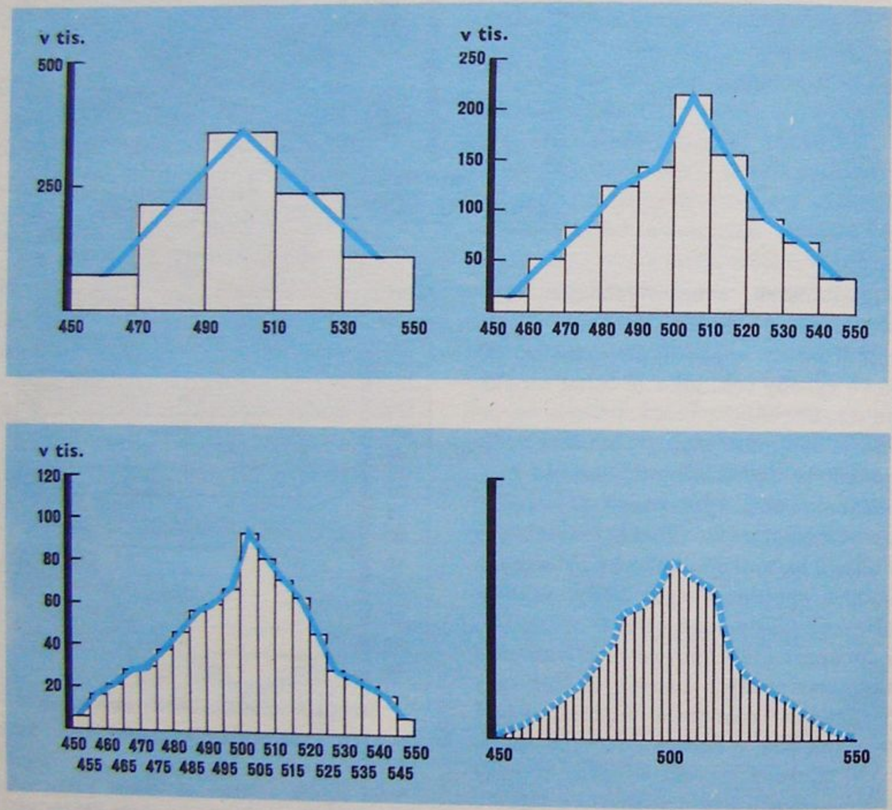
Cizinci v Mnichově

(osoby v tis.)



Počet cizinců v Mnichově během posledních 45 let. (Podle grafu mnichovského Úřadu pro statistiku a zpracování dat.)

Nejdříve však ještě jedno upozornění: u některých rozměrových čísel je zcela nemožné provést grafické vytřídění, v jiných případech je naopak možno provádět téměř neomezené třídění znaků. Předpokládejme např., že vyjadřujeme váhu jednoho miliónu kancelářských spínátek stejného druhu, která může činit v průměru $\bar{x} = 0{,}5$ g. Zde můžeme výkyvy ve váze při výrobě v pozorovaných četnostech vyjádřit mimo jiné v pěti, deseti nebo patnácti třídách. Můžeme však také přesnost vážení dále zvyšovat, až je počet tříd teoreticky neomezený, takže vyjadřujeme rozdělení hustoty pro jednotku „celková produkce".



Postupným členěním základního souboru do stále většího počtu tříd se stále menším třídním intervalem sloupky diagramu přecházejí pozvolna v jemné šrafování, zpočátku přímé linie polygonu se mění v křivku, četnost se mění v hustotu.

3 Normální rozdělení

3.1 Znaky a relativní četnosti

Statistika se nezabývá jednotlivými čísly, individui, nýbrž vždy *množstvím čísel, soubory, „velkými čísly"* nebo alespoň *výběrovými soubory*, které jsou však samy vytvořeny z větších celků a umožňují závěry o ještě větších souborech.

Je proto jen přirozené, že při sestavování statistického materiálu nestačí používat jen absolutní čísla, nýbrž že je třeba *tato čísla vyjadřovat ve vztahu k souboru;* vytvářet tedy relace, *vztahy mezi statistickým základním souborem* (nebo také výběrovým souborem) a *statistickým znakem*, a to tak, že se např. řekne: 10,3 % obyvatel Mnichova v r. 1969 byli cizinci a osoby bez státní příslušnosti. Tím se vytváří jasný poměr mezi základním souborem „obyvatelé Mnichova" a podskupinou, která je charakterizována znakem „cizinci a osoby bez státní příslušnosti".

U tzv. *diskrétních (nespojitých) znaků* je vytváření takového vztahu většinou velmi snadné. Diskrétní znaky jsou totiž ty, které lze zcela jednoznačně a jasně rozlišit — alespoň v teorii; např. muži a ženy (ačkoliv ojediněle se vyskytují i hermafroditi), cizinci nebo domácí (přičemž může být problémem dvojí státní příslušnost) atd. Pomocí výběrového souboru nebo celkového šetření lze poměrně jasně určit, kolik procent má znak A, který ostatní části nemají. Nejoblíbenějším způsobem zobrazení této skutečnosti je kruhová výseč (*„dortové řezy"*); plocha kruhu tvoří celek, kruhové výseče podíly (obr. na str. 62).

Pouze zdánlivým, nikoliv však skutečným protikladem diskrétních znaků jsou tzv. *spojité znaky*. Ty měříme („diskrétní" znaky se většinou sečítají) a jejich měření a číselné vymezení závisí koneckonců na přesnosti měřicích přístrojů. Při dotazu na mou váhu odpovídám: „70 kg". Na váze v koupelně mohu zjistit 70,5 kg, na přesnější váze 70 kg 450 g a konečně známe i tak přesné váhy, které zaznamenají i vdech a výdech.

V takových případech bude zjišťování spojitého znaku poznamenáno určitou nepřesností v důsledku měření a při statistickém zpracování zjištěných údajů bude asi nutno přikročit k třídění tak, jak jsme je načrtli již na konci předchozího oddílu. Jestliže však jsou třídy již vytvořeny, je možno provádět rozdělení četností uvnitř základního souboru stejným způsobem jako u diskrétních znaků. Pak se říká: „8,5 % dospělých žen váží mezi 50 a 55 kg" nebo „16 % rodin ve Zbohatlíkově má roční důchod méně než 15 000 tolarů".

Zvláštní případ představují *znaky, které zpočátku nejsou číselně vůbec zjistitelné.* Jak např. rozdělit zaměstnance na spokojené a nespokojené? Je samozřejmě možné se jich — přímo nebo méně přímo — na to zeptat, je možno získat o tom poznatky psychometrickými nebo sociometrickými postupy, nicméně jako propočtová základna zůstanou nakonec přece jen dosti vágní intenzitní znaky (velmi spokojen, spokojen, málo spokojen, nespokojen). Statistik se bude v takových případech snažit (jako každý vědec) dosáhnout přesné kvantifikace pomocí nějakých měření nebo testů, jako třeba v případě kvocientu inteligence IQ. Často však nezbývá než nějaká stupnice pořadí nebo preferencí, ze které se snad odvodí „nejčetnější hodnota", případně i „prostřední hodnota", ale stěží matematická střední hodnota.

Na tomto základě se pak zhruba vytvoří třídy a s pohledem na základní soubor se řekne: „32 % všech gymnazistů má ve škole uspokojivé výsledky ve fyzice."

Proč takové zdůrazňování *relativní četnosti*? Protože na ní je — v daleko větší míře než na absolutních měrných hodnotách — *vybudována celá stavba moderní statistiky*. Protože standardizací na jednotný vztažný soubor je umožněno strukturální srovnávání, které by se jinak ztratilo ve spoustě čísel rozdílné velikosti. Brzy poznáme normované normální rozdělení, ve kterém jsou všechna měření vztažena na jednotný pravděpodobnostní prostor 1 a toto normální rozdělení ovládá teorii výběru a s ní téměř všechny novější statistické postupy.

Také pro klasickou statistiku, pro popisnou statistiku vyčerpávajícího šetření a pro sčítání lidu je však *propočet procentních podílů* — zvláště pro srovnávání — účelný tehdy, jestliže jsou srovnávané základní soubory různě veliké: strukturální porovnání s dneškem umožňuje spíše *relativní* podíl rolníků na obyvatelstvu Skotska v roce 1900 než absolutní hodnota. Stejně tak lze jen pomocí procent porovnat negramotnost v malé Ghaně s negramotností ve velké Brazílii. Nutno však hned podotknout, že je jen velmi málo mezinárodních srovnání, která by byla úplně bez problémů. Přes veškeré snahy o jednotný obsah a rozsah pojmů má řada z nich, jako např. „mrtví při nehodách", „nezaměstnaní", „potrestaní" nebo „vysokoškoláci", různý obsah a nevyjadřuje všude přesně totéž. Mezinárodní statistiky se jen hemží odkazy a poukazy na rozdílný obsah pojmů.

Kruhové výseče („dortové řezy") se často používají k zobrazení podílů jednotlivých skupin na souboru: zde máme ženský svět rozdělen na blondýnky, černovlásky a rusovlásky. Kde zůstaly brunetky? Zařadili jsme je do ostatních skupin, abychom neměli tolik „řezů". Libovolné sdružování skupin není nic neobvyklého, zvláště jedná-li se o značně zjednodušené statistické výpovědi a zobrazení.

I tam, kde nejde o srovnání mezinárodní nebo s jiným časovým obdobím, se s velkou oblibou, zejména v populární statistice, používají čísla „na jednoho obyvatele", která spíše uvádějí v omyl než pomáhají. Tak např. údaj: „v Rakousku připadá na jednoho obyvatele 150 šilinků nákladů na výzkum", neříká laikovi vůbec nic. Stačí však, aby byl uveden slovy jako „jen" nebo „směšných", a laik bude pokládat uvedenou částku za nízkou. Kdyby dodatek zněl „přece jen" nebo „značná částka", zdálo by se mu stejné číslo mnohem větší. Až se později budeme zabývat demagogickými interpretacemi statistik, vrátíme se i k podobným jemným rozdílům.

Nyní je třeba poukázat ještě na jinou *poměrnou veličinu*. Není to vždy celé obyvatelstvo, na které něco připadá, mohou to také být jen „osoby výdělečně činné" nebo „majitelé osobních aut". Takové poměrné ukazatele mají často mnohem větší statistickou vypovídací schopnost (např. je-li počet porodů vztažen na počet žen v plodném věku, a nikoliv na celé obyvatelstvo). Někdy však jde jen o taškářské triky, s jejichž pomocí má situace vypadat jinak, než jaká je ve skutečnosti.

Prozatím se však nemusíme tolik zabývat poměrnými čísly, nýbrž *rozdělením znaků v základním nebo výběrovém souboru*. Brýlatí stejně jako řidiči mopedů, blondýny stejně jako běžkyně na krátkých tratích i nositelé jiných „znaků" jsou v obyvatelstvu nějak rozděleni, a to stejně jako je číslo 25 „rozděleno" ve sledu tažených čísel loterie nebo ve výsledcích rulety v souboru 49 nebo 37 čísel, stejně jako jsou rozděleny vadné žárovky nebo šrouby bez závitů v pásové výrobě. Svůj zájem však věnujme nejdříve základní struktuře tohoto rozdělení znaků, a tím vlastně elementární teorii pravděpodobnosti.

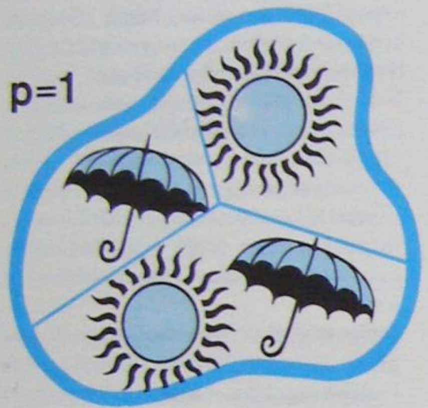
3.2 Počet pravděpodobnosti a statistika

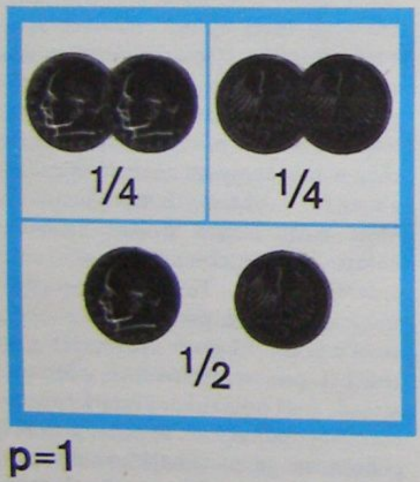
Někdy se říká, že *statistika je užitý počet pravděpodobnosti*, a na tomto tvrzení je nepochybně něco pravdy, pokud se nevezme doslovně. Zde musíme připomenout úvahu z úvodní kapitoly této knihy: *statistiku lze rozdělit na statistiku popisnou a induktivní*.

Popisná (deskriptivní) statistika, která je nejtypičtější pro sčítání lidu a podobná hromadná šetření, se zabývá především tím, jak přehledně uspořádat nepřehledné množství údajů, jak vyjádřit jejich znaky a provést zjednodušení všude tam, kde se tím nezmenší jejich vypovídací schopnost. Až do 20. století se statistika pěstovala převážně v tomto smyslu jako *popis veřejného života, politická aritmetika, matematická věda o státě*.

Stále více však nyní vystupuje do popředí zájmu *induktivní statistika*, statistická analýza. Ta sice také pracuje s měřením, sčítáním, šetřením, ale nemá k dispozici celkové soubory, často ani větší části těchto celkových souborů, opírá se o vzorky, které představují malé nebo někdy velmi malé podíly z daného základního souboru, jehož struktura a náplň mají být zjištěny a zobrazeny na základě výsledků získaných z výběrových vzorků.

Most mezi těmito dvěma hlavními oblastmi statistické práce tvoří *teorie pravděpodobnosti*. Poskytuje matematický základ pro posuzování spolehlivosti a přesnosti všech výběrových postupů i pro spolehlivost a přesnost metod, jimiž jsou získané vzorky zpracovávány. Skutečnost, že počet pravděpodobnosti je matematickou disciplínou, nás nesmí mýlit v tom, že důraz je na druhém slově — pravděpodobnost. Často lze počítat na mnoho desetinných míst, ale pravděpodobnost končí mnohdy již před desetinnou čárkou! „*Zákony pravděpodobnosti*" jsou proto zákony zcela zvláštního druhu — snášenlivé, pružné, nezavrhující pošetilé krajnosti, a přece dlouhodobě spolehlivé, nechybující a důvěryhodné. Proto se musíme alespoň krátce zastavit u některých základních pojmů počtu pravděpodobnosti.





Celkový prostor vymezující možnosti musí obsahovat vždy všechny alternativy: bude svítit slunce nebo bude pršet anebo bude „proměnlivo" či „zataženo". A jedna mince hozená dvakrát ukáže buď v obou případech orla nebo hlavu, nebo jednou hlavu a jednou orla.

Téměř každý výklad pojmu matematické pravděpodobnosti vychází z házení mincí, a i my zůstaneme této staré tradici věrni. (Autor byl této tradici věrný i ve své knize *„Sáhni po štěstí — náhoda, pravděpodobnost, jistota"*, ve které se počet pravděpodobnosti i oblasti jejího použití — od loterie až k moderní fyzice — probírají podrobněji.)

Hodíme-li si mincí, může se, jak známo, po dopadu ukázat buď hlava (H), nebo orel (O). Pro oba tyto jevy je pravděpodobnost a priori stejná, předpokládáme-li ovšem, že mince není falešná. Tato stejná pravděpodobnost se vyjadřuje tak, že se řekne: $p_H = p_O$, nebo, jak se rovněž často píše, $P(H) = P(O)$.

V tomto případě platí: $P(H) = 0{,}5$, stejně i $P(O) = 0{,}5$, neboť každý jednotlivý jev je z poloviny pravděpodobný, stejně jako absolutní jistota, že jeden z obou jevů nastane. Tato jistota, tento celkový prostor všech myslitelných možností se tedy vyjadřuje jednotkou (1). Součet jednotlivých pravděpodobností nemůže proto za žádných okolností být větší než 1.

I v hovorové řeči známe ostatně matematickou pravděpodobnost. Hodnoty se násobí dvěma a říká se: „naděje na úspěch je jedna ku jedné". Zde však jsme přesní: $P(H) = 0{,}5$, $P(O) = 0{,}5$. To jsou pravděpodobnosti pro jedno hození mincí.

Házíme-li mincí vícekrát, lze vypočítat pravděpodobnost jednotlivých výsledných kombinací (HOO, HOH, HHH atd.) násobením výchozích pravděpodobností ($p_H = 0{,}5$ atd.), protože tu jde o navzájem zcela nezávislé pokusy, neboť mince ani kostka nejsou ovlivněny dřívějšími výsledky. (Jinak by tomu bylo u předpovědi počasí: je-li dnes pěkně,

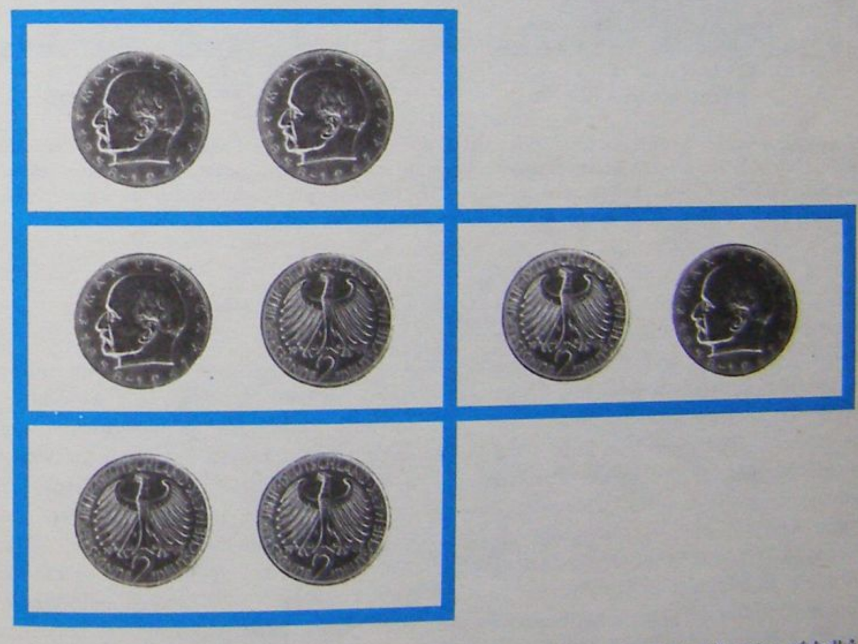
je též větší pravděpodobnost, že bude pěkně i zítra, než kdyby dnes pršelo — počasí dnes a počasí zítra nejsou navzájem nezávislé.)

Pro jev „dvakrát hlava" dostaneme: $(p_H = 0{,}5) \cdot (p_H = 0{,}5) = p_{HH} = 0{,}25$; stejně $p_{OO} = 0{,}25$. K tomu přistupují ještě obě možnosti: pro „jednou hlava, jednou orel", tzn. $p_{OH} = 0{,}25$, tedy nejdříve orel, pak hlava, a $p_{HO} = 0{,}25$ (nejdříve hlava, pak orel). Zajímá-li nás celkový výsledek, nikoliv však pořadí, dostaneme při dvojím hodu mincí tyto pravděpodobnosti: $p_{HH} = 0{,}25$, p_{OO} (v libovolném pořadí) $= 0{,}5$.

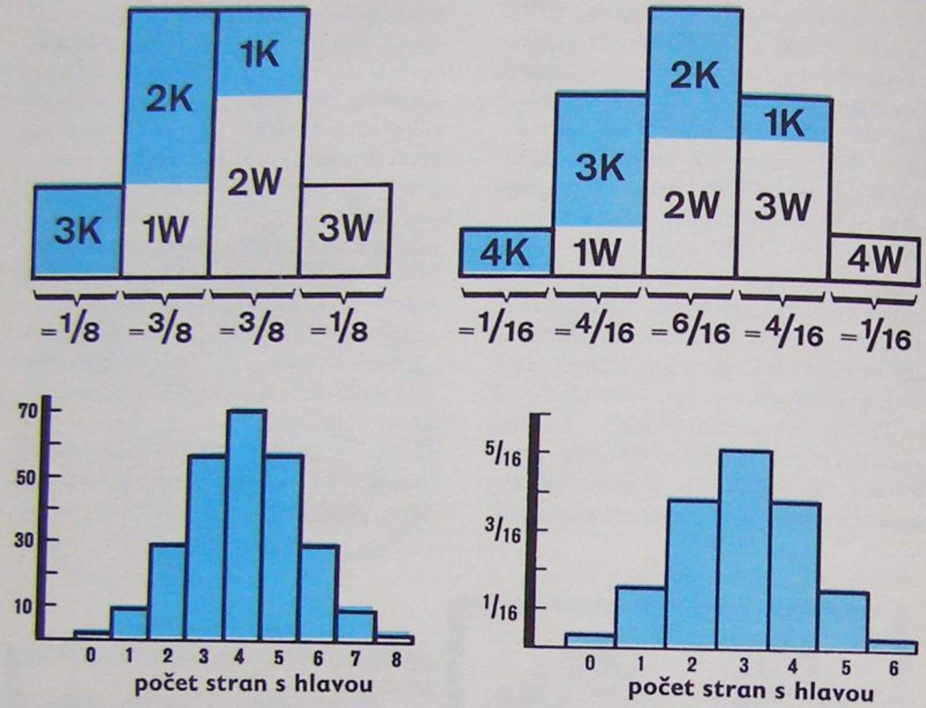
Podobným způsobem lze vypočítat očekávané četnosti při třech, čtyřech a více hodech mincí a získané rozdělení četnosti zobrazené jako *histogram* se vyvíjí způsobem uvedeným v obrázku na str. 66 nahoře. Z toho, co jsme uvedli, je zřetelně vidět, že se vytváří velmi charakteristický tvar, který se projeví ještě silněji, jestliže přes histogram promítneme *frekvenční polygon*, podobně jako jsme to udělali na konci oddílu 2. 7, a jestliže počet pokusů ještě zvýšíme (viz oba obrázky na str. 67).

Velmi dobrým názorným zobrazením takového rozdělení je také model *„římské kašny"*, kde voda přetékající z jednotlivých mís odtéká stejnoměrně vlevo i vpravo.

Další model, znázorňující posloupnost stejných pravděpodobností náhodných jevů, představuje tzv. *Galtonovo prkno*,



Rozdělení pravděpodobnosti při dvojím hození mincí. Při větší sérii pokusů dává čtvrtina výsledků dvakrát hlavu, polovina hlavu a orla, čtvrtina dvakrát orla. Pootočíme-li vyobrazením o 90 stupňů, dostáváme navíc histogram sestávající ze tří sloupků s relativními četnostmi 1, 2 a opět 1.

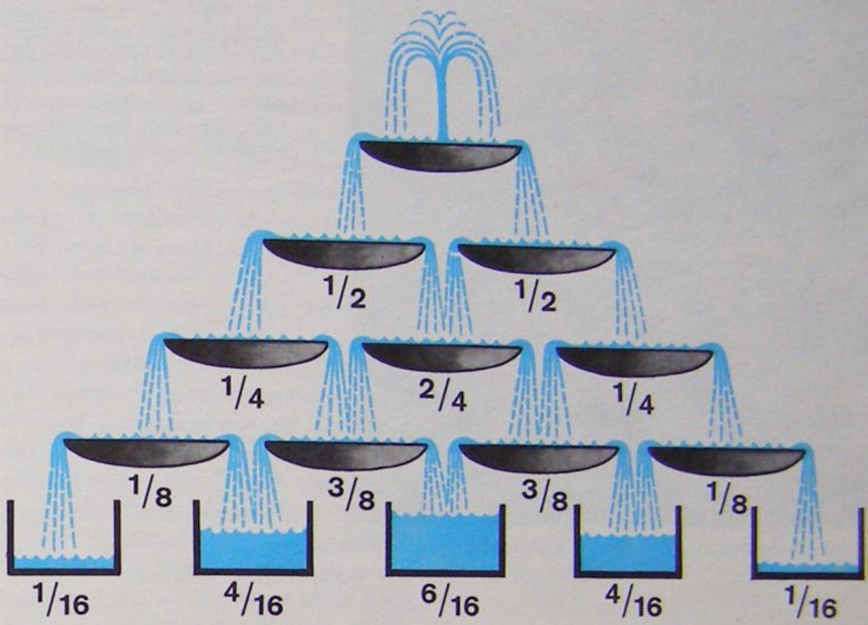


Pravděpodobnosti výsledků při troj, čtyř, šesti a osminásobném hození mincí. Relativní četnosti se doplňují vždy na celek 1 všech možností. V grafu osminásobného hození jsou naproti tomu uvedeny očekávané absolutní četnosti (celek = 256, relativní pravděpodobnosti se vypočítávají dělením: pravděpodobnost pro 4krát hlava a 4krát orel činí například 70/256).

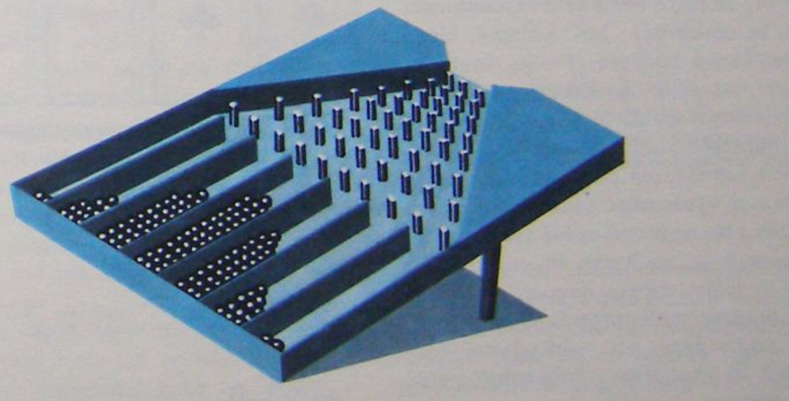
které kýčovitě a pseudoelektronicky vyparáděno stále ještě tvoří základní konstrukční princip četných naivních hazardních hracích automatů. Základní myšlenkou přitom je, že koule kutálející se po nakloněném prkně shora dolů narazí na kolíček, od něhož se odrazí vpravo nebo vlevo dolů. Tam narazí do dvou kolíčků, aby se znovu náhodně odrazila a narazila na trojici kolíčků atd. Po libovolném počtu řad kolíčků koule zapadne do přihrádky, nejčastěji uprostřed, do vedlejší přihrádky padne méně často a celkem zřídka padá do krajních přihrádek.

3.21 Binomické rozdělení

V uvedených případech je před-námi tzv. *binomické rozdělení*. Je pro ně charakteristické, že má jen *dvě rozdílná vyjádření znaků*, která však společně vyplňují celý pravděpodobnostní prostor. Při hodu mincí je to samozřejmé, protože jsou jen dva znaky (hlava nebo orel). Avšak již při tažení hrací karty máme čtyři „barvy". V takových případech, a ty jsou v praxi velmi časté, rozlišuje se prostě *„hledaný znak"* a *„všechny ostatní"*. Jak velká je pravděpodobnost, že z dobře zamíchaného balíčku karet



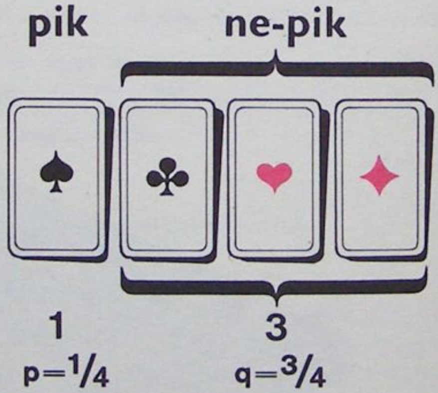
Binomické rozdělení zobrazené pomocí modelu římské kašny.



Galtonovo prkno s kolíčky tvoří dodnes „mechanismus náhody" pro četné hazardní hry. Jsou-li kolíčky zasazeny uvedeným způsobem, utvoří se pořadí kuliček podle Pascalova trojúhelníku (schéma římské kašny); v našem příkladu: 1 ku 6 ku 15 ku 20 ku 15 ku 6 ku 1.

„Galtonovo prkno“ je dnes známější než kdykoliv jindy, i když nikoli pod tímto názvem. Tento mechanismus náhodného rozdělení odchýlením kuliček na překážkách vpravo nebo vlevo je totiž základním konstrukčním principem hracích automatů.

vytáhneme pikovou kartu? Zřejmě jedna čtvrtina, zatímco zbytek — tři čtvrtiny — zůstává pro „nepiky“: „žaludy“, „srdce“ či „kára“ nebo jak se jim jinak říká.
Proto se často vedle pravděpodobnosti p, že očekávaný „jev“ nastane, klade doplňková hodnota $(1 - p)$, která se také označuje q. Potom $p + q = 1$. Proto čím menší je p, tím větší je q a obráceně. Jestliže máme znak „muži“, je v každém jen poněkud reprezentativním výběrovém souboru obyvatelstva p jen nepatrně menší než q. Jestliže znakem je „muži přes 40 let, kteří byli v minulém roce v Portugalsku“, pak p je velmi malé, protože q tvoří nejen všechny ženy, ale také všichni muži pod 40 let a všichni, kdo nebyli v Portugalsku.
Binomické rozdělení tedy udává, *jaká je pravděpodobnost pro určitý výsledek vý-*



Pravděpodobnost nějakého jevu se doplňuje pravděpodobností „nejevu“, a tím se dosáhne jistoty: buď vytáhnu z dobře zamíchaného balíčku karet pik, nebo „nepik“.

běrového souboru. Tvoří do jisté míry empirickou základnu pro teorii výběrových souborů. Umožňuje nám z přesně známého základního souboru, např. 52 karet, zjistit pravděpodobnost výsledku pro každé pořadí tahů. V praxi při vytváření výběrových souborů skutečný základní soubor *neznáme*, avšak víme, jak je pravděpodobné, že výběrové soubory poskytují jeho skutečný nebo zkreslený obraz.
Podívejme se např., kolik piků bude „pravděpodobně“ taženo, jestliže se z dobře zamíchaného balíčku karet táhne pětkrát za sebou po jedné kartě. (Tažená karta se přitom dá zpět a zamíchá se společně s ostatními; v daném případě máme před sebou model „*hypergeometrického rozdělení*“.)
Při jednom jediném tahu platí $p = 0{,}25$. Při 5 tazích je tedy početní „očekávaná hodnota“ $5\,p = 1{,}25$. Není však dosti dobře možné táhnout 1,25 piku; očekávaná hodnota slouží proto v tomto případě (i v mnoha jiných případech) jen jako jistý druh orientační pomůcky. Můžeme již teď předpokládat, že „1 pik mezi 5 kartami“ se objeví poměrně často. Nejjednodušší však bude, když vypočítáme nejdříve oba extrémy: 5 piků, případně žádný pik. Protože při tazích jde o navzájem nezávislé náhodné jevy (dali jsme přece taženou kartu vždy zpět a dobře se zamíchalo), platí pravidlo o násobení pravděpodobností: $p_{5\ \text{piků}} = (0{,}25)^5$. Tento výsledek by bylo možné očekávat jen jednou v 1000 sériích pokusů. Častěji se dá počítat s 5 nepiky: $q_{5\ \text{nepiků}} = (0{,}75)^5$. To tedy bychom věděli. Ale co teď? Jak často se vyskytne 1 pik a 4 nepiky?
Můžeme vypočítat, kolik je $(0{,}25) \cdot (0{,}75)^4$, avšak k tomuto výsledku lze dojít více než jedním způsobem, neboť 1 pik může být tažen jako první, ale také jako druhá, třetí, čtvrtá nebo dokonce pátá karta. Co lze v daném případě poměrně pohodlně vypočítat, nelze tak lehce zjistit ve složitějších případech (třeba 3 piky z 8). V dané obtížné situaci najdeme pomoc ve zvláštním uspořádání čísel, které je známo jako „*Pascalův trojúhelník*“, ačkoliv celkem neprávem, protože Pascal není ani jeho znovuobjevitelem, nýbrž poznal jej u svého učitele Hérigona. Podobné uspořádání nacházíme však už o 150 let dříve v knize „*Arithmetica integra*“ Michaela Stifela (Norimberk 1544); stejně tak se toto číselné uspořádání vyskytuje již v perské a čínské literatuře 13. a počátku 14. století, takže by bylo lépe, kdyby bylo označeno neutrálně, např. jako „*aritmetický trojúhelník*“. Jak tedy tato číselná sestava vypadá?

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1
1 7 21 35 35 21 7 1
. . . .
. . . .

Pascalův nebo aritmetický trojúhelník tvoří jakýsi přechod z číselné magie k teorii pravděpodobnosti. Z jednotlivých řádků trojúhelníku je možno vyčíst nahodilé pravděpodobnosti jevů všeho druhu.

Vrstva jedniček obklopuje trojúhelník, ve kterém je každé číslo složeno ze součtu obou nad ním šikmo, vpravo i vlevo, stojících čísel. To, co nejprve vypadá jako podivná hračka, nabývá rychle na významu, když znovu připomeneme římskou kašnu a necháme vodu stékat ze stupně na stupeň do dvou, tří, čtyř atd. mís. Předpokládáme-li v nejvyšší míse „jednotku" vody, pak v následujícím stupni se dělí na $^1/_2 + {}^1/_2$, na třetím stupni na $^1/_4 + {}^2/_4 + {}^1/_4$, v dalším na $^1/_8 + {}^3/_8 + {}^3/_8 + {}^1/_8$ atd., jak je vyznačeno v uvedeném schématu.

To přesně odpovídá výsledkům našeho házení mincí. Při jednom hodu je pravděpodobnost $^1/_2$ k $^1/_2$, při dalším je pravděpodobnost $^1/_4$ (*HH*) k $^1/_2$ (*HO*, resp. *OH*) a k $^1/_4$ (*OO*), při třech hodech $^1/_8$ *HHH*, $^3/_8$ *HHO*, $^3/_8$ *HOO*, $^1/_8$ *OOO* atd.

Je pravděpodobné, že tato sestava také vyvolává vzpomínky na polozapomenutou školní matematiku. V ní byla binomická poučka: $(a + b)^2 = a^2 + 2ab + b^2$, tzn. rozdělení $1(a^2)$ k $2(ab)$ k $1(b^2)$, stejně jako v Pascalově trojúhelníku, resp. v čitateli zlomků naší římské kašny. V případě $(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$ se to také potvrzuje, což obecně znamená, že *v Pascalově trojúhelníku jsou přesně a pečlivě uspořádány binomické koeficienty*. Pro laika v matematice zpočátku téměř neřešitelný problém, např. $(a + b)^5$, lze nyní vyřešit skoro bez námahy, prostě přečíst: na pátém řádku čteme 1 — 5 — — 10 — 10 — 5 — 1, tudíž: $(a + b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$.

Teď už není třeba nic jiného než za *a* a *b* dosadit *p* a *q*, a můžeme vypočítat pravděpodobnost všech kombinací pěti tahů z balíčku karet. Hledáme především případ pq^4, tzn. 1 pik a 4 nepiky. Protože $pq^4 = (0{,}25) \cdot (0{,}75)^4$, a $(0{,}75)^4 = 0{,}317$, dostaneme pravděpodobnost 0,08, kterou však musíme ještě násobit četností 5, protože v řádku je $p^5 + \ldots + 5\,pq^4 + q^5$. Protože $(0{,}08) \cdot 5 = 0{,}4$, je hledaná pravděpodobnost „1 pik v pěti tazích" asi 0,4. Podobným způsobem vypočítáme:

žádný pik:	$(0{,}75)^5$	$= 0{,}24$
jeden pik:	(viz nahoře)	$= 0{,}40$
dva piky:	$10\,(0{,}25)^2\,(0{,}75)^3$	$= 0{,}26$
tři piky:	$10\,(0{,}25)^3\,(0{,}75)^2$	$= 0{,}09$
čtyři piky:	$5\,(0{,}25)^4\,(0{,}75)$	$= 0{,}01$
pět piků:	$(0{,}25)^5$	$= 0{,}001$

Součet všech těchto výrazů dává zase 1, tedy prostor pravděpodobnosti všech výběrových souborů v rozsahu 5 tahů. Největší dílčí souhrn z nich je $5(0{,}25) \cdot (0{,}75)^4 = 0{,}4$, tzn. ten, který jsme vypočítali pro 1 pik, což je úplně správné, neboť „1 pik" zobrazuje situaci „očekávaná hodnota 1,25" nejlépe.

Binomické koeficienty jsou však pro teorii pravděpodobnosti velmi významné také z jiného hlediska: *pomáhají nám při výpočtu kombinací*. Jestliže chci např. umístit 3 osoby ve dvousedadlovém sportovním autě, jsou 3 možnosti pro vytvoření následujících kombinací dvojic: *AB*, *BC* a *AC*. Otázka formulovaná obecně i matematicky zní: kolik kombinací po *k* jednotkách lze vytvořit z *n* jednotek?

Odpověď: $\binom{n}{k}$, což se čte „*n* nad *k*" a vypočítává podle následujícího vzorce:

$$\boxed{\binom{n}{k} = \frac{n!}{k!\,(n-k)!}}$$

Uvedené vykřičníky se čtou jako „*faktoriál*" („*n* faktoriál"), což znamená, že příslušné číslo se běžně násobí svou hodnotou sníženou vždy o jednu : $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$ atd.

Když je tedy třeba $n = 5$ a $k = 2$, dostaneme:

$$\binom{n}{k} = \binom{5}{2} = \frac{5!}{2!\,.\,3!} = \frac{5 \,.\, 4 \,.\, 3 \,.\, 2 \,.\, 1}{2 \,.\, 1 \,.\, 3 \,.\, 2 \,.\, 1},$$

což po krácení dává 10, a to je těch 10,

Kdo s kým? Kombinační počet všedního dne. Jestliže dvě ze tří osob sedí ve sportovním voze, jsou možné tyto kombinace: $\frac{n!}{k!\,(n-k)!}$. Poněvadž v daném případě $n = 3$, $k = 2$, jsou tedy $\frac{6}{2} = 3$ možné kombinace.

které udávají možnosti kombinací $k = 2$ piky v $n = 5$ tazích.

Stejně při $k = 1$, $\binom{5}{1} = 5$. Ale jak je to při $k = 0$? Zde je nutno vzít v úvahu, že v definici „nulový faktoriál" (0!) není 0, nýbrž 1. Proto $\binom{5}{0} = \frac{5!}{0!\,.\,5!} = 1$, což je 1 z $1p^5$, resp. z $1q^5$.

Jestliže však *n* vzroste na více než 10, vzroste 11! atd. až na nepohodlně velké veličiny, které mohou být určeny pouze za pomoci přibližných vzorců. Naštěstí se však potom může provést rozdělení, a to nejen v případě $p = q = 0{,}5$, za pomoci mnohem pohodlnějšího normálního rozdělení, které je i při menších *p* plně uspokojivé, zatímco u velmi malého *p* se použije *Poissonova rozdělení*, jímž se také ještě budeme zabývat.

Tím chceme zakončit tento kurs první pomoci v počtu pravděpodobnosti a kombinatorice.

Poznámka pro *etymology-amatéry:* proč „*binomický*", „*binominální*", „*binom*"? Ex binis nominibus, „ze dvou výrazů" (totiž *p* a *q*, piky a nepiky atd.).

3.22 Zákon velkých čísel

Hráčská kasina zpravidla neznají úpadek. Je tomu tak z mnoha důvodů, z nichž dva nejdůležitější jsou: každá hra provozovaná v kasinu má včleněnu takovou možnost výhry, že bank musí v dlouhodobém výhledu dosáhnout zisku, a za druhé, že téměř všichni hráči jsou zatíženi pověrčivostí, která je pokřivena představou o „*zákonu velkých* *čísel*". V jejich představách se snoubí elementární počet pravděpodobnosti a naivní optimismus s chybnou matematikou do fantastické legendy, kterou lze vyjádřit takto: „Jestliže desetkrát po sobě padla červená, musí dříve nebo později padnout častěji černá, protože zákon velkých čísel tvrdě vyžaduje, aby červená a černá padaly koneckonců stejně často."

Stejnou představu by bylo možno uplatňovat, když v ruletě nebo ve Sportce dlouho nevyšlo určité číslo. Mnoho hráčů Sportky sází systematicky čísla, která dlouho nebyla tažena nebo jsou tažena nadprůměrně zřídka, na základě přesvědčení, že tato čísla jsou zralá, ba dokonce přezrálá k tahu. To je však legenda o „*dozrávání šancí*".

Podle vtipného výroku jednoho francouzského matematika nemá ruleta „ani svědomí, ani paměť". Moc málo lidí je však ochotno tomu věřit a své naděje upínají místo toho na špatně zažitou moudrost o „*zákonu velkých čísel*", který prý slibuje kompenzaci tím, že kategoricky vyžaduje, aby v dlouhodobém výhledu zcela nutně vyšla hodnota očekávaná podle teorie počtu pravděpodobnosti. Hodíme-li mincí stokrát, musí padnout hlava a orel stejně často, a nevyjde-li to ani při stém hodu, musí se tak stát při tisícím nebo desetitisícím nebo také při pozdějších hodech.

„Zákon velkých čísel" je ve svém původním znění znám již více než půl tisíciletí. Objevuje se v „*Ars coniectandi*" (Umění dohadu) Jakuba Bernoulliho, uveřejněném v roce 1713, po autorově smrti, a říká asi toto: Jestliže jsou pokusné řady dosti dlouhé a dostatečně často se opakují, lze dosáhnout vypočítané pravděpodobnosti v průměru těchto pokusů s libovolnou přesností. Důraz se přitom klade na *průměr četných* řad pokusů.

Jestliže to, co jsme zde uvedli, ilustrujeme na praktickém příkladu, vypadá to asi takto: pravděpodobnost čísla 4 při házení bezvadnou kostkou je — protože kostka má šest ploch — rovna $^1/_6$. Protože očekávaná hodnota navzájem nezávislých pokusů se rovná pravděpodobnosti správného hodu (p), násobené počtem pokusů (n), vychází při 6 pokusných hodnotách očekávaná hodnota 1 $(6 \cdot {}^1/_6)$, při 300 pokusech je očekávaná hodnota 50 $(p = {}^1/_6, n = 300)$, při 9000 je očekávaná hodnota 1500.

„Zákon velkých čísel" vůbec neříká, že při 9000 hodech zcela jistě hodím 1500 čtyřek bez ohledu na to, zda mám smůlu nebo štěstí. Především jde o to, že je možno stanovit uvedené číslo při třístovkových, tisícových nebo jiných sériích, jejichž výsledky vyjádřené jako aritmetický průměr se koneckonců neodchýlí od očekávané hodnoty o více než 1 nebo 3 nebo 7 anebo jinou libovolnou hodnotu. V jediné pokusné řadě nemůže být ani řeči o kompenzaci v průběhu řady.

To, co nastane — a to jsme měli možnost pozorovat v náznacích již v Pascalově aritmetickém trojúhelníku — je koncentrace četností v blízkosti očekávané hodnoty se současným stále širším celkovým rozptylem. Při dvou hodech mincí vyjde očekávaná hodnota 1 hlava $(p = {}^1/_2, n = 2)$ v polovině pokusné řady, při dvanácti hodech vyjde očekávaná hodnota 6 $(p = {}^1/_2, n = 12)$ jen ve $\frac{924}{4096} = 22{,}5\ \%$, neboli méně než v $^1/_4$ všech případů. Souhrn „blízkých" výsledků (totiž 4, 5, 6, 7 a 8 hlav) dává však celkově asi 70% pravděpodobnost, zatímco horní a dolní krajní případy

27. sázkový týden SPORTKY — II. tah

1 $\frac{2}{55}$	2 $\frac{4}{38}$	3 $\frac{6}{51}$	4 $\frac{2}{51}$	5 $\frac{4}{61}$	6 $\frac{1}{61}$	7 $\frac{4}{53}$
8 $\frac{2}{57}$	9 $\frac{2}{57}$	10 $\frac{2}{47}$	11 $\frac{4}{60}$	12 $\frac{1}{47}$	13 $\frac{3}{55}$	14 $\frac{1}{43}$
15 $\frac{5}{52}$	16 $\frac{2}{51}$	17 $\frac{4}{56}$	18 $\frac{4}{47}$	19 $\frac{7}{72}$	20 $\frac{4}{47}$	21 $\frac{4}{42}$
22 $\frac{5}{59}$	23 $\frac{4}{45}$	24 $\frac{2}{55}$	25 $\frac{5}{49}$	26 $\frac{3}{56}$	27 $\frac{4}{41}$	28 $\frac{3}{46}$
29 $\frac{5}{54}$	30 $\frac{5}{61}$	31 $\frac{0}{49}$	32 $\frac{6}{49}$	33 $\frac{4}{56}$	34 $\frac{3}{63}$	35 $\frac{1}{52}$
36 $\frac{4}{53}$	37 $\frac{5}{52}$	38 $\frac{4}{56}$	39 $\frac{7}{50}$	40 $\frac{4}{54}$	41 $\frac{1}{45}$	42 $\frac{3}{55}$
43 $\frac{2}{48}$	44 $\frac{1}{61}$	45 $\frac{1}{47}$	46 $\frac{3}{62}$	47 $\frac{3}{53}$	48 $\frac{4}{50}$	49 $\frac{0}{50}$

27. sázkový týden SPORTKY — I. tah

1 $\frac{2}{117}$	2 $\frac{2}{108}$	3 $\frac{0}{83}$	4 $\frac{3}{103}$	5 $\frac{2}{106}$	6 $\frac{2}{91}$	7 $\frac{0}{92}$
8 $\frac{1}{110}$	9 $\frac{2}{100}$	10 $\frac{5}{95}$	11 $\frac{1}{108}$	12 $\frac{5}{110}$	13 $\frac{3}{110}$	14 $\frac{3}{103}$
15 $\frac{3}{108}$	16 $\frac{3}{98}$	17 $\frac{4}{103}$	18 $\frac{1}{88}$	19 $\frac{4}{109}$	20 $\frac{5}{105}$	21 $\frac{5}{92}$
22 $\frac{2}{110}$	23 $\frac{5}{111}$	24 $\frac{2}{106}$	25 $\frac{2}{90}$	26 $\frac{2}{99}$	27 $\frac{5}{110}$	28 $\frac{5}{109}$
29 $\frac{4}{109}$	30 $\frac{2}{101}$	31 $\frac{6}{99}$	32 $\frac{6}{109}$	33 $\frac{2}{90}$	34 $\frac{6}{117}$	35 $\frac{4}{112}$
36 $\frac{4}{121}$	37 $\frac{3}{102}$	38 $\frac{4}{91}$	39 $\frac{3}{100}$	40 $\frac{1}{111}$	41 $\frac{2}{88}$	42 $\frac{6}{104}$
43 $\frac{2}{113}$	44 $\frac{5}{108}$	45 $\frac{3}{79}$	46 $\frac{4}{97}$	47 $\frac{4}{115}$	48 $\frac{7}{111}$	49 $\frac{5}{95}$

Četnost tahů jednotlivých čísel SPORTKY. Pro informaci sázejících je udána četnost tažení jednotlivých čísel, a to v I. a ve II. tahu od počátku roku (za 27 týdnů) a od počátku Sportky (v I. tahu to bylo 841 tažení a ve II. tahu 429).

Četnost tahu čísel Sportky ukazuje nahodilé rozdělení. V I. tahu bylo během 841 her taženo 5046 čísel, přičemž každé ze 49 čísel mělo očekávanou hodnotu četnosti tažení rovnou 5046 : 49 = 102,98, tj. 103. Ve II. tahu bylo během 429 her taženo 2574 čísel. Každé ze 49 čísel mělo očekávanou hodnotu četnosti tažení rovnou 52,5. V I. tahu se očekávaná hodnota objevila třikrát; většinou byla jednotlivá čísla tažena v rozmezí 90—110, výjimečné jsou extrémní hodnoty 79 a 121. Podobně tomu bylo i ve II. tahu, kde celá čísla sousedící s očekávanou hodnotou (52,5) se objevují šestkrát a zároveň se objevuje extrémně nízká četnost tahů u čísla 2 a extrémně vysoká u čísla 19.

V horní části zlomku jsou údaje o tom, kolikrát bylo které číslo sportky taženo letos, spodní část zlomku značí údaj, kolikrát bylo číslo taženo od začátku existence sportky.

(0, 1, 2, 3 a 9, 10, 11, 12 hlav) nedávají celkem víc než 30 %.

Co dělá „zákon velkých čísel" tak důležitým pro statistiku, není jeho užitečnost nebo bezcennost v hráčském kasinu, nýbrž jeho význam při využití výběrových souborů. Protože pokusná řada je prakticky výběrovým souborem, platí i v tomto případě, že *očekávanou hodnotu je možno určit (případně skutečnou hodnotu celkového množství) s libovolnou přesností*, jestliže jsou vytvořeny dostatečně velké vzorky. V části, kde se budeme zabývat výběry podrobněji, uvidíme, že v praxi se však musí dělat většinou velmi skromné kompromisy.

3.3 Normální rozdělení

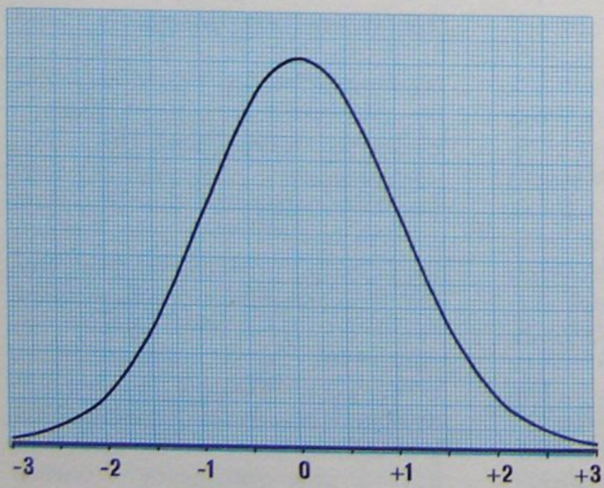
Jsou-li všichni havrani černí, není třeba vytvářet výběrové soubory a uvažovat o tom, kolik by mohlo být havranů šedých nebo světle modrých. Je-li pravděpodobnost, že havrani jsou černí, $p = 1$, jde o určitost.

Něco jiného je, zkoumám-li např. váhu sta, tisíce nebo ještě většího množství havranů. I když se přitom vyloučí mláďata, objeví se uvnitř zkoumaného souboru výkyvy. Podobně se navzájem odlišuje tělesná výška dospělých mužů, kvocient inteligence školních dětí, počet slov na plně potištěných stránkách knihy, přesná délka sériově vyráběných věšáků na zeď, životnost elektrických žárovek, ba dokonce navzájem nezávislá měření jedné a téže vzdálenosti. Jistě jste si vzpomněli i na rozdělení příjmů našich 25 abiturientů a obyvatel Zbohatlíkova, avšak právě tato rozdělení musíme z našeho pojednání prozatím vyloučit. Naproti tomu se sem velmi dobře hodí úvahy, které jsme rozvíjeli v souvislosti s binomickým rozdělením a zákonem velkých čísel. Výsledky dvaceti hodů mincí stejně jako téměř každého druhu výběrových šetření vykazují totiž také charakteristické znaky tzv. normálního rozdělení, jímž se nyní budeme zabývat.

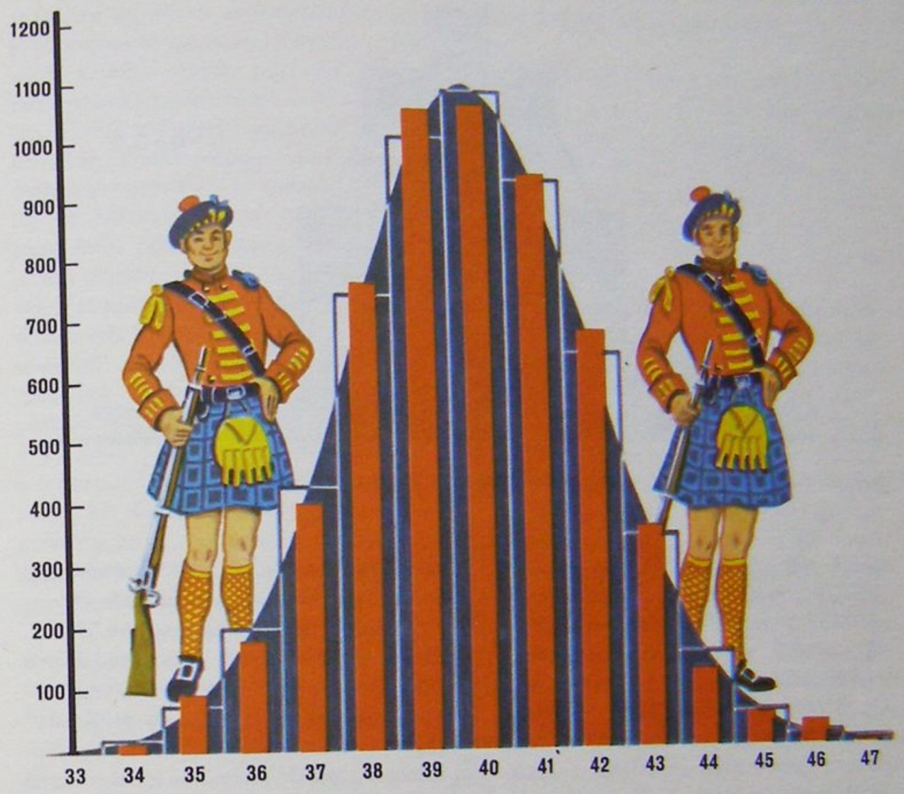
Může být sporné, zda označení „*normální rozdělení*" je zvoleno šťastně (kritika někdy uvádí, že slovo „normální" naznačuje jakési souhlasné hodnocení), nicméně ve všech světových jazycích se přesto mluví o „*normálním rozdělení*", „*Normalverteilung*", „*normal distribution*", „*distribution normale*" — tzn. že statistická terminologie je bez tohoto slova nemyslitelná —, aniž by statistika proto napadlo, že jiná rozdělení než normální jsou „abnormální". Nebude také ani požadovat, aby se daná čísla a měřené hodnoty daly ve všech případech úzkostlivě uvést do souladu s ideálním obrazem normálních rozdělení. *Normální rozdělení* je jako každé jiné statistické rozdělení *především myšlenkovým modelem a početní pomůckou*, tedy nikoliv exaktním přírodním zákonem, který by musel být naplněn s malichernou přesností.

Jestliže se statistika vždy, kdykoli vystupuje jako induktivní statistika, snaží usuzovat o celku na základě vzorků a dílčích pozorování, a proto může poskytnout jen více nebo méně pravděpodobné odhady, bylo by dvojnásob absurdní, kdyby se předstírala přesnost, která se především ve skutečnosti v takovém stupni nikdy nevyskytuje a která je dále vyloučena samým nahodilým charakterem každého výběrového souboru.

Teorie a praxe však již prokázaly správ-



Normální křivka nebo lépe jedna z normálních křivek vzhledem k tomu, že by ji bylo možno nakreslit strmější nebo mnohem plošší. Podstatný je jen vztah výše křivky v bodech, které vymezují směrodatnou odchylku. Body ležící na křivce ve vzdálenosti + nebo — 2 σ směrodatné odchylky se nacházejí v $^1/_8$ výšky, kterou má křivka normálního rozdělení ve vrcholu (nad průměrem 0).

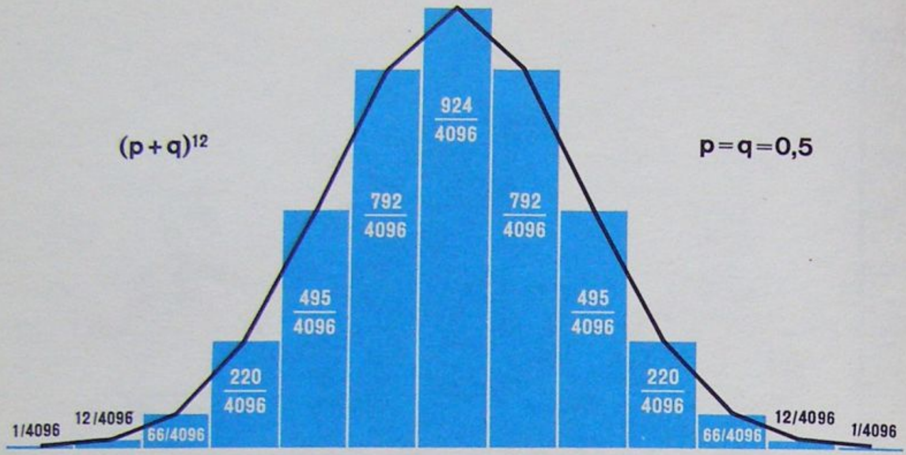


Obvod hrudi skotských vojáků podle Quételetovy statistiky. Shoda pozorovaných hodnot s normálním rozdělením je až zarážející (μ = 39,8 palce).

nost domněnky, že normální rozdělení platí pro téměř všechny výběry a pro velmi mnohá rozdělení podchytitelných souborů.

Normální rozdělení má především velmi příjemnou vlastnost, která lehce vysvětluje jeho oblibu: ať už jde o jakékoliv objekty, úkazy, měření nebo sčítání, *jsou vždy jednoznačně určeny střední hodnotou a rozptylem*. Pokud jde o „střední hodnotu", má se zpravidla, a nikoliv neprávem, na mysli aritmetický průměr $\bar{x}$ nebo μ, a to platí i v případě normálního rozdělení. Avšak stejnou hodnotu jako aritmetický průměr mají i modus a medián (nejčetnější hodnota a prostřední hodnota). Krátká úvaha hned prokáže, že normální rozdělení s největší četností „uprostřed" musí být symetrické. Jak taková normální křivka, která je grafickým znázorněním normálního rozdělení, ve skutečnosti vypadá, ukazuje obrázek na str. 74.

Základem normálního rozdělení je zkušenost, že bezpočetné znaky a hodnoty jsou rozloženy tak, že jeden vý-



Binomické rozdělení $(p + q)^{12}$ dovoluje již jasně poznat podobu normální křivky.

sledek měření nebo sčítání je „nejčetnější" a „na obě strany od něho" jsou výsledky ponenáhlu stále méně četné, až konečně vykazují jen ojedinělou extrémní hodnotu. Jeden z prvních příkladů takového normálního rozdělení podal Quételet na základě měření obvodu prsou 5738 skotských vojáků. Nejčetnější hodnota činila (zaokrouhleno na celé coule) 40 coulů, skoro stejnou četnost vykazovalo 39 coulů, 41 a 38 coulů se vyskytovalo již vzácněji, 42 a 37 byly ještě vzácnější a konečně zjištěných 33, resp. 48 coulů představovalo jen ojedinělé extrémní hodnoty.

Podobné uspořádání vykazují i výsledky dalších četných měření, např. váha cigaret vyráběných cigaretovým automatem: v nejčetnějších případech vážily 1,18 až 1,20 g nebo 1,20 až 1,22 g, jen málokteré byly lehčí než 1,08 g nebo těžší než 1,32 g.

Podle rozsahu rozptylu i podle měřítka zvoleného pro grafické znázornění vzniká v idealizované formě plošší nebo strmější křivka.

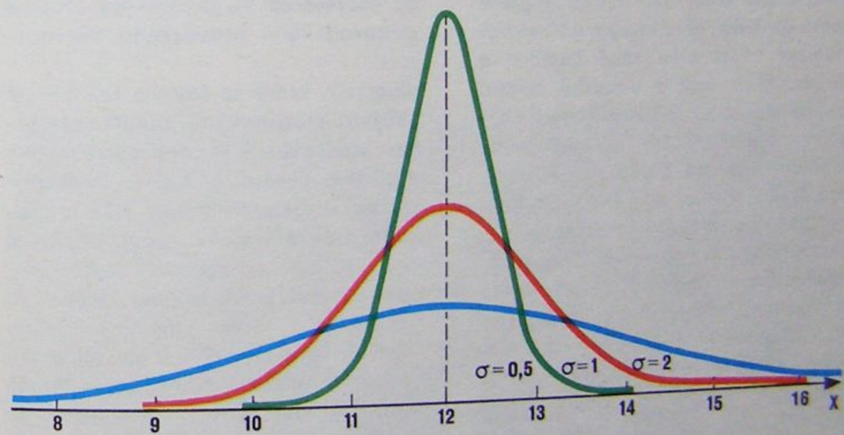
Mluví se o „*normální křivce*", v němčině o „Glockenkurve" *(zvonovité křivce)* se zřetelem na středně vysokou křivku, která se také uvádí jako model pro normovanou normální křivku, ve Francii o „*courbe en chapeau de gendarme*" (křivku policejního klobouku) se zřetelem na plošší normální křivku. Jako vědecká označení se také používají názvy „*Gaussova křivka rozdělení chyb*" a „*de Moivrova stochastika*" s odvoláním na oba praotce zvonovité křivky.

Ať už se zvolí jakékoliv měřítko a ať je rozptyl k průměru v jakémkoliv poměru, normální křivka má vždy některé charakteristické znaky, z nichž uvedeme ty, které jsou pro praxi nejdůležitější. Jestliže pozorujeme plochu ležící pod „zvonem" jako soubor, leží *na obě strany od maxima* (střední hodnota, nejčetnější hodnota) *vždy přesně stejné části této plochy*, a to v úseku mezi $+\sigma$ a $-\sigma$ leží 68,26 %, tj. nepatrně více než $^2/_3$ celkové plochy, v úseku mezi $+2\sigma$ a -2σ skoro přesně 95 % a mezi $+3\sigma$ a -3σ již 99,7 % plochy.

Hodnoty za třemi směrodatnými odchylkami se proto berou v úvahu již jen velmi zřídka, ačkoliv normální křivka se teoreticky rozkládá od $-\infty$ po $+\infty$, tzn. v jisté míře od nekonečna do nekonečna. Tímto *postupem do nekonečna a plynulým průběhem* se normální křivka liší od binomického rozdělení. Přesto však mezi těmito oběma rozděleními je tak těsná souvislost, že *normální rozdělení* je možno v téměř všech prakticky důležitých případech pokládat za dostatečně přesné *vyjádření binomického rozdělení*, čímž si lze ušetřit svízelné početní operace, které jsme alespoň v náznaku poznali při našich úvahách o binomickém rozdělení. V historii normální křivky nelze zamlčet její původ z teorie hazardní hry. U její kolébky stál stařešina počtu pravděpodobnosti Abraham de Moivre.

3.31 Dějiny normální křivky

Normální křivka oslavila své 250. narozeniny a na rozdíl od jiných vědeckých objevů je lze dokonce přesně zjistit: bylo to 12. listopadu 1733. To je totiž datum uveřejnění malého spisku Abrahama de Moivre, matematika — hugenota vyhnaného z Francie, který se všelijak protloukal v Londýně, mimo jiné také tím, že radil hazardním hráčům. De Moivre započal své úvahy experimenty s házením mincí, které jsme uvedli v části o binomickém rozdělení, a uskutečnil myšlenkový skok od sloupků histogramu k plynulé křivce, až konečně objevil i křivkovou rovnici normální křivky, kterou však zde neuvádíme, protože matematika laika spíše zmate, než by mu něco vysvětlila. De Moivre už také pozoroval stejno-



Tři normální rozdělení kolem střední hodnoty $\mu = 12$. Má-li rozdělení silný rozptyl ($\sigma = 2$), je křivka plochá a rozložená; je-li rozptyl malý ($\sigma = 0,5$), je strmá a vysoká. Střední křivka vykazuje proporce „normovaného normálního rozdělení".

rodost ploch uvnitř směrodatných odchylek. Pro teorii hazardní hry nebylo však možno z jeho matematických poznatků získat mnoho nového a jiné praktické možnosti použití elegantně zahnutých křivek nebylo v polovině 18. století ještě možno nalézt; křivka i rovnice upadly v zapomenutí. Až když se vynořil praktický problém, přímo prahnoucí po početním zvládnutí, byla normální křivka znovuobjevena jako Gaussova-Laplaceova „*křivka chyb*“.

Proč křivka chyb? Na konci 18. a počátku 19. století byli astronomové neustále v nepříjemné situaci, která vznikala tím, že při svých měřeních získávali hodnoty, které se více nebo méně navzájem odlišovaly. Příčina spočívala v nedokonalosti používaných přístrojů, a tak se na věci nedalo prozatím nic měnit. Bylo proto nutné nalézt cesty, jak by se ze spousty měřených dat získala pravděpodobně správná hodnota. Vzpomeneme-li si na obrázek ukazující zásahy kolem černého středu střeleckého terče, lze si trochu představit problém, před kterým stáli astronomové, mezi nimi také Laplace a Gauss, kteří měli z umístění zásahů pokud možno co nejpřesněji vypočítat polohu cíle. Z množství měření navzájem více nebo méně odchylných, např. vzdáleností stálic, se měla určit hodnota nejvíce se blížící skutečnosti.

Gauss se nejdříve domníval, že za takovou správnou hodnotu lze považovat aritmetický průměr všech měření, potom však došel, podobně jako Laplace, k představě rozdělení četností výsledků měření: zcela nesprávné jsou ojedinělé extrémy, pak následují měření stále více si podobná a zároveň četnější, doslova pak vrcholící ve střední hodnotě a zároveň nejčetnější hodnotě normální křivky.

Jak Laplace, tak i Gauss se museli mimo jiné vypořádat s otázkou, *zda nelze nalézt stejné uplatnění pro kladné i záporné výsledky*. Je zřejmě nesmyslné, aby chyby v měření +2 a +5 se navzájem prostě zrušily chybami —2 a —5 a přitom se tvrdilo, že per saldo se vlastně žádná chyba nestala. Laplace postavený před tuto otázku se rozhodl pro absolutní hodnotu chyby tím, že uvažoval asi takto: jestliže dvě více nebo dvě méně, *obojí je „dvě špatně“*. Gauss volil jiné východisko, a sice to, které jsme již poznali při výpočtu směrodatné odchylky: *chyby umocnil na druhou*, a tak dostal vesměs kladné hodnoty.

Vítězné tažení normálního rozdělení zahájil však teprve Adolphe-Lambert Quételet, všestranný belgický vědec, který objevil normální rozdělení pro biometrii, vědu o měření člověka. Quételet byl jedním ze zakladatelů Královské statistické společnosti v Londýně, hlavním povoláním astronom, ze záliby statistik, jehož vášeň přecházela až do rozvernosti — jednou prý spočítal celkovou váhu bruselského obyvatelstva.

Quételet, který se často a tak přesně zabýval statistickými tabulkami, během studia zjistil, že překvapivě mnoho výsledků sčítání a měření vykazuje takové rozdělení četnosti, jaké je charakteristické pro „křivku chyb“, která mu byla jako astronomovi dobře známa. Provedl proto mnohá měření na vlastní pěst, mezi nimi jako jedno z prvních proslulé měření obvodu prsou skotských vojáků. Na tomto základě dospěl k závěru, že — nepřihlíží-li se k poměru mezi střední hodnotou a rozptylem — uspořádání četností vykazuje při těchto a jiných biometrických měřeních skutečně přesně stejnou strukturu, jaká je v tabulkách chyb měření: *křivku normálního rozdělení*. A pokud je známo, Quételet byl první, kdo mluvil o „*normálním*“ *rozdělení*.

Normální rozdělení podle Quételeta neznamenalo však nic jiného než to, že příroda se snaží vytvořit ideální typ, „*homme moyen*“, avšak v různé míře chybuje. V této souvislosti použil také obrazu se střeleckým terčem — „... hodně smělá metafora“, jak kriticky komentoval později logik a teoretik pravděpodobnosti John Venn. (Je to týž Venn, kterému moderní matematika vděčí za Vennovy diagramy pro zobrazení množin.)

„Homme moyen“ se může doslova přeložit jako „průměrný člověk“ nebo „střední člověk“, avšak bez onoho znevažujícího spodního tónu, který byl úplně cizí Quételetovu myšlení. Pro něho byl „homme moyen“ vznešenou, vysněnou představou, která svědomitému statistikovi umožňovala přesně nahlédnout do dílny přírody a objevovat její nejtajnější cíle. Tato představa narazila na rozhořčený odpor, zároveň však získal vášnivé a nadšené stoupence. K těmto stoupencům patřila také sociální pracovnice Florence Nightingalová, nadšená amatérská statistička, a biolog, kriminolog a afrikanista Francis Galton, který s veškerou energií jako jeden z prvních zavedl kvantitativní metody do biologie a mimo jiné navrhl měrné stupnice pro všechny možné tělesné znaky, ba dokonce i pro ženskou krásu.

Galton — ačkoliv vlastně nebyl básníkem — se nechal svým nadšením pro normální rozdělení strhnout až k téměř básnickému projevu a mimo jiné napsal: „Stěží znám něco, co může naši fantazii uchvátit tak jako obdivuhodná forma kosmického pořádku, který vyjadřuje »zákon o četnosti chyb«. Kdyby byl znám Řekům, byli by jej personifikovali a uctívali jako božstvo. V nejdivějším zmatku šíří harmonický klid; čím větší anarchie, tím svrchovanější je jeho vláda. Za závojem chaosu se objevuje jeho netušený a nádherný tvar pravidelnosti.“

THE
NORMAL
LAW OF ERROR
STANDS OUT IN THE
EXPERIENCE OF MANKIND
AS ONE OF THE BROADEST
GENERALIZATIONS OF NATURAL
PHILOSOPHY + IT SERVES AS THE
GUIDING INSTRUMENT IN RESEARCHES
IN THE PHYSICAL AND SOCIAL SCIENCES AND
IN MEDICINE AGRICULTURE AND ENGINEERING +
IT IS AN INDISPENSABLE TOOL FOR THE ANALYSIS AND THE
INTERPRETATION OF THE BASIC DATA OBTAINED BY OBSERVATION AND EXPERIMENT

Americký statistik Youden „zanotoval“ typografickou pochvalnou píseň na normální rozdělení: Je „vedoucím nástrojem výzkumů v přírodních a sociálních vědách...“, „nepostradatelným nástrojem analýzy a výkladu výsledků pozorování a pokusů“.

Karl Pearson, otec moderní matematické statistiky, byl rovněž obdivovatelem normální křivky, ačkoliv znal i její nedostatky: stanovil, *že v přírodě jsou i nenormálně rozdělené veličiny*, a také se pokusil vypracovat specifická schémata rozdělení pro tyto případy. Přitom však zjistil, že z mnohých, na první pohled nenormálních rozdělení se po pečlivém rozboru skutečností vyklubou spletence dvou nebo i více normálních rozdělení.

Vášnivé debaty o významu a bezcennosti, oprávněnosti a nesmyslnosti normální křivky utichly až ve 20. století, i když úplně nevymizely. Přijímá se taková, jaká je — *jako cenná pomůcka statistické práce*. O její filozofické nebo i vědeckoteoretické podstatě se však příliš nemluví. Fyzik Gabriel Lippman prý k tomu jednou zlomyslně poznamenal: „Celý svět pevně věří na normální rozdělení — matematici, protože jsou přesvědčeni, že jde o experimentálně dokázanou skutečnost, experimentátoři, protože jsou přesvědčeni, že se jedná o matematickou poučku."

Mezi nejoblíbenější „vysvětlení" normální křivky stále ještě patří Gaussovo: nesčetné dílčí vlivy vyvolávají větší nebo menší odchylky od „průměru", který všude nacházíme, a tato náhodná kombinace náhodných vlivů podléhá nakonec „zákonům" hazardní hry, pravidlům binomického rozdělení s téměř nekonečným počtem „pokusů". Tato úvaha nachází matematický výraz v „*centrální limitní větě*", jíž se na tomto místě nebudeme blíže zabývat. S pomocí této věty však lze ukázat, že vždy se musí dojít alespoň přibližně k normálnímu rozdělení, jestliže je znak určen působením většího počtu navzájem nezávislých vlivů, ať už je každý z těchto jednotlivých faktorů rozdělen jakkoli.

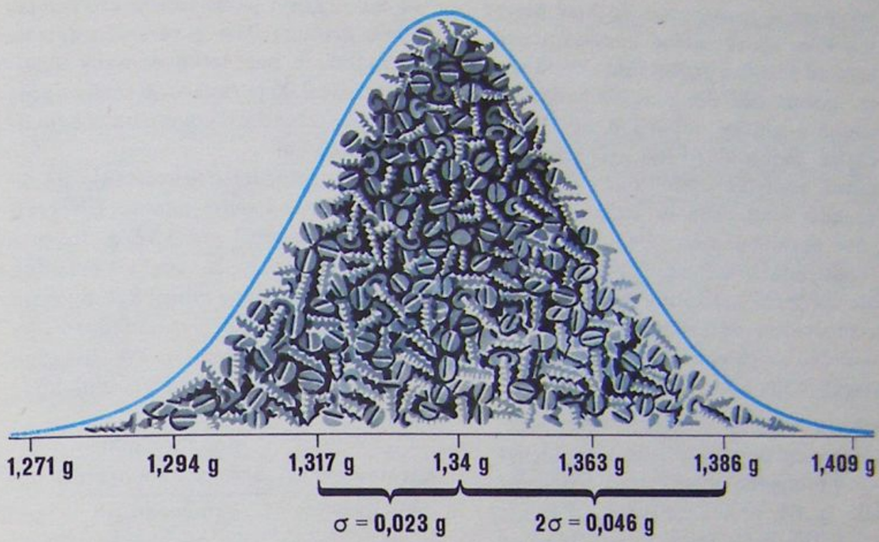
S jemnou ironií se dá snad říci, že volná hra náhody je to jediné opravdu normální na tomto světě.

Einstein prý jednou řekl: „Bůh nehraje v kostky." Normální rozdělení však nemůže posloužit jako důkaz opaku. Je to koneckonců otázka individuálního postoje, jestliže se v něm spatřuje jen matematický výraz nahodilých kombinací nemajících zdánlivě smyslu nebo pravěký zákon stvoření a důkaz božského řádu v kosmu, jak kdysi soudil de Moivre.

3.32 Normované normální rozdělení

I když jsou normální křivky pravidelné, symetrické a stejnorodé, získávají velký praktický význam teprve dalším procesem *standardizace (normování)*. K tomu, abychom porozuměli procesu standardizace, musíme uvést ještě několik příkladů nestandardizovaného normálního rozdělení. Charakteristické chování křivky je vždy stejné: bod obratu křivky leží vždy ve vzdálenosti $+\sigma$ a $-\sigma$, tečna v bodu obratu vždy protíná souřadnici ve vzdálenosti $+2\sigma$ a -2σ, teoreticky je vždy křivka rozložena na obě strany donekonečna, avšak již při $+3\sigma$ a -3σ se prakticky dotýká osy úseček (souřadnice x).

Nanášené měrné hodnoty jsou však nestejné. Váha šroubů se může někdy odchylovat o směrodatnou odchylku 0,023 g od průměru 1,34 g; jindy mohou psychometrická šetření skupiny studentů vykazovat rozptyl 36 ($\sigma = 6$) kolem kvocientu inteligence 112; tělesná váha, výsledky sklizně, stejně tak jako libovolný počet sčítaných a měřených výsledků mohou být rozděleny přibližně normálně kolem střední hod-



Průměrná váha určitého druhu šroubů je 1,34 g, směrodatná odchylka 0,023 g; protože jde o normální rozdělení, budou se velmi vzácně vyskytovat šrouby, které váží méně než 1,294 nebo více než 1,386 g.

noty se směrodatnou odchylkou, která může jednou činit půldruhého dne, jindy 0,85 kg, 3,2 cm, 0,8 ohmů, 35 marek nebo 3,5 mm.

Mají-li být všechna tato rozdílná měření a sčítání opravdu účelně zobrazena normálním rozdělením, je žádoucí, aby byl k dispozici *standardizovaný soubor nástrojů*, které umožňují odpověď na velmi rozličné otázky: „Kolik % šroubů je mimo toleranční meze ± 0,06 g?" — „Odchyluje se výrazně rozptyl inteligenčního kvocientu určité skupiny od rozptylu přibližně dvojnásobně velké skupiny, s níž byl proveden stejný test?" — „Vytvoříme-li výběrový soubor 50 žárovek, bude v něm s pravděpodobností větší než 68 % nejméně jeden zmetek?" — „Jak velká je pravděpodobnost, že v ruletě padne v průběhu nejbližšího tisíce sázek číslo 13 přesně třináctkrát?"

Převedení všech těchto rozmanitých měření, otázek a odpovědí na jediné schéma je možné jen tehdy, když je normální rozdělení *jednoznačně určeno* směrodatnou odchylkou a průměrem a když struktura normálního rozdělení *se nemění*, ať už jde o centimetry, hektolitry, ohmy anebo čísla v ruletě.

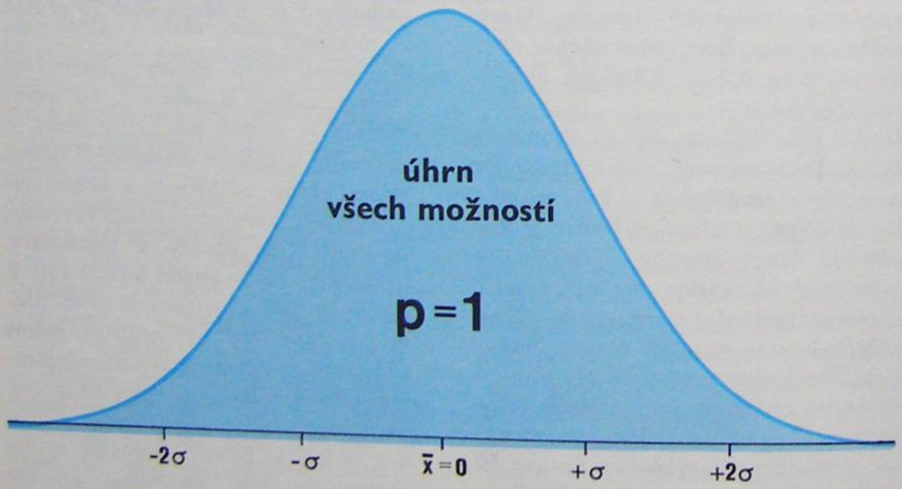
Použijme pro objasnění daného problému např. šroubů: nejdříve máme průměrnou hodnotu 1,34 g a směrodatnou odchylku 0,023 g. První krok: vynecháme gramy a zůstává $\mu = 1{,}34$ a $\sigma = 0{,}023$. Během druhého kroku provedeme další abstrakci: průměr $\mu = 1{,}34$ je pro standardizované normální rozdělení právě průměrem a průměr standardizované normální křivky se zásadně rovná nule. To není pouhá libovůle, nýbrž účelná konvence, protože od kterého jiného čísla než od nuly je možno tak lehce zjistit zrcadlově stejné odchylky nahoru a dolů? Odchylce „nahoru" (třeba +2) odpovídá stejná odchylka „dolů" (—2). Z těchto důvodů také tabulky standardizovaného normálního rozdělení, které jsou nepostradatelnou pomůckou při statistické práci, nerozlišují mezi kladnými a zápornými odchylkami. Přesto však je možno z nich vyčíst pravděpodobnost kteréhokoliv jevu a četnosti.

Jak to vypadá v praxi? Tak např. chceme zjistit, kolik šroubů se odchyluje v tom nebo onom směru o více než 0,06 g od průměrné váhy. Známe: $\sigma = 0{,}023$. Víme také, že rozdělení ve standardizovaném normálním rozdělení závisí již jen na směrodatné odchylce, protože jsme průměr posunuli na nulu. Zcela jednoduše proto porovnáme hledané 0,06 g se směrodatnou odchylkou a bez velké námahy vypočítáme, že 0,06 je rovno 2,6 směrodatné odchylky standardizovaného normálního rozdělení.

Danou skutečnost lze označit také jinak. Třeba takto: šrouby nemají být těžší než 1,40 g a lehčí než 1,28 g. Jaká je pravděpodobnost, že odchylky budou přesahovat uvedené váhy? Pak bychom měli podle vzorce pro standardizaci normálního rozdělení uvést hledané hodnoty do spojitosti se střední hodnotou:

$$\text{normovaná hodnota} = \frac{\text{hodnota minus aritmetický průměr}}{\text{směrodatná odchylka}}$$

Jestliže tuto „normovanou hodnotu" označíme z, může být vyjádření ještě kratší:



Standardizace normálního rozdělení: průměrná hodnota je vždy 0, odchylky se už neudávají v gramech, litrech nebo absolutních četnostech, nýbrž pouze v blíže nespecifikovaných směrodatných odchylkách.

$$\boxed{z = \frac{x - \mu}{\sigma}}$$

Dosadíme-li naše čísla, dostaneme:

$$\frac{1{,}40 - 1{,}34}{0{,}023} = 2{,}6, \text{ a stejně}$$

$$\frac{1{,}28 - 1{,}34}{0{,}023} = -2{,}6.$$

Záporné znaménko označuje odchylku pod střední hodnotu.

Nyní tedy víme, že $z = 2{,}6$. Ale co vlastně je z (nebo ať už tuto hodnotu nazveme jakkoliv, protože se vyskytuje pod různými názvy)? Chceme-li to tak vyjádřit, je to standardizovaná směrodatná odchylka. Stejně jako jsme zprvu přeměnili reálnou střední hodnotu $\mu = 1{,}34$ g na $\mu = 0$, tak i nyní jsme přeměnili reálnou směrodatnou odchylku $\sigma = 0{,}023$ na standardizovanou (normalizovanou) odchylkovou veličinu $z = 2{,}6$. A teď si vezmeme k ruce tabulky normálního rozdělení, které obsahuje ve více nebo méně přesné formě každá učebnice statistiky. Zacházení s nimi není pro začátečníka vždy lehké, jak hned poznáme. Najděme nyní hodnotu 2,6 a v tabulkách začněme hledat „*výsledek*".

První pokus. Použijeme „Metod statistiky" od Wallise a Robertse a najdeme: Normální proměnná 2,6 .... 0047. Dobře. To bychom měli — nebo z opatrnosti provedeme ještě jedno porovnání?

Vezměme Pfanzaglovu knihu „*Allgemeine Methodelehre der Statistik II*" (Všeobecná nauka statistických metod II) a najdeme:

x 2,6 0,495

Hleďme — ale co to je?

Ale je tu ještě Gotkinova a Goldsteinova programovaná učebnice „*Grundkurs in Statistik*" (Základní kurs statistiky[1]):

$\frac{x}{\sigma}$ nebo z 2,6 0,4953

O jedno desetinné místo přesnější, ale jinak naštěstí stejný výsledek. Proč jsme však u Wallise a Robertse našli něco jiného?

Sáhněme ještě po „*Mathematische Formeln*" (Matematické vzorce), kterou vydal Bartsch[2]:

λ 2,6 0,9953

Teď je zmatek úplný. Je snad ještě nějaké příslušné dílo v dosahu? Ale ano, Gnědenkův a Chinčinův „*Elementární úvod do počtu pravděpodobnosti*" s tabulkou hodnot v příloze:

a 2,60 0,991

Pět tabulek — čtyři různé hodnoty. Komu můžeme, komu smíme věřit?

3.33 Jak zacházet s tabulkami pravděpodobnosti — kvartily a percentily

Pozorný čtenář už asi s nedůvěrou pozoroval, že hodnota 2,6 má vždy jiné označení: jednou *a*, pak *x*, pak zase λ, potom „normální proměnná", pak $\frac{x}{\sigma}$ nebo z. Je samozřejmé, že již proto musí vždycky vyjít něco jiného. Ale ne, kupodivu v tom nedorozumění nespočívá. Ať jsou tyto hodnoty v tabulkách označeny jakkoliv, ve všech případech je myšleno totéž. Už dříve jsme

[1] Pozn. překl.: Podobně R. Reisenauer, Metody matematické statistiky (Praha 1970), str. 196.

[2] Pozn. překl.: Podobně J. Hátle — J. Likeš, Základy počtu pravděpodobnosti a matematické statistiky (Praha 1974) u 2,60...0,99534.

krátce upozornili na to, že terminologie statistiky (a především zkratek a symbolů) není vždy zcela jednotná.
Jak však potom vzniká rozdílnost hodnot? Doporučuje se nejprve důkladná orientace v tabulce, které hodnoty se vlastně vztahují k hodnotě *x* nebo *z*, event. *t* (jak se také někdy označuje). Tak např. *v prvním případě* (0.0047): „Pravděpodobnosti, že dané standardní normální proměnné budou překročeny; pravděpodobnosti se udávají pro horní konec." V *druhém případě* (0,495) se říká: *Φ* (x)*, v *třetím případě* se stejnou hodnotou (0,4953): „Podíl celkové plochy pod normální křivkou mezi středovou pořadnicí a pořadnicí v dané vzdálenosti z od střední hodnoty." Ve *„Vzorcích"* (0,9953) je stručně: *Φ (λ)*. Konečně v posledním případě (0,991) je obdobně stručně: *Φ (a)*. K tomu poznamenáváme, že třikrát uvedené *Φ* je velké řecké písmeno fí. Je-li bez hvězdičky, označuje plošný podíl pod normální křivkou až k hledanému bodu — na obě strany, a na jednu stranu, uvádí-li se jako *Φ**. A tento plošný podíl, jak si hned ukážeme, symbolizuje pravděpodobnost hledaného jevu.
Všechny čtyři výsledky jsou samozřejmě správné, jak lze od statistických tabulek právem očekávat. Část tajemství se podhalí upozorněním, že údaje se vztahují na *plochu pod normální křivkou* a že tato plocha v souladu s definicí *zachycuje celý prostor pravděpodobnosti*. Pravděpodobnosti se však prostírají jen od $p = 0$ = absolutní nemožnost až po $p = 1$ = absolutní jistota. Prostor pravděpodobnosti („jevové pole"), který se nám zde představuje jako pravděpodobnostní plocha, zahrnuje tedy pouze jednotku 1 celé pravděpodobnosti, nebo — v běžné řeči, násobeno 100 — oněch 100 %, která zahrnují všechny možnosti, všechny myslitelné jevy.
Podíváme-li se na poslední číslo ($\Phi (a) = 0{,}991$), bude nám nápadné, že leží již těsně u 1; říká, že 991 promile všech možných případů leží v této oblasti; tedy jen 9 promile mimo ni. Nemělo by snad těchto 9 promile zachycovat naše šrouby s vyšší váhou?
Jestliže se však znovu podíváme na číslo 0,991 a srovnáme-li je s 0,495, které se nám vyskytlo už dvakrát, ukáže se, že 0,991 je téměř přesně dvojnásobkem 0,495. A v třetím případě se přece uvádí: „Podíl ... celkové plochy ... mezi *středovou* pořadnicí a pořadnicí vzdálenou z ...". Zde se tedy uvažuje jen jedna polovina plochy, zatímco ve druhé tabulce *obě* poloviny. Názorně to ukazuje obrázek na str. 86.
Když však již víme, že většina tabulek operuje jen s polovinou plochy, pak i poslední číslo (.0047) nabývá nového významu: 0,0047 + 0,4593 (druhý a třetí příklad) tvoří součet 0,5, tedy přesně polovinu plochy pod křivkou. Všechno je v pořádku a jen někdo zlomyslný může prohlašovat, že nejen se statistikou, ale dokonce i s jednotlivými statistickými tabulkami je možno dokázat všechno, co se chce.
Zůstává ještě hodnota 0,9953, která se našemu cvičenému oku nyní ukazuje jako 0,4953 + 0,5 — k polovině pod křivkou se připojuje druhá polovina.
Ale co to vše říká o našich šroubech s vyšší vahou? Na naši otázku: „Kolik procent váží více než 1,4 g?", dostaneme tuto přesnou odpověď: „Pravděpodobnost výskytu šroubů s vyšší vahou je $p = 0{,}0047$" — tedy 4,7 promile bezpochyby bude překračovat stanovenou váhu. Dvě další odpovědi říkají: „V pravé (horní) polovině normálního rozdělení, která zabírá plochu 0,5, bude za-

Hodnoty *Φ* podle tří různých tabulek

Normální rozdělení

x → Φ* (x)	
x	Φ* (x)
0,0	0,000
0,1	0,040
0,2	0,079
0,3	0,118
0,4	0,155
0,5	0,191
0,6	0,226
0,7	0,258
0,8	0,288
0,9	0,316
1,0	0,341
1,1	0,364
1,2	0,385
1,3	0,403
1,4	0,419
1,5	0,433
1,6	0,445
1,7	0,455
1,8	0,464
1,9	0,471
2,0	0,477
2,1	0,482
2,2	0,486
2,3	0,489
2,4	0,492
2,5	0,494
2,6	0,495
2,7	0,496
2,8	0,497
2,9	0,498
3,0	0,499

Pravděpodobnosti, že standardizované směrodatné proměnné byly překročeny. Pravděpodobnosti jsou udávány pro horní konec

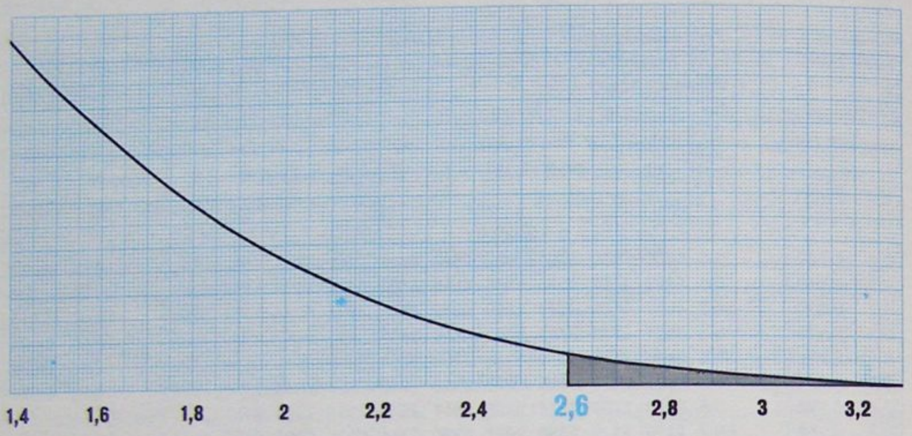
Normální proměnná	0	1	2	3	4	5	6	7	8	9
0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641
0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2777
0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021.	0021	.0020	.0019
2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010

Hodnotová tabulka veličiny *Φ* (a)

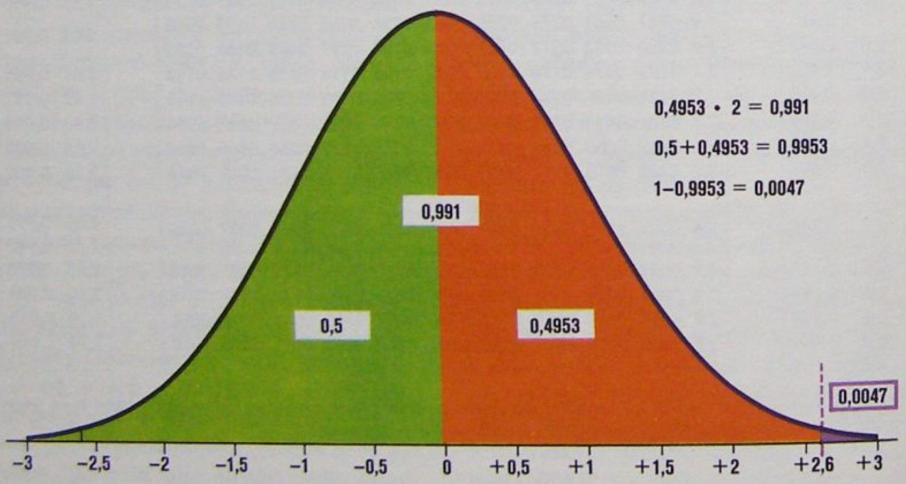
a	Φ (a)
0,00	0,000
0,01	0,008
0,02	0,016
0,03	0,024
0,04	0,032
0,05	0,040
.	.
.	.
.	.
2,50	0,988
2,51	0,988
2,52	0,988
2,53	0,989
2,54	0,989
2,55	0,989
2,56	0,990
2,57	0,990
2,58	0,990
2,59	0,990
2,60	0,991
2,61	0,991
2,62	0,991
2,63	0,991
2,64	0,992
2,65	0,992
2,66	0,992
2,67	0,992
2,68	0,993
2,69	0,993
2,70	0,993
2,72	0,993

Tabulky normálního rozdělení uvádějí na první pohled různé zmatené údaje. Naučíme-li se však účelně s nimi zacházet, můžeme z každé z nich vyčíst to správné.

hrnuto 0,4953 lehčích šroubů; k tomu přistoupí ještě celá levá polovina, která obsahuje šrouby lehčí než střední hodnota se svými 0,5. Je tedy celkem 0,9953 šroubů lehčích a zbytek, 0,0047, těžších než 1,4 g." Smysl této odpovědi je tedy stejný jako předchozí: asi 5 promile překračuje danou váhu.



Vpravo od 2,6 je ještě jen maličký zlomek celkové plochy pod normální křivkou. Proto zde ukazujeme pouze horní hraniční plochu vymezenou křivkou. Celá zvonovitá křivka od —3,3 do + 3,3 by byla asi pětkrát širší než zobrazený úsek.



Pro hodnotu $z = 2{,}6$ je možno vyčíst z tabulek různé hodnoty. Graf ukazuje, jak se doplňují na celek pod normální křivkou. Hodnota 0,991 se většinou nazývá $D(z)$, 0,9953 je $\Phi(z)$ a 0,0057 $\Phi(-z)$.

Jak však uvedeme do souladu 0,9953 šroubů lehčích než 1,4 g s číslem 0,991, které je podobné, ale ne stejné? Vzpomínáme si, že toto číslo zahrnuje plochu na obě strany od střední souřadnice — vynechává však extrémní hodnoty na obě strany! Jestliže však je 0,0047 na jedné straně pro těžší šrouby, je 0,0047 na druhé straně pro šrouby lehčí; dvakrát 0,0047 dává 0,0094, zaokrouhleno 0,009, a to se nyní doplňuje s 0,991 opět na celkovou plochu pod křivkou. Zde jsme tedy vlastně dostali odpověď: „99,1 % výrobků se neodchyluje o více než 2,6 σ od střední hodnoty, zbývajících 0,9 % (přesně 9,4 %) spadá za hranici tolerance, a to samozřejmě symetricky, vždy polovinou na obou koncích křivky."

Tolik o tom, jak zacházet s tabulkami pravděpodobnosti. V případě pochybnosti nikdy neuškodí, když si pozorně přečteme, co je napsáno v záhlaví tabulky (někdy je připojen i vysvětlující nákres normální křivky).

S normálním rozdělením souvisí také *kvartily, decily a percentily*, které se někdy vyskytují v odborné literatuře. Ještě blíže jsou však spřízněny s centrální hodnotou, mediánem. Stejně jako tato hodnota — vzpomínáte si? — udávají, že polovina měrných hodnot leží nahoře a polovina dole, takže *kvartily* tvoří dělítko mezi čtvrtinami, *decily* mezi desetinami a *percentily* mezi setinami. Ale pozor: protože tyto „-ily" jsou jen myšlená dělítka, „neobsahují" vůbec nic. Nejsou nejvyšší čtvrtinou nebo poslední desetinou plochy, nýbrž čarou, která odděluje čtvrtinu či desetinu. Jsou proto jen tři a nikoliv čtyři kvartily, jen 9 decilů a jen 99 percentilů. Výrazy jako „v devátém centilu" se sice používají, ale jsou nesprávné.

3.4 Jiná rozdělení

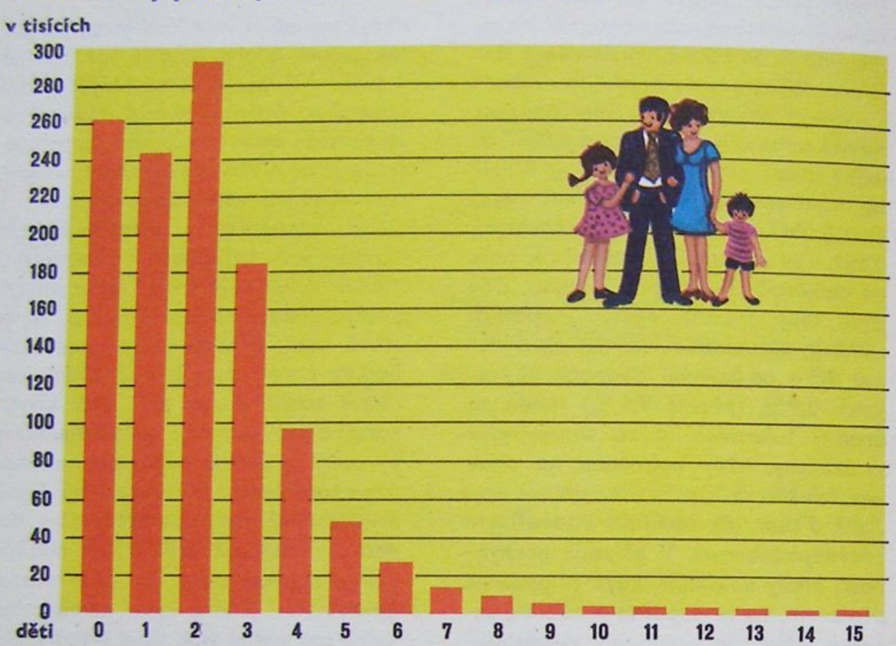
I když normální rozdělení je na pohled tak pěkné a pro použití tak účelné a může být proto často plným (nebo alespoň značným) právem použito, přesto nemůže zachytit všechny skutečnosti. Je mnoho věcí mezi nebem a zemí, o nichž se normální křivce ani nesní.

Velmi jednoduchý a názorný příklad vzniká, je-li nejčetnější hodnota poměrně blízko nuly, avšak nelze přitom předpokládat záporné výsledky: to se stává např. u počtu dětí v rodinách. Rodiny bezdětné, s 1 a 2 dětmi jsou téměř stejně četné; pak však křivka padá na pravou stranu způsobem, který by mohl přibližně odpovídat normálnímu rozdělení: 3 děti jsou méně časté, 4 ještě méně atd., až v úseku 11 a více dětí se dosahuje extrémních hodnot. Co je naproti tomu *na „levé" straně křivky*? Nic. Žádná žena nemůže mít méně než žádné dítě. Graf švýcarského statistického úřadu to ukazuje velmi názorně (obr. na str. 88).

Normální rozdělení se v tomto případě nehodí, protože naráží na nulové hranici na bariéru, která je — obrazně řečeno — zešikmuje. „*Šikmost*" rozdělení je ostatně statistický vědecký termín, kterým se měří horizontální odchylka rozdělení od normálního rozdělení. Doplněk k tomu tvoří „*exces*", jenž křiví zvonovitý tvar normální křivky ve svislém směru, to znamená, že zvon je příliš strmý nebo příliš plochý.

Je-li šikmost tak výrazná, že křivka vykazuje při nejnižších hodnotách nejvyšší četnosti (bylo by tomu tak v případě, že bezdětná manželství by byla nejčastější), vzniká takzvaná *křivka L.* O rozdělení L, které se doprava zplošťuje, koluje melancholické rčení, že prý

Vdané ženy podle počtu živě narozených dětí



Výrazně „strmé“ rozdělení ukazuje mimo jiné počet dětí na vdanou ženu (nebo také na rodinu): značné četnosti při 0, 1 a 2 dětech jsou zhruba stejně vysoké a potom se silně snižují.

je charakteristickým rozdělením všech krásných věcí. Téměř každé rozdělení příjmů ukazuje — udává-li se výše příjmů na souřadnici a počet jejich příjemců na pořadnici — řídkost velkých příjmů a četnost příjmů malých. Chce-li někdo např. považovat vysoký věk za žádoucí příjemnost, udává grafické znázornění úmrtnosti rovněž klesající křivku.

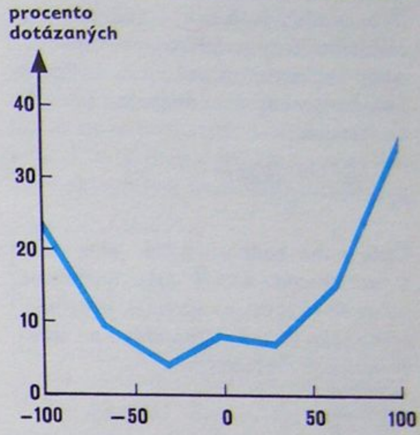
Cokoli je zvlášť cenné, vyskytuje se jen u mála osob. Kdyby se měl graficky znázornit počet barokních soch, případně počet knih připadajících na jednu domácnost, zase by vzniklo rozdělení L, rozdělení krásných věcí.

Výrazné křivky J jsou vzácnější, protože neomezená možnost rozšíření vpravo vede přece jen dříve nebo později ke zploštění. Mohou vzniknout pouze tam, kde je také vpravo pevná hranice, která vlevo je většinou na nule.

Konečně existuje *kombinace křivky J a křivky L; výsledkem je dvouvrcholové rozdělení U*, pro které mimo jiné uvádí Haseloff charakteristický příklad: 326 pokusných osob je dotázáno, za jak pravděpodobné pokládají, že lidé někdy navštíví vzdálené hvězdy Mléčné dráhy. 34 % dotázaných odpovědělo „určitě“, 15 % „pravděpodobně“, 7 % „možná“, 8 % „nevím“, 4 % „asi ne“, 9 % „jistě nikdy“ a 23 % „to není možné“. Polarizace názorů dává v grafickém znázornění rozdělení U.

3.41 Poissonovo rozdělení

Jednoho rozdělení si však musíme všimnout podrobněji, a to takzvaného *Poissonova rozdělení*, které dostalo název podle francouzského matematika z počátku 19. století. Pokládá se za „*speciální rozdělení*“ *pro nepatrné pravděpodobnosti*, za početní schéma pro řídké události. Již v binomickém rozdělení existují řídké jevy — např. výskyt určitého sledu čísel v ruletě nebo ve Sportce. Pro binomické rozdělení je charakteristické *doplňování jevu a nejevu*. Když $p = 0{,}2$, pak $(1 - p) = q = 0{,}8$; ptáme-li se po počtu piků, platí všechno ostatní jako nepiky. Zkrátka vždy byla nějaká protihodnota k hledané pravděpodobnosti nebo četnosti. Mnohé procesy však pomocí uvedeného binomického schématu vyjádřit nelze. Jestliže je na 10 cm dlouhém obrobku jedno vadné místo — kolik míst není vadných? Narodí-li se v době od 10 do 12 hodin 4 děti — kolik se jich v té době nenarodilo? Jestliže během jedné bouřky jsme napočítali 10 blesků, kolik blesků nebylo? Nejzřetelněji, i když velmi podivně, je povaha problému charakteristického pro Poissonovo rozdělení vyjádřena v tabulce, kterou sestavil německý statistik Bortkiewicz na počátku našeho století. Zjišťoval po deset let ve dvaceti německých armádních sborech (podle jiné verze: v deseti armádních sborech po dvacet let) počet vojenských osob zabitých úderem koňského kopyta a zjistil, že u 109 ukazatelů „armádní sbor/rok“ nebyl žádný smrtelný úraz, 65krát byl jeden mrtvý, 22krát dva mrtví, 3krát tři mrtví, jedenkrát dokonce čtyři. Celkový počet smrtelných případů byl $122 = (65 + 44 + 9 + 4)$, „očekávaná hodnota“ smrtelných případů na rok a armádní sbor činí tedy 0,61.



Budou lidé cestovat až ke hvězdám Mléčné dráhy? Odpověď na tuto otázku poskytlo „U - rozdělení“. K vytvoření oceňovací stupnice byla odpověď „určitě“ dosazena jako + 100, „pravděpodobně“ jako + 75, odpověď „není možné“ jako —100. Mezní hodnoty měly nejvyšší podíl: 34 % dotázaných odpovědělo „určitě“, 23 % „to není možné“.

Vyjdeme-li z této očekávané hodnoty, můžeme porovnat, zda tyto nehody mohou být přibližně správně zachyceny Poissonovým rozdělením. Nejdříve prokážeme pomocí vzorce, že tomu tak skutečně je, a teprve potom pojednáme blíže o struktuře Poissonova rozdělení.

$$\boxed{p_x = \frac{e^{-\lambda} \cdot \lambda^x}{x\,!}}$$

Přitom platí: λ = očekávaná hodnota, x = počet hledaných jevů.

Očekávanou hodnotu λ známe: 0,61. Hledejme tedy nejdříve pravděpodobnost, že jev (smrt po úderu koňským kopytem) nastane nulkrát ($x = 0$). Dostaneme:

$$p_0 = \frac{e^{-0,61} \cdot (0{,}61)^0}{0\,!}$$

Číslo e má hodnotu 2,718, jeho vznik si načrtneme. Každé číslo umocněné nulou dává 1; 0!, jak již víme, je definicí určen jako 0! = 1. Vzpomínka na středoškolskou matematiku:

$$a^{-c} = \frac{1}{a^c},$$

to je v našem případě:

$$e^{-0,61} = \frac{1}{e^{0,61}}\,;\ e^{0,61} = 1{,}841, \text{ tedy}$$

$$e^{-0,61} = \frac{1}{1{,}841} = 0{,}543; \text{ to je souhrnně}$$

$$p_0 = \frac{0{,}543 \cdot 1}{1} = 0{,}543.$$

Z celého pole pravděpodobnosti $p = 1$ je tedy více než polovina zaplněna nulovými jevy. Podle Poissonova rozdělení by měl jev „žádný mrtvý úderem kopyta" nastat v 543 z 1000 ukazatelů „armádní sbor/rok", respektive ve 109 z 200 pozorování.

Další krok: $x = 1$, jak často nastane případ úmrtí?

$$p_1 = \frac{e^{-0,61} \cdot (0{,}61)^1}{1\,!} = \frac{0{,}543 \cdot 0{,}61}{1} = $$
$$= 0{,}331$$

Odpověď: v 331 z 1000 případů, tj. asi 66 z 200.

Otázka po 2 mrtvých ($x = 2$):

$$p_2 = \frac{e^{-0,61} \cdot (0{,}61)^2}{2\,!} = \frac{0{,}543 \cdot 0{,}372}{2} =$$
$$= 0{,}101$$

Odpověď: ve 101 z 1000 případů, tedy 20 z 200.

Tímto způsobem dostaneme konečně $p_3 = 0{,}021$ (4 případy ze 200); $p_4 = 0{,}003$ (0,6 případu ze 200).

Porovnejme nyní skutečné hodnoty a čísla vypočítaná podle Poissonova rozdělení:

	Skutečnost	Propočet
0 mrtvý	109	109
1 mrtvý	65	66
2 mrtví	22	20
3 mrtví	3	4
4 mrtví	1	0,6

Jen stěží si lze představit dokonalejší důkaz, že Poissonovo rozdělení výstižně zobrazuje stav věcí.

Dříve než podáme další důkazy o schopnostech Poissonova rozdělení, podívejme se krátce na pozoruhodné číslo e, které je také známo jako *„základ přirozených logaritmů"*. Toto e je souhrnem nekonečné řady a zároveň „mezní hodnotou". Totiž e lze zároveň pokládat za extrémní výsledek výpočtu úroků z úroků. Částka, např. 1000 korun, se má ročně zúročit 100 %. To by při placení na konci roku činilo 2000 korun. Jestliže se však úroky připisují půlročně, úročí se od začátku července nikoliv 1000, ale 1500 korun a na konci roku je tedy v bance 2550 korun. Jestliže se úrok počítá čtvrtletně, úročí se od 1. dubna 1250 korun, od 1. července a 1. října odpovídající vyšší částky, a na konci roku je to již 2440 korun.

Výpočet úroků z úroků lze teoreticky stále více zužovat: měsíčně, týdně, denně, hodinově. Částka připadající na konec roku tak bude stále vzrůstat, nikoli však donekonečna, nýbrž ke konci ve stále menších krocích tak, jak se budeme blížit k hranici — více než 2718 korun 28 haléřů to v žádném případě nemůže být: *našich původních 1000 korun násobených číslem e.*

Číslo e lze také vyjádřit jiným způsobem, a sice touto exponenciální řadou:

$$e = \frac{1}{0\,!} + \frac{1}{1\,!} + \frac{1}{2\,!} + \frac{1}{3\,!} + \frac{1}{4\,!} + \frac{1}{5\,!} + \ldots$$
$$= 2{,}7183\ldots$$

Při této příležitosti se znovu setkáváme s čísly s vykřičníkem (vyslov: „faktoriál"), které jsme poznali u binomického rozdělení a o nichž už víme: 0! = 1; 1! = 1; 2! = 2.1 = 2; 3! = = 3.2.1 = 6; 4! = 4.3.2.1 = 24 atd. Na tomto místě nechceme popisovat matematické odvození Poissonova rozdělení z binomického, ale pouze poukázat na základní myšlenky. Rozložíme našich 200 ukazatelů „armádní sbor/rok" na 200 . 365 = 73 000 jednotlivých dnů (nebo zhruba 1,8 mil. hodin atd.), takže v takovém časovém úseku nastane nejvýše jeden jev. Máme tu jakousi „karetní hru" složenou ze 73 000 dnů, ve které je 122 dní se smrtí úderem kopyta. Z takové karetní hry pak mohou být zásadně zjištěny pravděpodobnosti tahu jako při tahu 1 piku z normální karetní hry, ovšem s mnohem vyššími čísly.

Takové téměř nekonečné další dělení časového úseku je vyjádřeno v čísle e, k jehož vysvětlení jsme rovněž použili nekonečného postupného členění časového úseku.

Pro naše úvahy v této knize není ani tak důležitá souvislost Poissonova rozdělení s binomickým, jako spíše skutečnost, že — na rozdíl od normálního rozdělení — je určováno jedinou proměnnou λ. Zatímco pro stanovení normálního rozdělení potřebujeme jak střední hodnotu, tak rozptyl, v tomto případě postačuje zjištěná průměrná četnost jevů, pravděpodobnost zásahů, která vstupuje do vzorce.

Tato skutečnost umožňuje při použití Poissonova rozdělení další zjednodušení. Protože prakticky záleží jen na mé libovůli, na kterou základní veličinu chci četnost jevů vztáhnout, mohu dosadit $\lambda = 1$ (nebo přibližně rovno 1), z čehož vyplývá zjednodušený vzorec:

$$\boxed{p_x = \frac{1}{e} \cdot \frac{1}{x\,!}}$$

Při tomto zjednodušení vyjdou však někdy pozoruhodné poměrné veličiny. V případě obětí úderem kopyta by bylo nutno při $\lambda = 1$ zkoumat četnost smrtelných případů pro armádní sbor v průběhu 19 měsíců. Pak by vyšlo:

$$p_0 = \frac{1}{e} = 0{,}368$$

Pro jeden smrtelný případ bychom rovněž dostali stejnou pravděpodobnost:

$$p_1 = \frac{1}{e} = 0{,}368$$

Prodloužením časového úseku se totiž zvýšila pravděpodobnost pro více než jednu takovou nehodu:

$$p_2 = \frac{1}{e} \cdot \frac{1}{2\,!} = 0{,}184$$

$$p_3 = \frac{1}{e} \cdot \frac{1}{3\,!} = \frac{1}{6e} = 0{,}061$$

A nyní pokračuje zajímavá řada: p_4 je čtvrtinou výsledku p_3; p_5 je pětina p_4 atd., jak už bylo p_2 polovinou p_1 a p_3 třetinou p_2.

Předpokladem pro použitelnost Poissonova rozdělení je vždy závislost náhody na λ. Počet předmětů nalezených a odevzdaných denně v obchodním domě např. zřetelně ukazuje Poissonovo rozdělení. Typickými příklady jsou dále počet telefonních hovorů v určitém

časovém úseku (srovnej k tomu později „teorii front" v osmé kapitole) nebo počet úrazů na dělníka za rok. Mnohý problém binomického rozdělení lze podstatně rychleji vyřešit Poissonovým vzorcem — např. otázku po jednom piku při 5 tazích, kdy je ovšem třeba velké opatrnosti, protože pravděpodobnost $p = 0,25$ je vlastně příliš velká.
Bude však docela zajímavé, jestliže to zkusíme a zjistíme rozdíl. Očekávaná hodnota je — jak už víme — 1,25 a tu nyní označíme λ. A hned můžeme dosazovat do vzorce:

$$p_{1\,pik} = \frac{e^{-1,25}\cdot(1,25)^1}{1} = \frac{1}{e^{1,25}}\cdot 1,25 = $$
$$= \frac{1,25}{3,5} = 0,36$$

Při tomto výpočtu dostáváme tedy $p = 0,36$ proti $p = 0,40$ při binomickém rozdělení. Ostatně jsme varovali, že v tomto případě není Poissonovo rozdělení nejvhodnější, protože nejde o „řídký" jev. K tomu však přistupuje ještě něco jiného: Poissonovo rozdělení mlčky předpokládá, že n je velké, a proto se toto n ve vzorci vůbec nevyskytuje. Početářsky je tedy možno tímto vzorcem vypočítat pravděpodobnost tahu *šesti* piků při *pěti* tazích:

$$p_{6\,pik} = \frac{1}{e^{1,25}}\cdot\frac{(1,25)^6}{6!}$$

Dostáváme jen asi tisícinu, a to je přece zcela nemožné! Nedokazuje to však, že Poissonovo rozdělení je nesmyslné, nýbrž jen, že musíme vážit a vědět, kdy a jak se má použít. A znovu to dokazuje, že přesnými početními postupy lze dospět k velmi působivým číslům, která však při své kráse mají jen tu malou chybu, že nemají žádný význam. *Statistika bez použití rozumu dává nesmysly* — a to neplatí jen o statistice.

Teď ještě několik slov o zvláštní formě rozdělení a jeho zobrazení — o součtových křivkách.

3.5 Součtové a koncentrační křivky

V tabulkách k normálnímu rozdělení je ke každé hodnotě na souřadnici (osa x) uveden odpovídající podíl plochy pod normální křivkou — a to, jak už víme, podle tabulky nejdřív podíl od záporného konce až po x, pak podíl od x až ke kladnému konci křivky — a vše ostatní, co jsme se o této hře naučili. Celková plocha pod křivkou je tedy dohodnuta na 1, a všechny její díly se udávají v promile nebo v částkách menších než 1. Na vrcholu křivky se dosahuje přesně podílu 0,5.
Na základě těchto tabulek můžeme dát normální křivce zcela jinou podobu tím, že ji změníme na součtovou křivku. Jednice 1 se dosadí za rovných 100 %, což je podle smyslu totéž, a vynese se na pořadnici; proti obvyklému znázornění normální křivky se abscisa vůbec nemění a sahá asi od -3σ až po asi $+3\sigma$. Nyní vyhledáme pro některé charakteristické hodnoty x v tabulkách podíl ploch a naneseme je na osu y. K tomu účelu nejdříve použijeme tabulky, která udává pro průměr $\mu = 0$ hodnotu 0,5.

Pak např. dostaneme:

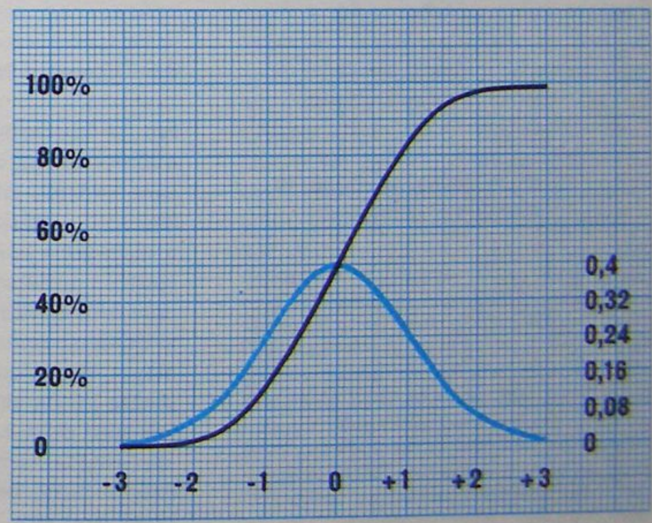
	podíl plochy	naneseme na osu y v mm
$x = 3$	0,0013	0,1
$x = 2,5$	0,0062	0,6
$x = 2$	0,0228	2,2
$x = 1,5$	0,0668	6,6
$x = 1$	0,1587	15,8
$x = 0,5$	0,3085	30,8
$x = 0$	0,5	50,0

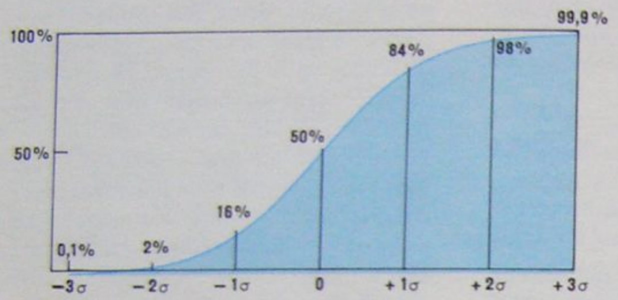
Na pravé polovině křivky odečteme „levé" hodnoty od celku 1 a tak dostaneme:

$x = 0,5$	0,6915	69,2
$x = 1$	0,8413	84,1
$x = 1,5$	0,9332	93,3
$x = 2$	0,9772	97,7
$x = 2,5$	0,9938	99,4
$x = 3$	0,9987	99,9

Při pečlivém provedení na milimetrovém papíře může taková křivka v určité míře dokonce nahradit tabulku. Tam, kde 25%, resp. 75% čára protíná křivku, jsou na ose x hodnoty kvartilů, přibližně +, resp. — 0,675 σ.
Součtová křivka je především užitečná tam, kde se chceme na první pohled přesvědčit, jak se určitá naměřená hodnota umístila v celkovém obraze četností. (K tomu je ovšem účelnější nanést na osu x hrubé měrné hodnoty po výpočtu standardních odchylek, a nikoliv standardizované hodnoty —3, —2 atd.)
Naneseme-li uvedeným způsobem z jednoho našeho dřívějšího případu výsledky měření kvocientu inteligence ($\mu = 112$, $\sigma = 6$), okamžitě vidíme, kolik procent pokusných osob má IQ pod 100, pod 110, nad 120 atd.
Je však ještě jiný druh součtové křivky, která by se vlastně měla jmenovat „*součtová přímka*". Nemá nic společného s normálním rozdělením ve statistickém slova smyslu, ale velmi mnoho s normálním rozdělením ve smyslu přesného rozdělení spravedlnosti. Když totiž každý má stejně — ať už peněz, pozemků nebo radiopřijímačů — pak v případě každé nově zahrnuté osoby rovnoměrně stoupne celkový počet peněz, pozemků i radiopřijímačů. Musí však přitom jít o částky nebo věci, které se mohou sečítat. Nemohu např. říci: A má kvocient inteligence 100, B a C také, tzn. že všichni tři mají tedy IQ 300.
Protože však pozemské statky nejsou skoro nikdy tak rovnoměrně rozděleny,

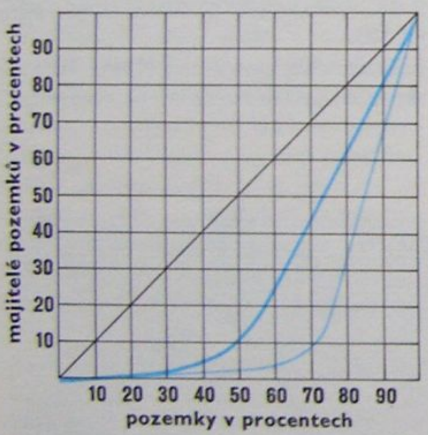
Ze součtové křivky normálního rozdělení (černá křivka, levá stupnice) lze vyčíst, kolik procent plochy pod normální křivkou (modrá křivka) „se spotřebuje" až po každý daný bod standardní odchylky. K — 1 je to např. 16 %, k + 1 asi 84 %. Čísla stupnice vpravo udávají relativní hustotu normálního rozdělení: ať už normální křivka bude strmá nebo plochá, tyto číselné poměry musí být v každém případě dodrženy. Tak je křivka např. v bodě + 1,2 a — 1,2 vždy poloviční ve vztahu k vrcholu nad bodem 0.





Zobrazení samotné součtové křivky. Změnou měřítka vypadá proti předchozímu vyobrazení plošší. Ale také normální křivka by byla relativně plošší.

nevzniká zpravidla žádná přímka, ale pozoruhodná, propadlá křivka, kterou lze nejsnáze porovnat s rozdělením J. (V tomto případě může vzniknout pravá křivka J, protože na pravé straně se po dosažení 100 % táhne uzavírací čára.) Tyto křivky měří „koncentraci“ peněz, pozemků atd., a proto se nazývají *„koncentrační křivky“*; nejznámější je tzv. *Lorenzova křivka.* Na příkladu z egyptské statistiky lze zcela jasně vidět princip takové Lorenzovy křivky. Na souřadnici je procentní sazba pozemkového vlastnictví, na pořadnici procentní sazba majitelů pozemků. Obě křivky udávají hodnoty z let 1952, resp. 1965, a výrazně ukazují úspěch pozemkové reformy v Egyptě. Na první pohled snadno zjišťujeme, že v roce 1952 10 % majitelů pozemků vlastnilo 70 % půdy (v roce 1965 jim patřilo již jen 50 %).

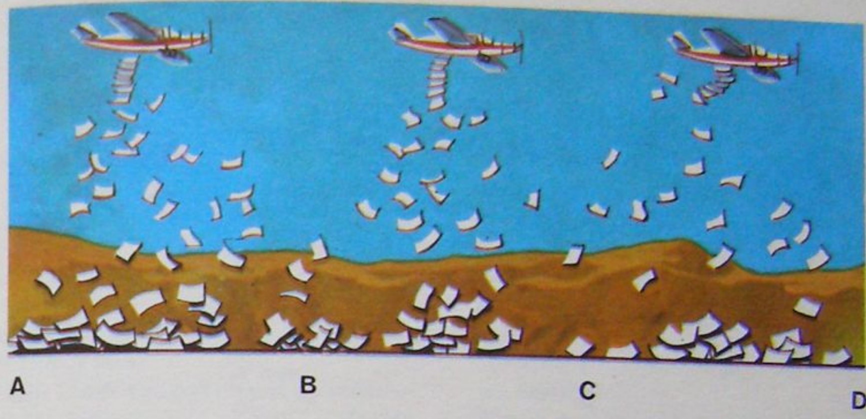


Kdyby každý majitel pozemků vlastnil stejný rozsah půdy, vyjadřovala by diagonála toto „ideální rozdělení“. Egyptská statistika umožňuje ukázat jak krajně nespravedlivé rozdělení půdy v roce 1952 (pravá křivka — 10 % majitelů pozemků vlastnilo 70 % půdy), tak důsledky opatření zemědělské reformy uskutečněné do roku 1965 (střední křivka), které přece jen podstatně přiblížilo skutečnost k ideálu. Tento způsob zobrazení se nazývá koncentrační nebo Lorenzova křivka.

3.6 Četnost a hustota

Na začátku tohoto oddílu jsme mluvili o diskrétních a spojitých znacích a pak jsme ukázali, jak z diskrétního rozdělení při rostoucím počtu pozorování a možných výsledků ponenáhlu vzniká spojité rozdělení, které lze vyjádřit normální křivkou.

K určitému bodu normálního rozdělení nelze již přiřadit žádnou četnost, je možno již jen určit *podíly četnosti*, které leží mezi dvěma body. Je to zcela logické, protože mezi $+3\sigma$ a -3σ je nekonečně mnoho bodů. V každém z těchto bodů jsou sice 2 křivky a (z tabulek)

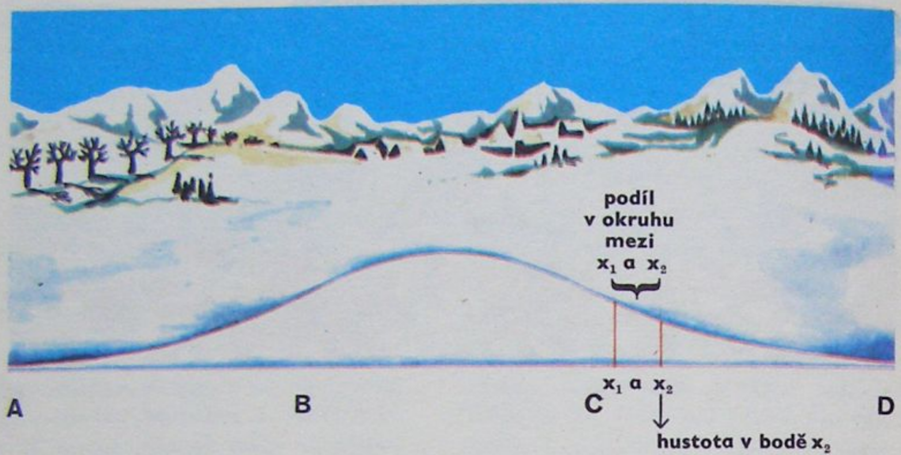


Letáky shozené nad krajinou mohou být na některých místech i hustě na sobě, lze je však vždy spočítat.

zřejmá „hustota“, ale již *žádná četnost.* Rozdíl lze objasnit na následujícím příkladu. Jestliže letadlo shazuje bomby, je možno později říci: nad osadou A spadlo 20 bomb, mezi A a B 48, nad B 65, mezi B a C 30, nad C 22, mezi C a D 8 a nad D 2. Pro získání porovnávací základny nezávislé na velikosti osad by bylo možno shození bomb určovat také na čtvereční kilometr. Zřejmě jde o jednoznačně diskrétní (chceme-li mluvit v souvislosti s bombami o „diskrétním“) rozdělení spočitatelných četností.

Jestliže letadla — naštěstí — místo bomb shazují letáky, je možno zjistit ještě toto: nad A bylo shozeno 360, mezi A a B 500, nad B 1200 letáků atd. Lze si však také představit krajinu souvisle pokrytou letáky: v tomto případě leží rozsypané jednotlivě, někde ve velkém počtu, dál od sebe nebo na sobě — rozdělení přechází z *„diskrétního“* do *„spojitého“, jež je zjistitelné jen odhadem.*

Pokročme ještě dál do mírových podmínek. Teď již nepadají letáky, ale kroupy. Mezi B a C leží na polích v centimetrové vrstvě, neboť zde byly srážky nejprudší; odtud na obě strany se četnost krup (což se nyní nazývá „hustota vrstvy krup“) zmenšuje. Teoreticky bych mohl jít ještě na určitý lán pole mezi osadami B a C a říci: na čtverečním decimetru, který tvoří nejzazší severozápadní kousek řepného pole, leží přesně 598 krup, na čtverečním decimetru vlevo vedle dokonce 634. Mohl bych v této pošetilé hře pokračovat až podle čtverečních centimetrů pozemku. Ale teď ať sněží — a skončeme s diskrétním rozdělením, vítězí hustota. V každém jednotlivém bodě krajiny mohu bez námahy zjistit: zde leží sníh 5 cm, 18 cm nebo 38 cm vysoko. Ale v každém bodě nemohu říci: nad touto špendlíkovou hlavičkou leží přesně 325 852 sněhových vloček. Mohu však říci: mezi tímto bodem a bodem druhým leží tisícina nebo osmina nebo sedma-



Sněhové vločky, které spadnou nad krajinou, nelze v žádném bodě už spočítat, je možno pouze změřit výšku (hustotu) sněhové pokrývky. Podíl na délkovém rozdělení sněhu lze určit pouze v okruhu mezi dvěma místy měření.

třicetina celkového množství sněhu. V dalším průběhu našeho pojednání se pojmu „hustota" úplně vzdáme. Protože však normální rozdělení, přísně matematicky řečeno, je funkcí hustoty a protože čtenář možná někdy v nějaké učebnici na pojem „hustota" narazí, chtěli jsme závěrem této kapitoly ještě ukázat, jak se *četnost mění v hustotu* a proč nelze nikdy v jednotlivém bodě normálního rozdělení zjistit podílovou četnost, nýbrž vždy jen četnost *mezi* dvěma body tím, že se vyříznou více nebo méně úzké díly z celkové četnosti.

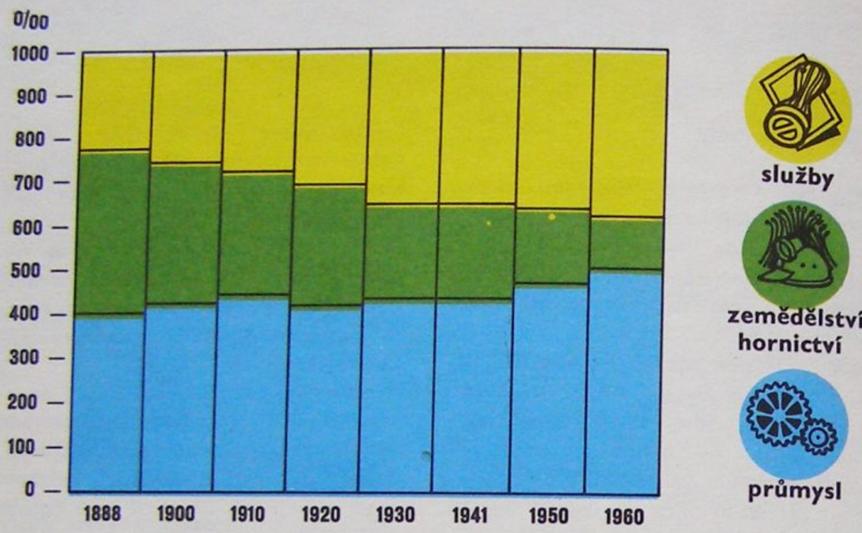
4 Body, procenta, indexy

4.1 Časové řady

„Prouděním života, vichrem dění, / vzhůru a dolů a sem a tam / vám a tkám / zrození — hrob — / věčný oceán, / ohnivé vání, / kolotání a proměňování./ Tak času jsem tkadlec, tak hrčí můj stav."
Další řádek pak pokračuje: „a tkám — tkám božstva živoucí háv" (překlad O. Fischera podle vydání SNKLHU, Praha 1955) — málem bychom se začali domnívat, že Duch z *„Fausta"* začne zpívat hymnus na časové řady, které jsou opravdovým tkaním hávů životních procesů, neboť graficky a v tabulkách zachycují veškeré kolotání a proměňování mezi zrozením a hrobem. Jsou to především časové řady, s nimiž se občan setkává často jak v denním tisku, tak i v odborných časopisech, více nebo méně originální — s podivnými figurkami, které šplhají nahoru nebo zase zděšeně sviští po nějaké křivce do hlubin. Časová řada je také to, co bývá v kreslených vtipech na pozadí ředitelských kanceláří a zasedacích síní — většinou prudce klesá a některý z přítomných, zpravidla zachmuřených pánů k tomu chce podotknout něco velmi kousavého.
Jestliže tento druh statistiky tak „nadprůměrně" vystupuje do popředí, není to samozřejmě bez důvodu. *Časová řada vnáší totiž do statistiky dynamický prvek.* Nejenže umožňuje poznat vzorek utkaný na tkalcovském stavu času, ale přímo vyzývá pozorovatele, aby s pohledem do budoucnosti kreslil nové vzory.

I když časové řady nemají u pozorovatele vyvolat futurologické vize plné fantazie, upoutávají však více než poměrná čísla nebo nehybná rozdělení četností, protože uvádějí *do hry dimenzi času*, ukazují několika čarami nebo čísly vývoj, nepředpokládaný nebo skličující, který jsme většinou zpravidla jen nejasně tušili.
Přesto není zásadní rozdíl mezi časovou řadou a jednotlivými statistickými výběry nebo vyčerpávajícím šetřením. Stejně jako se film skládá z jednotlivých nehybných obrázků, je i časová řada složena z takových jednotlivých snímků, a chceme-li zůstat u filmového obrázku a pomyslíme na rozsáhlá časová období, která jsou v časových řadách překlenuta, můžeme bezmála mluvit o metodě časové nenasytnosti.
Časové řady v zásadě vytvářejí spojení mezi stejnorodými údaji (zjištěními, výpověďmi) z různých dob, avšak stejného věcného obsahu. Může jít nejen o plynulá porovnávání, jako je tomu v případě sledování ročních dovozů a vývozů za posledních patnáct let (str. 101), ale i o porovnání jednotlivých vybraných údajů, jako např. porovnání struktury povolání ve Švýcarsku v letech 1888, 1920, 1940 a 1960 (str. 98). Plynulá porovnání nejsou však často vůbec možná. Časová řada z výsledků hromadných sčítání lidu nejenže přeskakuje desetiletá období, nýbrž poskytuje mimoto jen bodové údaje rozhodného okamžiku — zjištěné hodnoty, které přísně vzato platily pouze jediný

Zaměstnaní podle hospodářských odvětví od roku 1888



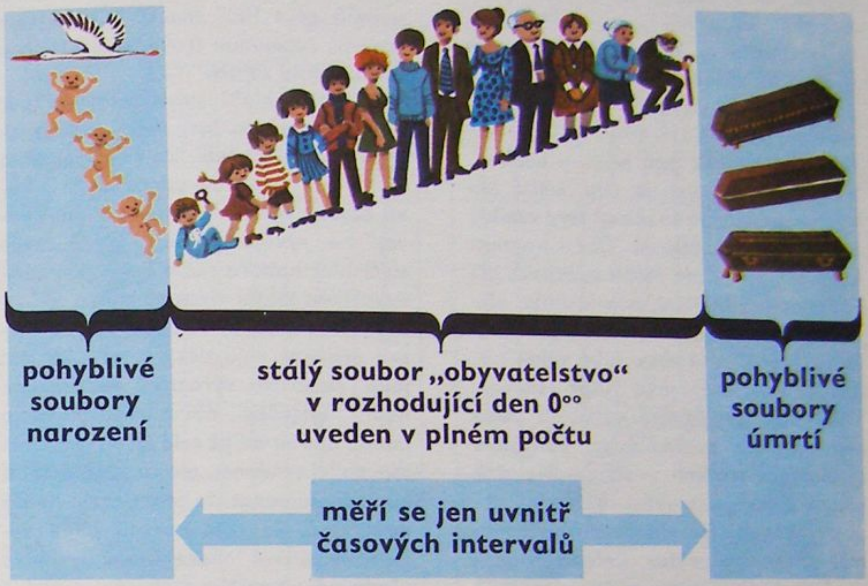
Grafické zobrazení časových řad ulehčuje poznání vývoje, jako v tomto příkladu ze švýcarské statistiky. Procentní podíl zaměstnaných v „sekundárním sektoru" (průmysl, zpracovatelské závody) od roku 1888 poněkud vzrostl — s nepatrným kolísáním, zatímco „v primárním sektoru" (prvovýroba) lze pozorovat stálý pokles a „v terciárním sektoru" (služby apod.) plynulý vzestup. Při tomto způsobu srovnání časových řad se ukazují relativní četnosti, tedy nikoliv absolutní hodnoty — sloupky pro každý rok jsou stejně vysoké („1000 promile"). To znamená, že není vidět, zda bylo v roce 1960 více nebo méně zaměstnaných než v roce 1930 nebo 1900. To by se dalo vyjádřit různou výškou sloupků, pak by však byl obtížně vidět trend struktury zaměstnanosti.

den. Tím jsme se dostali k rozdílu, který si laik zřídkakdy uvědomuje: *jsou soubory stálé a soubory pohyblivé.* Stálé soubory (nazývané také *prostorové* soubory) jsou soubory, jejichž jednotky mají určitou dobu trvání.

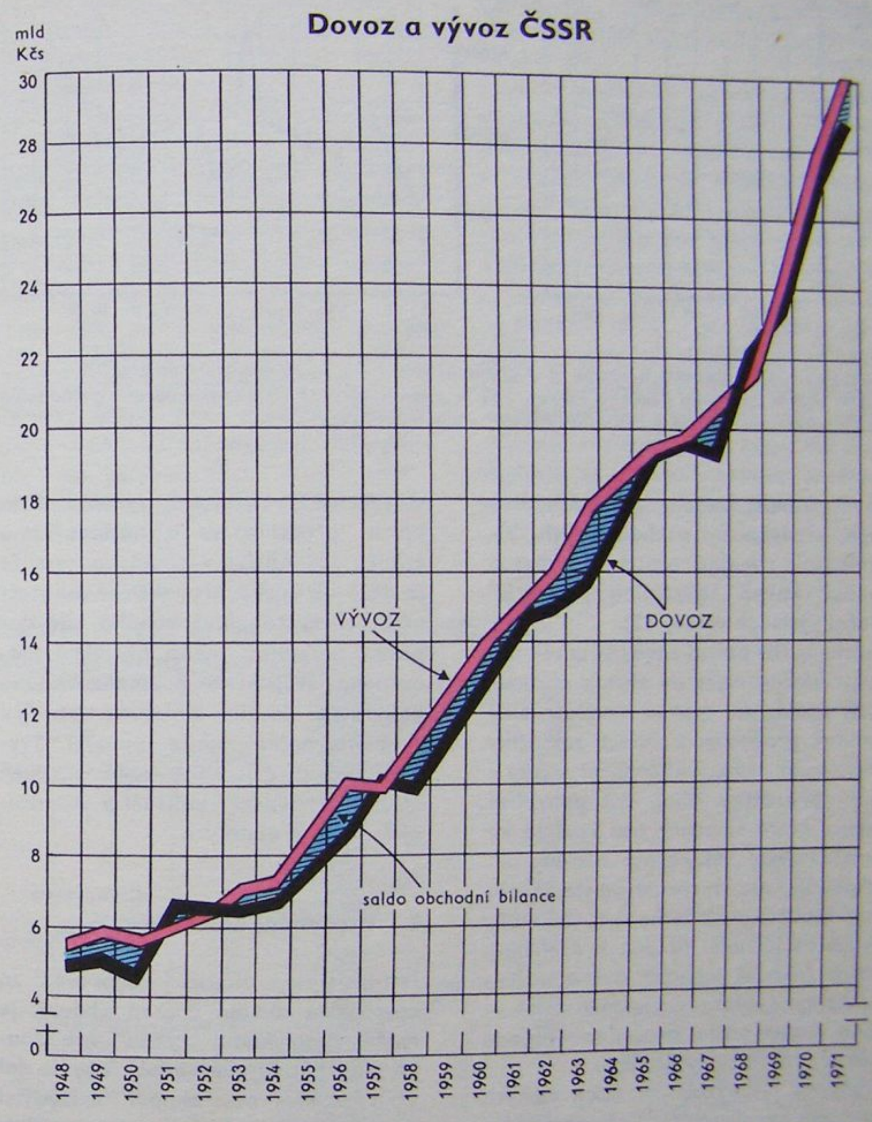
Nejjednodušším příkladem je obyvatelstvo nějaké země — jednotlivci se rodí a umírají: celek obyvatelstva je tím z dlouhodobého hlediska dotčen jen tehdy, když trvá zřejmá převaha narození nebo úmrtí. Každá jednotka

Obr. vpravo nahoře: Stálé soubory mohou být zjišťovány (alespoň teoreticky) ke zcela určitému časovému bodu — jako třeba obyvatelstvo země v den sčítání lidu. Pohyblivé soubory (jako narození nebo úmrtí) se zjišťují jen uvnitř časových intervalů („v průběhu včerejšího dne...", „během minulého roku...").

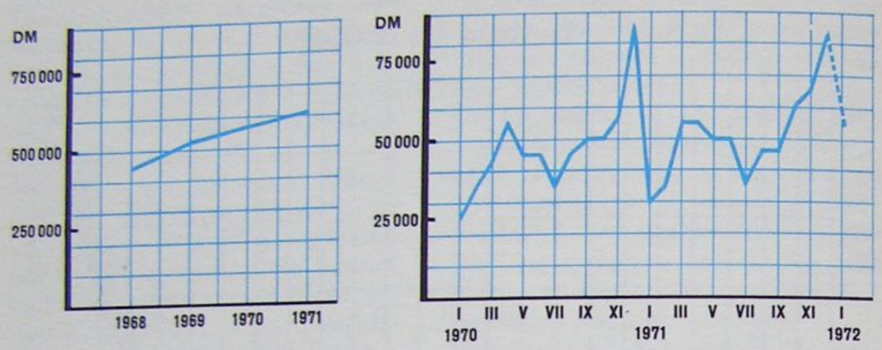
Obr. vpravo dole: Příjezd, pobyt na dovolené a odjezd se podobají narození, životu a smrti. Proto přijíždějící a odjíždějící tvoří „pohyblivé soubory", prázdninoví hosté jsou však „stálým souborem", ačkoliv se v osobní skladbě denně mění.



uvedených souborů má určitou „*dobu setrvání*" — u lidí je to individuální délka života, u klikové hřídele doba uložení ve skladu dílny a u hosta na dovolené doba pobytu v prázdninovém hotelu. (Obr. na str. 99.)
Pohyblivé soubory jsou soubory událostí. Vznikající události se dají měřit jen tím způsobem, že se sčítají jevy vzniklé během daného období. Oběti kopnutí koněm, které jsme mohli politovat při Poissonově rozdělení, jsou krajním případem řídkých jevů — stěží v průměru jeden ročně. Jsou však také velmi pohyblivé soubory, jako např. sňatky. Také oba jevy, které vymezují dobu setrvání, lze pokládat za pohyblivé soubory; narození a smrt, případně příjem do skladu a výdej ze skladu.
Stálé soubory lze zjišťovat *k určitému rozhodnému dni:* v den posledního hromadného sčítání obyvatelstva v ČSSR r. 1970 bylo tolik a tolik mužů a žen, tolik a tolik rodin, nezletilých dětí nebo také televizorů a praček.
Pohyblivé soubory musí být naproti tomu zjišťovány a *srovnávány za určité období.* Počet všech televizorů v soukromých domácnostech může být — alespoň teoreticky — zjištěn k určitému rozhodnému okamžiku: pozítří v poledne úderem 12. hodiny. Tři hodiny před tímto termínem bude asi v zemi o několik přístrojů méně, tři hodiny později naopak o několik přístrojů více, ale v rozhodném okamžiku „dvanácti hodin" bude zjištěn přesný obraz zjišťované situace.
Počítání pohyblivého souboru, např. sňatky ke stejnému rozhodnému okamžiku (pozítří úderem dvanácté), by poskytlo pravděpodobně nulu, ledaže by někde připadla oddávací formule přesně na dvanáctou hodinu. Časový interval „pozítří od 6 hodin ráno do 20 hodin večer" dovoluje naproti tomu uzavřít přes 1000 sňatků a poskytuje alespoň rozumnou srovnávací hodnotu pro „sňatky za den".
Jaký je tedy hlubší smysl časových řad? Jde o zbytečnou hru málo zaměstnaných statistiků nebo za tím vězí něco víc? Ihned se vnucuje odpověď, že taková časová řada ukazuje většinou „vývoj" — vývojovou linii, která může směřovat nahoru nebo dolů, případně vykazovat jakýsi vlnovitý pohyb.
Je samozřejmé, že se takový vývoj nesmí přijímat nekriticky, i když se jeví jako sebevíce výrazný. Porovnejme např. výsledky obratu obchodního domu. Budou asi již celá léta vykazovat stoupající tendenci, ale to ještě nemusí mnoho znamenat. Je přece zcela dobře myslitelné, že růst obratu bude výsledkem plíživé inflace, která vyvolává představu plynulého zvyšování, zatímco „reálný" obrat (při jeho propočtu se přihlíží k poklesu kupní síly měny) stagnuje nebo dokonce mírně klesá. V daném případě by takové zobrazení nebylo „*relevantní*" — nevypovídalo by to, co by vypovídat mělo: neukazovalo by vývoj obratu podniku, nýbrž spíše rozsah inflace. Z těchto důvodů je nutno údaje o obratu „očistit" od tohoto rušivého faktoru.
Avšak takové roční porovnání může nejen nadhodit, ale i zakrýt některé další problémy. Zatímco křivka celkových ročních výsledků směřuje plynule — jak jsme se domnívali — „nahoru", změní se tento obraz velmi rychle, když rok rozdělíme na čtvrtletí nebo dokonce na měsíce: z původně hladké, klidné čáry se stane hektické cikcak. Začnou se totiž projevovat všechny vlivy, které jsou pro obraty obchodních domů zpravidla typické: mdlá kupní nálada na začátku roku, oživení ob-



Časové řady zavádějí do statistického dění dynamiku vzhledem k tomu, že zřetelně ukazují vývoj srovnáváním výsledků měření prováděných v pravidelných intervalech, a současně jako by vyzývaly k dalšímu promítnutí tohoto vývoje do budoucnosti, což ovšem není vždy oprávněné. Typickou časovou řadu ukazuje tento graf ze Statistické ročenky ČSSR: údaje zahraničního obchodu od r. 1948 do r. 1969 jsou postaveny vedle sebe a umožňují srovnání jak vývozu, tak dovozu a přebytku (schodku) obchodní bilance během celého časového období.



Co se jeví v delším časovém intervalu jako klidná křivka, může se po rozdělení do kratších srovnávacích období jevit jako hektické výkyvy. Tak je tomu v případě smyšlených obratů obchodního domu, které jsou rozděleny podle let (vlevo) a podle měsíců (vpravo).

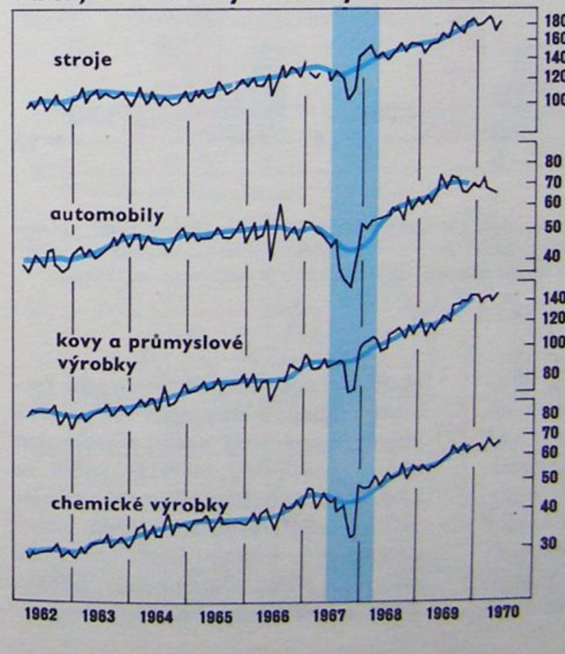
chodu k velikonocům, sezónní výprodeje, předvánoční obchodní ruch. Vidíme zcela zřetelně rozptyl jednotlivých měsíců kolem „měsíčního průměru" (roční výsledek dělený 12).
Svátky a vliv počasí zavádějí prvek nepravidelnosti nejen do obratu obchodních domů, ale mohou vyvolat velký zmatek především v číslech statistiky cestovního ruchu. Velikonoční svátky a jarní prázdniny, obojí tak pohyblivé volno, které se stále více využívá ke krátkodobým rekreačním cestám, nepřipadají každým rokem na stejné měsíce. Statistika cizineckého ruchu může proto od jednoho března k druhému, stejně jako od jednoho dubna k druhému, vykazovat podivuhodné skoky — a to vše jen proto, že uvedené volno je někdy dříve a jindy později.
Dokonce i s výrobními údaji velkých podniků bývají potíže: jiné uspořádání dovolených v letních měsících vnáší nepořádek do údajů za červenec a srpen a takový měsíc, jakým byl např. květen 1970, na který připadl 1. máj, církevní svátky a k tomu ještě na jeho začátek a konec připadly konce týdnů, podstatně sníží absolutní výrobní čísla oproti předchozímu a následujícímu měsíci. Jak nalézt východisko, jestliže se stále ještě čeká na praktické uskutečnění tak často ohlašovaného, ale doposud nerealizovaného projektu jednotného světového normalizovaného kalendáře? Je jím *syntetické vytvoření normalizovaného měsíce* pomocí „vyrovnání" a přirozeně také zároveň „*sezónní vyrovnání*" kolísání v jednotlivých ročních obdobích.

4.2 Vyrovnání kalendáře

Jestliže jsou k dispozici pozorování za dostatečně dlouhá časová období, je možno postřehnout „*cyklus*". Jak dlouhé musí být časové období, aby se dal určitý cyklus postřehnout, nelze říci a priori, to může vyplynout jen ze získaných údajů.
Zřetelný cyklus v průběhu asi 24 hodin vykazuje např. příliv a odliv, který je možno zjistit řadou pozorování v hodinových nebo dvouhodinových intervalech. Kdybychom však měřili stav

vody na pobřeží každý čtvrtek přesně ve 12 hodin, trvalo by asi velmi dlouho, než bychom z těchto měření mohli učinit správný závěr.
Vědecká statistika se na cykly dívá zpravidla skepticky. Na speciálním úseku ekonomické statistiky sice trvá snaha zachovat pojem *konjunkturních cyklů*, ovšem ani v tomto případě nikoliv ve strnulé formě rytmu, který se opakuje po pěti, sedmi nebo dvaceti letech, nýbrž spíše v podobě charakteristického pohybu křivky, který se opakuje v přibližně stejných intervalech.
Celkový vývoj, vlivy ročních období, cykly — co ještě kromě toho může působit na časovou řadu? Nuže např. *jednorázové mimořádné jevy*. Tyto jevy mohou být rozeznatelné již předem, např. revalvace (zhodnocení) nebo devalvace měny.
Je zcela možné, že teprve na základě statistických údajů se zjistí, že — zcela neočekávaně — nastal určitý obrat v myšlení lidí, který vedl k náhlému zvýšení nebo snížení odbytu určitého výrobku. Jestliže jev, který vznikl jednorázově, působí trvale, lze mluvit o jakémsi *zlomu struktury*, který vede ke změně *dalšího vývoje*. Tak např. vynález syntetických vláken postavil textilní průmysl před zcela novou situaci. Jednorázový jev a vývoj jsou tedy někdy v úzkém spojení, cykly mohou mít pochybnou vypovídací hodnotu. Sezónní vlivy nejsou vždy tak jasně prokazatelné,

Export Velké Británie v miliónech angl. liber, sezónně vyrovnaný



Zcela mimořádné a neobvyklé jevy mohou být znázorněny grafem časové řady. Tento příklad je převzat z knihy „Economic Trends" (Hospodářské trendy) britského statistického úřadu: snížení hodnoty libry z podzimu 1967 lze zřetelně rozpoznat ve vývoji vývozu: stupnice na pravé straně je rozdělena pro každou skupinu zboží a má logaritmické členění.

jako je tomu např. při prodeji zmrzliny nebo ve stavebnictví a v cizineckém ruchu. Proto je každý pokus o „vyrovnání" prvotních údajů s cílem získat nezkreslený obraz vývoje provázen nebezpečím, že *dojde jen k zdánlivému vyrovnání, zatímco ve skutečnosti dochází k novému zkreslení*. Kdo začne manipulovat s výsledky měření, vystavuje se nebezpečí, že se v nich velmi lehce utopí. Nejméně škodlivé je, když se určí srovnatelné období. V daném případě se pak často říká: „proti srovnatelnému měsíci předchozího roku...", přičemž základní myšlenkou je, že se rok rozpadá na rozdílná sezónní období, která se pak mohou nejlépe vzájemně porovnat. Tak např. lze účelně srovnat údaje o cizineckém ruchu v září jen s údaji měsíce září, a nikoliv s údaji za předchozí měsíc srpen nebo za následující měsíc říjen. Takové relativně jednoduché porovnání se „srovnatelným měsícem" může ovšem vést k potížím, jestliže právě zvolený srovnatelný měsíc je netypický. Předpokládejme, že v předchozím roce bylo září krásné, teplé a nádherné a následovalo po deštivém a chladném srpnu. V tomto roce je naopak normálně teplý a krásný srpen, po kterém následuje rovněž průměrně mírné, ale přece jen znatelně chladnější září. Co ukáže porovnání se „srovnatelným obdobím předchozího roku"? Letošní údaje za cizinecký ruch v srpnu vyskočí o 30 a více procent a naproti tomu v září se ukáže zdánlivě katastrofální pokles.

Abychom se všeobecně vyhnuli takovým nahodilostem, pak v zájmu opravy časové řady od rušivých, ale nepodstatných vlivů se použije *metody klouzavých průměrů*. Podle příkladu z Pfanzagla zvolme třeba následující řadu měření: 12,3 — 12,9 — 13,6 — 14,4 — 15,3 — 16,3 — 17,2 — 18,0.

To mohou být — podle dané situace — absolutní hodnoty nebo relativní četnosti, může jít o bezprostředně po sobě jdoucí časové body anebo srovnatelná období rozdělená intervaly. Má-li se z této řady vypočítat klouzavý průměr, určí se nejdříve počet činitelů, které chceme brát v úvahu (bude závislý na celkovém počtu disponibilních údajů). Zvolme pětičlenný klouzavý průměr

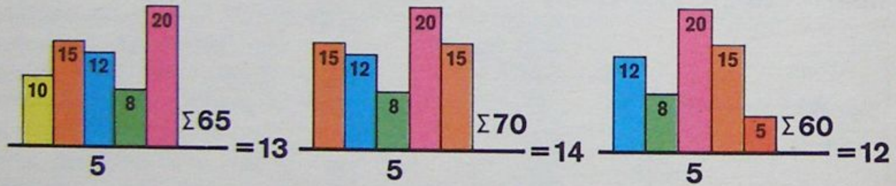


Schéma klouzavého průměru. Vždy se připojuje nejnovější měsíční nebo roční, resp. týdenní výsledek a nejstarší se vylučuje. Vlevo je první propočet průměru, uprostřed odpadá „starých" 10 a „nových" 15 se připojuje, vpravo rovněž odpadá nejstarší číslo a je nahrazeno nejnovějším.

a utvořme nejdříve aritmetický průměr prvních pěti hodnot:

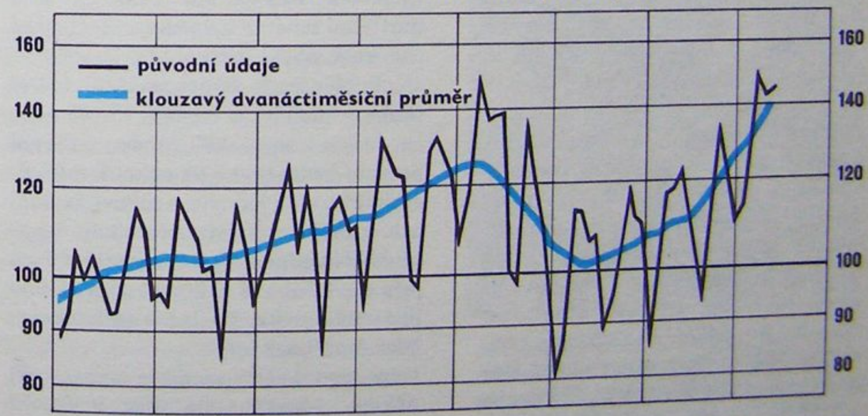
$$\frac{12{,}3 + 12{,}9 + 13{,}6 + 14{,}4 + 15{,}3}{5} = 13{,}70$$

Pak se vynechá první, „nejstarší" hodnota a připojí se další; součet v čitateli shora uvedeného zlomku se tedy zmenší o zcela vlevo stojících 12,3 a zvýší o 16,3 (šesté číslo naší řady). Vyjde další průměr 14,50, následuje 15,36 a 16,24. Tyto klouzavé průměry mohou pak posloužit jednak pro porovnání s nejnovějšími údaji a také k vypracování hladké vyrovnané křivky vývoje, která ovšem — jako každý průměr — má sklon k tomu, že nechává neobvyklé extrémy spadnout pod stůl.

Pomocí klouzavých průměrů procentních podílů lze provést jednoduché sezónní vyrovnání např. tímto způsobem: během několika let pozorování obratu určitého podniku nebo oboru se zjistí, že např. na leden připadá asi 7 % obratu, na únor 6 %, na březen 11 % atd., zatímco podíl každého měsíce v pravidelném rozdělení za celý rok je dvanáctina (= $8^1/_3$ %). *Skutečnému obratu* např. 5000 kusů v lednu odpovídá pak *sezónně „vyrovnaný" obrat* $5000 \cdot \frac{8^1/_3}{7} = 5900$, skutečnému obratu 8000 v březnu přizpůsobený údaj $8000 \cdot \frac{8^1/_3}{11} = 6050$. Březen je i v tomto případě relativně lepší než leden, protože obraty v březnu jsou vždy absolutně vyšší.

Existuje ještě řada jiných a mnohem složitějších postupů eliminace (nebo také vyjadřování) jednotlivých složek, ze kterých se skládá časová řada. Žádný však není docela bez problémů, protože v samé podstatě vyrovnávání je obsažena nutnost odchylky od daných údajů tím, že se posuzují — s nutně subjektivním zabarvením — faktory, které se mohou koneckonců jen odhadnout, avšak statistika se vždy velmi

Vnitrostátní obrat v automobilovém průmyslu 1962 = 100



„Klouzavý" průměr také v grafickém znázornění vykazuje kolísavou křivku. Hektický sestup a vzestup během roku se uklidňují a dostáváme klidnější křivku vývoje. Čísla jsou za léta 1962 až 1969. Stupnice indexu je opět logaritmická.

lehce dostává do choulostivé situace, když se nespokojuje s tím, že vyjadřuje skutečnost, anebo určuje s náležitou opatrností pravděpodobnosti, ale pustí se do problematiky příčiny a následku, vedlejšího vlivu a vedlejšího účinku. Bude ještě příležitost říci k tomu více, až se budeme v dalším oddíle zabývat korelacemi.

Nejdříve se však věnujme *výpočtu indexů*, metodě, která se pro laika stala souhrnem statistické práce a která plně vychází z podstaty časové řady.

4.3 Index a koš zboží

Index je hospodářský ukazatel, indikátor pokroku nebo neúspěchu. Na tom není nic divného, protože už etymologicky je index přesně totéž co ukazovatel (nebo ukazováček) a indikátor, což značí také jen „ukazatel". V italštině se ukazováček řekne „indice", v angličtině „index finger".

Co tedy indikuje tento index? Ukazuje nám průběh nějakého vývoje tím, že zaznamenává změny oproti dřívějšímu období. Není to však obyčejná časová řada, nýbrž v jisté míře *časová řada koše zboží*. Tento podivuhodný pojem ihned vysvětlíme.

Zjistíme-li, kolik stála tuna surové mědi v roce 1966, a vedle toho uvedeme její cenu v červnu 1967, 1968 atd., dostáváme sice časový vývoj určité ceny, nikoli však cenový index. Nebude tomu jinak ani v tom případě, když pro získání snadného přehledu stanovíme cenu za červen 1966 jako základní cenu rovnou 100, takže pro následující léta dostáváme 102,4 % nebo 95,4 % anebo 121,7 %, protože stále ještě nejde o index v běžném slova smyslu.

Index totiž musí *charakterizovat celkovou situaci*, nejen situaci jednotlivého výrobku. Index nám musí říci, jak se změnily životní náklady, kursy akcií, daňové zatížení apod. Na celém světě nejznámější a také nejvíce napadaný je *index spotřebitelských cen*, většinou nazývaný *indexem životních nákladů*.

Proti tomuto indexu se často namítá, že se v něm skutečný vzestup životních nákladů zrcadlí jen nedostatečně. Zkrátka, statistika zase jednou lže. Z hlediska jednotlivce je to často zcela oprávněné. Ovšem index nezobrazuje ani zdaleka cenové změny tím způsobem, jak se projevily v rozpočtu domácnosti pana X nebo paní Y. To je vlastnost společná všem statistickým údajům; ve statistice jde *o soubory, četnosti, nikoliv o jednotlivce.*

Ale — argumentuje se dále — také moji přátelé a známí se domnívají, že je všechno mnohem dražší, ačkoliv v minulém měsíci stoupl index jen o 0,5 %. V daném případě je především nutno upozornit na dvě okolnosti. Za prvé může u tohoto okruhu známých jít o osoby s podobnou příjmovou nebo spotřební strukturou, aniž by toto spotřební schéma bylo charakteristické pro celé obyvatelstvo. Za druhé, ve všech oblastech života existují určité skoky vnímání, při nichž si člověk něco „s hrůzou uvědomí". Index stoupal pomalu bezpochyby již několik měsíců, zřejmě se však jednalo o takové zvýšení cen, kterého si v rozpočtu nikdo nepovšiml. Najednou však chcete třeba na jaře navštívit své příbuzné a při koupi jízdenky zjistíte, že jízdné je letos daleko dražší než loni!

Něco podobného se může ukázat při nákupu nábytku, ale také u častěji kupovaného zboží, třeba u potravin. Mouka už zase stojí o 5 feniků více a index životních nákladů se nemění!

Životní náklady nelze tak lehce převést na stejného jmenovatele. Rozdíly v příjmu i individuální způsob života vyvolávají značné rozdíly ve vydáních za jednotlivé druhy zboží.

A protože cenový index se sleduje většinou jen tehdy, když už toho trápení se zvyšováním cen je hodně, zůstává mimo pozornost řada měsíců, během nichž index registroval zvýšení cen, o čemž by se poctivě mělo říci: „No tohle! A já jsem zvyšování cen ani nepozoroval!" Vzniká zcela opačná situace než při čtení horoskopů, kdy člověk vzpomíná na to, co se stalo. Při čtení statistik a zpráv o počasí si naopak říká: „Tak se podívej — už zase úplně naopak!"

Existuje ještě jeden důvod, proč index působí tak často nesprávně. Je zkonstruován na základě spotřebního schématu („*spotřebního koše*"), které přesně neodpovídá snad pro žádného spotřebitele.

Spotřební koš musí totiž přihlížet ke *spotřebě každého jednotlivce*, musí sečítat a pak dělit takřka všechny nákupy a výdaje všech domácností v celé zemi. Jinak by se muselo přihlížet k nesčetnému počtu druhů nákupů vzhledem k tomu, že někdo chodí častěji do divadla, zatímco druhý sedí raději před obrazovkou, jeden vypije týdně láhev whisky, a druhý je abstinent a vegetarián, jeden bydlí ve zděděné vile a jiný platí vysoké nájemné v novostavbě, někdo blouzní o hlavičkách chřestu, druhému se hnusí. Výčtem podobných protikladů by bylo možno zaplnit mnoho stran, ale jedno je již zcela

Ve „spotřebním koši" úřední statistiky nejsou zcela určitě ty druhy, které zde byly nakoupeny, a pravděpodobně v žádném jiném nákupním košíku. Přesto je úřední „spotřební koš" dobrou srovnávací základnou pro vývoj životních nákladů — je jakýmsi „váženým průměrem" všech spotřebitelských vydání.

INDEX MALOOBCHODNÍCH CEN A ŽIVOTNÍCH NÁKLADŮ V ČSSR

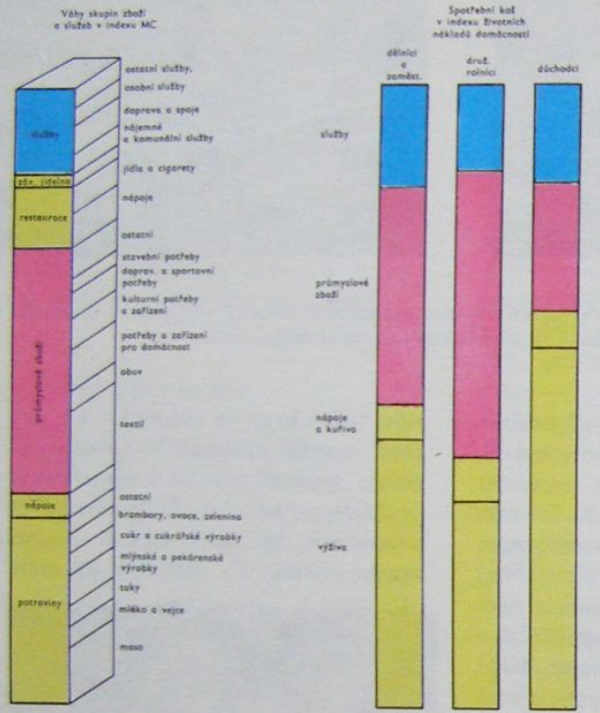


Diagram ukazuje současnou strukturu spotřeby pro výpočet indexu maloobchodních cen a indexu životních nákladů v ČSSR.

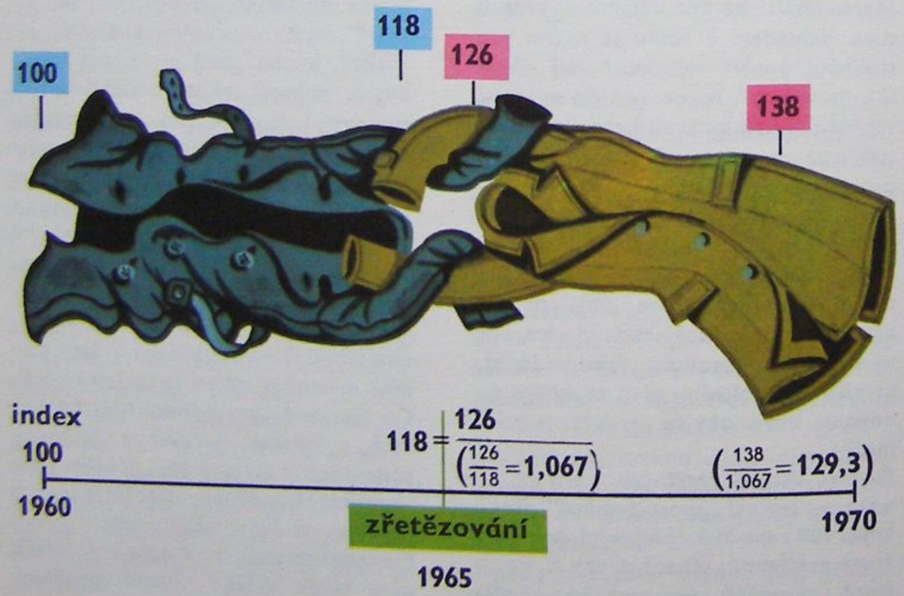
zřejmé — v indexu není možno zachytit všechny uskutečněné nákupy a výdaje, nýbrž jen některé zvlášť typické. To vypadá velmi prakticky a účelně, avšak jde-li o realizaci takového průměrného spotřebního koše, je už situace mnohem obtížnější. První otázka: *Co se vlastně kupuje?* Protože chceme budovat na solidních základech, nemůžeme měřit a odhadovat od oka, ale musíme mít podklady. Takovými podklady jsou např. rodinné účty, v nichž se pečlivě zaznamenává každé deko tuku, každá jízdenka z tramvaje, každá vodová ondulace, každé kilo brambor a každá krabice pracího prášku.

Vedete vy nebo vaše manželka takový rodinný účet? Moje manželka je vzorná hospodyně, ale takový účet nevede. Vaše asi také ne. Mimoto chceme zjistit průměrnou spotřebu, a nikoliv výdaje v rodinách, kde hospodyně nebo manžel mají zvláštní zálibu v počítání. Proto by se musel nejdříve vybrat výběrový soubor a takto „vylosovaným" domácnostem by bylo třeba slavnostně předat rodinné účty a říci: „Pěkně prosíme, zaznamenávejte nyní po dobu jednoho měsíce přesně veškeré své výdaje, dáme vám také malou odměnu jako uznání."
Někdo okamžitě odpoví: Ne, děkuji, to je moc namáhavé, nechci. Jiní budou

toužit po odměně a dají se do práce, za tři dny však ztratí chuť a nakonec tam mrzutě něco napíší. Někteří povedou rodinné účty nanejvýš pečlivě a svědomitě. Druzí začnou stejně svědomitě, ale najednou je přepadnou výčitky: „Už zase jsme koupili tři láhve piva. Jak to potom vypadá! Ale co, napíšeme místo toho jablečný mošt."
Při sestavování jedné anglické úřední statistiky bylo výslovně uvedeno, že údaje v rodinných účtech byly upravovány, „protože bylo možno dokázat, že některé výdajové položky nebyly uvedeny v plném rozsahu". Jednalo se o alkohol, tabák, cukrovinky a podobné výdaje.
U většiny rodinných účtů lze zjistit, zda byly nebo nebyly vedeny svědomitě. Máme-li však k dispozici jen číselné údaje relativně malého okruhu zvlášť pečlivých (a tím i snad zvlášť šetrných a zdatných) lidí, pak zase jejich spotřební koš nemusí být bezpodmínečně charakteristický pro většinu obyvatelstva. Vzhledem k tomu je nutno pro srovnání použít velkoobchodní obrat tak, že se např. řekne: jestliže se ročně vyrábí přes 70 miliónů hektolitrů piva (přičemž dovoz a vývoz považujeme pro jednoduchost za vyrovnané), ale z našich rodinných účtů vychází jen 50 litrů na osobu ročně, bude asi správnější použít údajů pivovarského průmyslu, který přece platí ze své produkce daň a jistě neudává více, než se vyrobilo a prodalo. Předpokládat, že více než polovina piva se vyleje do odpadu, místo aby se prodala, je stěží možné.
Pomocí těchto obtížných postupů se nakonec sestaví „spotřební koš" obsahující 100 nebo 200 výdajových položek, které zastupují veškeré druhy a sortiment skutečně kupovaného zboží.

Např.: 2 balíčky mraženého špenátu reprezentují (zastupují) mražené ovoce a zeleninu všeho druhu. Nebo 3 libry jablek reprezentují čerstvé ovoce všeho druhu, $^1/_2$ libry čočky a malá plechovka fazolí reprezentuje luštěniny atd. Větší předměty musí být někdy pomyslně rozděleny na menší díly, takže např. na jeden měsíc připadá polovina gramofonové desky, rozhlasový poplatek a „1 % televizoru" z komplexu „rozhlas, televize, magnetofony". Lze také pracovat s globálnějším ukazatelem a jako reprezentanta skupiny „zábava" použít vstupenky do kina, rozhlasový poplatek, lístek do divadla a gramofonovou desku.
Tento stručný nárys problematiky spotřebního koše umožňuje poznání dalšího problému: *Co stojí libra jablek,* jedna gramofonová deska nebo plechovka fazolí? K tomu musíme nejdříve *pokud možno přesně popsat* předměty zařazené do spotřebního koše: „libra jablek domácího původu, střední jakosti", nebo „normální gramofonová deska", anebo „celá plechovka fazolí, bílých, nejlepší jakosti". Dokonce lze stanovit i to, jde-li o dlouhohrající desku, i čistou váhu fazolí. To nám poněkud pomůže, ale jen málo. Koupím televizor se slevou nebo za ceníkovou cenu? Jaký přístroj si vyberu? „Střední jakost" je ještě dost široký pojem. A kde koupím libru jablek? V lahůdkářství na hlavní třídě nebo v obchodním domě či v zelinářském krámě malého města? U oděvů je to ještě horší. Co znamená „pár pánských vycházkových polobotek, černých s gumovou podrážkou"? Strojně šité, dobře — ale i v tomto případě jsou stále ještě možné odchylky až dvojnásobku ceny.
Předpokládejme, že i tento problém jsme nějak vyřešili. Námi pověření

K mnoha obtížím výpočtu správného indexu patří také volba obchodů, v nichž se zjišťují ceny. I stejně kvalitní zboží se někdy kupuje za rozdílné ceny, takže mnohý spotřebitel zaplatí za stejný spotřební koš více než druhý.



nákupčí rozsetí po celé zemi nakupují vždy ve stejných obchodech a podle možnosti stejnou jakost. Tímto způsobem však bylo možno vypočítat věrohodný index na mnoho let dopředu tak asi koncem minulého století. Dnes je však situace taková, že určité zboží dané kvality a vyrobené z určitého materiálu prostě není k dostání, protože se již nevyrábí. Co teď?

Matematicky lze tento problém do jisté míry řešit. Jen se musí určit doba, kdy bylo staré a nové zboží na trhu k dostání, a pak provedu zřetězování indexu. Zjistím např., že „pár dámských flórových punčoch" již prostě neexistuje, jsou k dostání jen „nylonky". Poslední cena flórových punčoch byla 1,56 DM, zatímco nylonové (tolik a tolik denier) stály 1,75 DM. Nyní mohu pokračovat v propočtu indexu s novým zbožím, a přesto nemusím měnit zvolený teoretický spotřební koš. Násobím pár dámských punčoch 156/175 a tímto způsobem vytvořím fiktivní flórové punčochy. (Viz obr. na str. 110 dole.)

Tím jsme se dostali k dalšímu problému: *Spotřební zvyklosti se mění*. K čemu jsou počtářské kousky s flórovými punčochami, které se nenosí už desítky let? A nejsou to jen punčochy. Jsou to i holicí čepelky a elektrické holicí strojky, fotoaparáty a filmovací kamery na barevný film a další tisíc a jeden druh zboží, které se kupují a které je nutno brát v úvahu. Dnes již není možné vyjádřit položku „dopravní prostředky" prostřednictvím duše na jízdní kolo, týdenní jízdenkou a cenou jízdenky za 200kilometrovou trať jízdy po železnici. Do indexu se musí dostat také výdaje za benzín a garáž.

Zkrátka a dobře: *Spotřební koš se musí v několikaletých intervalech vždy znovu revidovat*, aby se nestal směšným. Nelze však vždy znovu začínat se zcela novým indexem, vždy je nutno hledat spojení se starými údaji — provádět zřetězování, které je však stále volnější. Srovnání dnešního indexu životních nákladů s indexem z roku 1938 je již bezmála k ničemu a jestliže se s dalším zřetězováním počítá zpět až do roku 1908, je to sice matematicky zcela možné, ale jinak zcela nesmyslné. (Tím se zabývají jen historikové — podivíni, kteří nám ještě dnes pečlivě a přesně vypočítají, jakou hodnotu měl sestercius ve starém Římě.)

Jsou tím již vyřešeny všechny problémy indexu? Vůbec ne, i když jsme je naznačili jen letmo. Zcela jsme pominuli závažný problém: *vážení*. Do spotřebního koše sice dáváme 3 libry jablek, ale nikoliv 3 libry soli nebo čokolády anebo pánských bot černých. I v množství se řídíme pokud možno průměrnými spotřebními zvyklostmi.

Tak tomu však nebylo vždy. V počátcích sestavování indexu nebyl ještě „vá-

Princip zřetězování indexů: část oděvu, která byla dříve ve „spotřebním koši", zmizí v roce 1965 z trhu a musí být nahrazena podobnou, která přijde na trh nově. Ve vztahu k výchozímu roku 1960 činí index ceny starého zboží v době záměny zboží 118, nové zboží je (snad proto, že je vyrobeno z lepšího materiálu) poněkud dražší: stojí 126/118 starého zboží, tedy 1,067násobek. Za dalších 5 let stoupne cena nového zboží ze 126 na 138. Aby se udržela propočtová základna původního spotřebního koše, musí se cena starého zboží nějak fingovat. Tak by bylo v roce 1970 dosaženo indexu 129,3, totiž ceny nového zboží (138) dělené tehdejší přepočtovou relací (1,067). Podobným způsobem mohou být navzájem zřetězovány nejen jednotlivé druhy zboží, nýbrž celé „spotřební koše".

žený" spotřební koš. Myšlenka měřit snižování hodnoty peněz časově průběžným vzestupem cen sahá až do počátku 18. století. Anglický biskup Fleetwood se jako první tímto problémem zabýval v díle „*Chronicon Preciosum*", vydaném v r. 1707, které řešilo morální hlediska ztráty kupní síly peněz. V polovině 15. století bylo založena jedné anglické univerzitě založeno stipendium s tou podmínkou, že jeho držitel je musí vrátit, jakmile by jeho majetek přesáhl 5 liber. Vznikla otázka, zda se má dodržet doslovný text zakládací listiny nebo zda je spravedlivé, aby se přihlíželo k poklesu kupní síly peněz, který mezitím v průběhu půldruhého století nastal. Biskup Fleetwood na základě starých dokumentů porovnal ceny a zjistil, že prý v době od r. 1505 do r. 1707 stouply ceny obilí a masa na šestinásobek, „nápoje" a „oděvy" na pětinásobek, a proto — uzavíral — je prý možno dnes (roku 1707) přiznat stipendistovi majetek 28 liber, aniž by tím byl změněn úmysl zakladatele stipendia.

O 30 let později se Francouz Dutot zabýval otázkou, zda Ludvík XV. se svým příjmem ve výši 100 miliónů franků byl na tom lépe než Ludvík XII. o dvě stě let dříve s necelými 8 milióny. Také on sestavil pro porovnání seznam zboží, který obsahoval mimo jiné i kozu, holuba, náklad sena a — což je pozoruhodné — též „služby", a to mzdu služky a nádeníka. Na tomto základě určil znehodnocení peněz na jednu dvaadvacetinu a tak usoudil, že Ludvík XV. byl na tom mnohem hůře než kdysi Ludvík XII.

Francouzský král Ludvík XV. měl příjem 100 miliónů franků. O dvě stě let dříve měl Ludvík XII. příjem jen 8 miliónů. Dutot prokázal v 18. století pomocí primitivního „spotřebního koše", že se hodnota peněz v těchto dvou stech letech zmenšila na 22/100 — Ludvík XII. měl tedy vyšší reálný příjem.

Nejznámějším se však stal propočet indexu provedený italským historikem a finančníkem Gianrinaldo Carlim, který roku 1764 uveřejnil studii o úbytku hodnoty peněz a načrtl cenový vývoj v Itálii od r. 1500 do r. 1750. Jeho „spotřební koš" se skládal z vína, pšenice a oleje, tedy nejcharakterističtějších produktů země. Ale ani on ještě „nevážil", nýbrž vypočítal pouze aritmetický průměr relativního vzestupu cen. Řekl např.: litr vína stál v roce 1500 10 lir, o 20 let později však 15 lir; 1 kilogram pšenice stál v roce 1500 8 lir, o 20 let později 14 lir; 1 litr oleje stál nejdříve 60 lir a později 80 lir. Z toho vyplynul následující propočet indexu:

$$\frac{\left(\dfrac{\text{nové ceny}}{\text{staré ceny}}\right)}{\text{počet druhů zboží}} \qquad \frac{\dfrac{15}{10}+\dfrac{14}{8}+\dfrac{80}{60}}{3} =$$

$$= \frac{\dfrac{18}{12}+\dfrac{21}{12}+\dfrac{16}{12}}{3} = \frac{55}{36} = 1{,}525$$

Při základním roku 1500 by se měl jeho „index" zvýšit ze 100 na 152,5.

Přitom se však vůbec nepřihlíželo k dvěma důležitým faktorům. Především chybí jakýkoliv údaj o „*zvyklostech spotřeby*": je přece veliký rozdíl, jestliže průměrný Ital spotřeboval 10 litrů vína, 5 kg pšenice a 3 litry oleje, nebo 5 litrů vína, 10 kg pšenice a 1 litr oleje. Pouze v případě, že cenový vývoj probíhal u všech tří druhů zboží stejně,


DELLE MONETE
E DELL' INSTITUZIONE
DELLE ZECCHE D'ITALIA
DELL' ANTICO, E PRESENTE SISTEMA D'ESSE:
E DEL LORO INTRINSECO VALORE, E RAPPORTO CON LA PRESENTE MONETA.
DALLA DECADENZA DELL' IMPERO SINO AL SECOLO XVII.
Per utile delle Pubbliche, e delle Private Ragioni.
DISSERTAZIONI
DEL CONTE
DON GIANRINALDO CARLI RUBBI
CAVALIERE, E COMMENDATORE DELLA SACRA RELIGIONE, ED ORDINE MILITARE
DE' SS. MAURIZIO, E LAZARO.
IN MANTOVA,
MDCCLIV.
CON LICENZA DE' SUPERIORI.

K úvahám o kupní síle peněz došel Gianrinaldo Carli ve své studii o italském mincovnictví. Hrabětem Carlim začínající propočty indexů životních nákladů.

by nebylo nutné vážení. Za druhé je třeba přihlédnout k tomu, že se mohou změnit spotřební zvyklosti. Co když se v jednom roce, v němž v důsledku neúrody oliv je olej mimořádně drahý, spotřebuje více chleba a vína, ale méně oleje?

Trvalo ještě značnou dobu, než se vytvořila představa o spotřebním koši. Je pozoruhodné, že je úzce spjata se zjišťováním existenčního minima chudých vrstev obyvatelstva. Joseph Lowe sestavil v roce 1822 přehled výdajů rodiny zemědělského dělníka a londýnské měšťanské domácnosti a porovnal zvy-

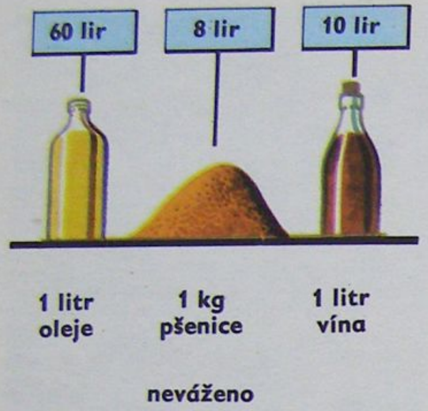


Schéma prvního primitivního propočtu indexu: cena tří základních potravin se stanoví za váhovou jednotku a porovná v průběhu roku — bez ohledu na množství spotřeby: „neváženo".

Propočet indexu jednoduchého „spotřebního koše": důsledky cenových změn se váží přesně odpovídajícím spotřebovaným množstvím: „vážení podle spotřeby".

šování cen pro tento „spotřební koš" (ve spotřebním koši měšťanské rodiny byla zahrnuta také mzda služebné, výdaje na výchovu, almužny a opravy). Později to byli Le Play a pruský statistik Engel, kdo pozorovali důsledky vzestupu cen na životní náklady chudých rodin prostřednictvím výpočtu indexů. Pokud jde o problematiku změn spotřebních zvyklostí a o přechod na výhodnější zboží, povíme o tom hned něco více. Dříve však ještě ukážeme jednoduchý příklad *výpočtu moderního indexu* na základě staré metody Carliho.

Nejdříve zvolíme základní rok — Carli zvolil rok 1500, zůstaňme také u něho. V tomto roce, jak uvádí Carli, stál 1 litr vína řekněme 10 lir, 1 kilogram pšenice 8 lir, 1 litr olivového oleje 60 lir. Carlimu to postačovalo. My se však ptáme dál: Jaké množství těchto potravin se tenkrát spotřebovalo za měsíc (nebo také za rok, případně čtvrt roku)? Zjistíme, že by to mohlo být na hlavu a měsíc asi 5 litrů vína, 8 kilogramů pšenice a 2 litry oleje.

A nyní vypočítáme náklady tohoto spotřebního koše v roce 1500. Je to: 5 litrů vína po 10 lirách = 50 lir, 8 kg pšenice po 8 lirách = 64 lir, 2 litry oleje po 60 lirách = 120 lir, celkem 234 lir. Zde by bylo možno namítnout, že Italové snad také konzumovali ještě ryby, maso, ovoce a jiné potraviny, avšak tato námitka není v tomto stadiu našeho zkoumání rozhodující. Vybrali jsme tři charakteristické potraviny k tomu, abychom měli možnost alespoň do jisté míry sledovat na vývoji jejich cen všeobecný vývoj.

Uvedené 234 liry považujeme za rovné 100 — to není žádný taškářský trik, nýbrž jen zjednodušení, které nám umožní lépe pozorovat pozdější procentní zdražení. To proto, že nyní před-

pokládáme, že v roce 1501 podražilo víno o 1 liru, to znamená, že nezměněný spotřební koš nestojí nyní 234, nýbrž 239 lir.

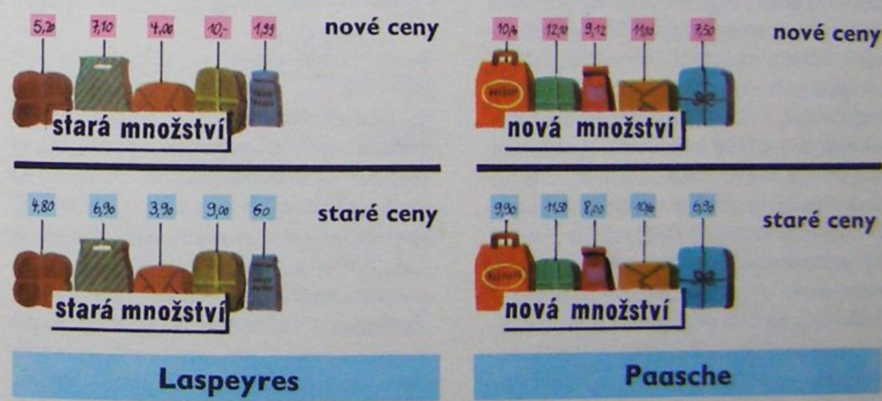
V daném případě se „index" roku 1501 rovná $\frac{239}{234} = 102{,}1$, a tudíž ve vztahu k základnímu roku (1500 = 100) můžeme říci, že životní náklady v tomto jednom roce stouply o 2,1 %.

V roce 1502 podražila také pšenice a stojí 9 místo 8 lir a náš spotřební koš stojí tedy 274 lir, index 105,5. O kolik stouply životní náklady proti předchozímu roku?

Ve srovnání se základním rokem je situace jasná: vzestup o 5,5 %. Proti předchozímu roku mohu vzestup vypočítat tak, že počítám poměr 239 : 247 lirám (= 102,1 : 105,5 indexního čísla). To pro naše zkoumání stačí, ale nemůžeme si nevšimnout dalšího předpokladu. Víme-li, že v důsledku neúrody pšenice nejen stoupla cena, ale také klesl odbyt, můžeme pak s dobrým svědomím dále pracovat s naším starým spotřebním schématem? Ne. Statistika proto používá poněkud složitějších přepočítávacích vzorců.

Protože tyto vzorce mají základní význam pro výpočet indexů, musíme je alespoň načrtnout. Předpoklad: víme, že v důsledku výsledků sklizně, cenového vývoje a změn spotřebních zvyklostí v roce 1510 bylo spotřební schéma z roku 1500 již nesprávné, protože nové schéma se skládalo z pěti (místo čtyř) litrů vína, deseti (místo osmi) kilogramů pšenice a dvou litrů oleje (nedošlo ke změně). Ceny tohoto zboží v roce 1510: vína 12 lir, pšenice 7 lir, oleje 65 lir. Jak nyní provést nějaké rozumné porovnání?

Jsou dvě možnosti: buď provedeme porovnání na základě starých množství, nebo na základě množství nových.



Mezi četnými cenovými indexy nabyly zvláštního významu dva: Laspeyresův index (porovnání cen na základě původně spotřebovaného množství) a Paascheho index (porovnání cen na základě nové spotřeby).

V prvním případě to vypadá takto:

$$\frac{\text{nové ceny} \times \text{stará množství}}{\text{staré ceny} \times \text{stará množství}};$$

$$\text{tedy } \frac{60 + 56 + 130}{50 + 64 + 120} = 105{,}1.$$

Ve druhém případě:

$$\frac{\text{nové ceny} \times \text{nová množství}}{\text{staré ceny} \times \text{nová množství}};$$

$$\text{tedy } \frac{48 + 70 + 130}{40 + 80 + 120} = 103{,}3.$$

Výsledek je tedy značně rozdílný. Kdo chce v daném případě „statisticky dokázat", že životní náklady poměrně značně stouply, zvolí metodu, podle které vychází index 105,1. Kdo má opačný zájem, zvolí metodu druhou. Správné jsou obě. První (stará množství jako základna) se jmenuje *Laspeyresův index*, druhá (nová množství) *index Paascheho.* Paasche a Laspeyres byli němečtí národohospodáři z konce 19. století.

(V ČSSR se index maloobchodních cen počítá podle Laspeyresova vzorce jako vážený aritmetický průměr individuálních indexů cen zboží — reprezentantů, přičemž vahami jsou podíly jednotlivých skupin zboží na maloobchodním obratu roku 1968 a základem je cenová hladina z ledna 1968. Kromě toho se v ročním intervalu počítá i index Paascheho, a to jako vážený harmonický průměr týchž individuálních indexů cen zboží — reprezentantů, přičemž vahami jsou odpovídající údaje o struktuře maloobchodního obratu v běžném roce.)

Vědecky neudržitelný je naproti tomu *„index laiků"*, který však byl ještě v minulém století prosazován se vší vážností i vědci a který zní:

$$\frac{\text{nové ceny} \times \text{nová množství}}{\text{staré ceny} \times \text{stará množství}};$$

$$\text{tedy } \frac{48 + 70 + 130}{50 + 64 + 120} = \frac{248}{234} = 105{,}9.$$

Na tento index se instinktivně odvolává většina hospodyň, když si stěžují, že „všechno podražilo", že již nevyjdou s penězi na domácnost a vůbec, co povídat. Porovnávají měsíční výdaje před třemi nebo pěti roky s tím, co vydávají dnes — a přitom snadno zapomínají, že to není jen zvýšení cen, které se projevuje ve zvýšených výdajích, nýbrž také větší požadavky, o jejichž vznik se postará čilá reklama. Jak silné mohou být takové změny spotřeby za delší dobu, hned ještě uslyšíme od Jeana Fourastié.

Nejdříve ještě několik slov k Laspeyresovu a Paascheho indexům. Nejsou zdaleka jediné, které věda zná, avšak jsou nejčastěji používány. Americký národohospodář Irving Fischer navrhl jejich spojení do *„ideálního indexu"*, který měl vypadat takto:

$$\boxed{I_F = \sqrt{I_P \cdot I_L}}$$

Z Paascheho indexu a Laspeyresova indexu měl být získán geometrický průměr — Fisherův index. Většinou se však zůstává u obou „klasických" indexů, přičemž je třeba ještě dodat, že Laspeyresův index téměř vždy dosahuje vyšších hodnot než index Paascheho vzhledem k tomu, že při neproporcionálním zdražení jednotlivých druhů zboží spotřebitel většinou přechází na jiné, lacinější druhy, takže „nová množství" zachytí část zdražení. To zní poněkud rozporně, vzpomeneme-li na právě kritizovaný „index laiků" a na zjištění, že se dnes častěji kupují dražší věci než před několika lety. Především je nutno rozlišovat záměnu zboží, které již tu vždy bylo (např. třešně za meruňky atd.), a výdaje za zboží, které se nově objevilo ve spotřebním koši. Za druhé — a to nejen ve statistice a ekonomické vědě — není při bližším pohledu vždy všechno tak zcela jasné a nesporné, jak bychom si přáli.

Od vzorců Laspeyresova a Paascheho indexů vede jen krátká cesta do houštin národního hospodářství, k teorii hodnoty, psychologického uspokojování potřeb, k teorii mezního užitku atd. Nemůžeme jít touto cestou, protože by nás zavedla příliš daleko od našeho hlavního tématu. Chtěli jsme však naznačit, o co jde.

Své potřeby mohu v jistých mezích zaměnit, aniž bych pociťoval ztrátu. Např. jím rád jablka, ale ta jsou právě velmi drahá, zatímco broskve jsou velmi laciné. Protože vlastně jím rád i broskve, koupím raději 1 kg broskví za 2 DM než $^1/_2$ kg jablek za 2,30 DM. Mám stejný „užitek", a přesto jsem 30 feniků ušetřil.

Proti indexům se někdy argumentuje takto: Jsou reálné proto, že v koši mají stále stejné zboží, i když každý jen trochu rozumný spotřebitel přechází při zvýšení ceny na jiné zboží. Předpokládejme, že v mém indexovém spotřebním koši je v měsících květnu až červenci po 1,5 kg třešní na týden; pak ale přijde neúroda, která vyžene ceny třešní na dvojnásobek: můj index stoupá, zatímco spotřebitel pravděpodobně přejde na dovážené pomeranče nebo banány anebo jiný druh ovoce.

Ragnar Frisch popsal celou problematiku životních nákladů a jejich propočtu lehce srozumitelnou formou v článku publikovaném v časopise *„Econometrica"*. Uvádí tam mimo jiné, že všechny indexové vzorce se nedají uvést do souladu s jinými koncepty hospodářské teorie, protože jsou příliš „mechanistické" — ani s oním základním principem, že se změnou cen se mění i struktura spotřeby. Také se tam zabývá potížemi, které se vyskytují při mezinárodních srovnáváních cen životních potřeb.

Ideální řešení se dosud nenašlo. Praxe indexového propočtu se úzkostlivě brání pojetí pružné spotřeby, a to z toho pádného důvodu, že by pak byla nehybná a solidní základna spotřebního koše vydána napospas kdejaké libovůli. Jestliže se chce manipulovat spotřebním košem, musí se na to jít jinak. Stát, který má silný vliv na tvorbu cen, může to dělat třeba tak, že intervenuje to zboží, které je ve spotřebním koši. Zůstaňme u našeho příkladu. Je-li málo třešní, mohou být ve velkém množství dováženy v zájmu udržení nízké ceny, a domácímu zemědělství se současně zaručí absolutní zákaz dovozu jablek, meruněk, hrušek a veškerého dalšího domácího ovoce, které ve spotřebním koši není.

To je poměrně snadné tehdy, když oficiální „spotřební koš" úřední statistiky obsahuje jen málo položek — úsilí se pak může soustředit na udržení nízké úrovně relativně malého množství cen. Naopak je velmi nemilé, jestliže se nešťastnou náhodou zdraží právě to zboží, které má v indexu významnou váhu. Tak např. na podzim roku 1969 rakouská vláda zoufale bojovala o maloobchodní cenu určitého modelu „brouka" (Volkswagenu), protože ovlivňovala index. Kupující a spotřebitel může a bude zpravidla měnit nákupní zvyklosti, jestliže se některá

Struktura výdajů jednotlivých domácností v ČSSR

(v ‰)

	Výdaje domácností		
	dělníků a zaměstnanců	družstevních rolníků	důchodců
Celkem	1.000,0	1.000,0	1.000,0
Výživa	425,3	334,4	583,0
Maso a masné výrobky	116,5	77,1	162,3
Mléko a mléčné výrobky	40,0	30,9	65,2
Vejce	16,2	2,1	25,6
Tuky	49,8	34,6	78,1
Mlýnské, pekárenské a těstárenské výrobky	53,0	71,6	89,2
Brambory, ovoce, zelenina a výrobky z nich	42,1	28,1	64,2
Cukr, cukrovinky a cukrářské výrobky	29,7	35,9	40,3
Ostatní potraviny	26,7	25,7	36,1
Jídla ve veřejném stravování	51,3	28,4	22,0
Nápoje a kuřivo	58,0	69,6	56,0
Nealkoholické nápoje	6,0	7,0	8,0
Pivo	13,3	22,1	16,0
Víno	6,8	6,8	7,0
Lihoviny	10,9	17,1	8,0
Kuřivo a kuřácké potřeby	21,0	16,6	17,0
Průmyslové zboží	349,6	459,0	207,0
Odívání	109,6	133,8	64,6
Ostatní textilní zboží	23,7	28,2	15,8
Obuv	25,8	28,9	25,0
Galantérie	3,9	2,9	2,6
Zařízení a potřeby pro domácnost	45,5	74,4	13,0
Kulturní zařízení a potřeby	37,4	32,7	19,3
Osobní dopr. prostředky a sportovní potřeby	53,3	58,2	1,5
Drogistické a zdravotnické zboží	24,6	24,3	32,5
Paliva	10,0	17,7	29,2
Stavební potřeby	15,8	57,9	3,5
Služby	167,1	137,0	154,0
Nájemné a komunální služby	54,4	19,6	69,5
Doprava a spoje	33,1	25,0	31,4
Osobní služby	13,3	9,3	15,0
Opravy a zakázkové práce	28,8	55,9	24,2
Vzdělání, kultura, zábava	11,3	9,1	5,0
Rekreace, léčebná péče, dětská zařízení	21,4	11,2	4,0
Ostatní služby	4,8	6,9	4,9

Váhy ve spotřebních koších indexů životních nákladů jsou pro domácnosti jednotlivých výdajových skupin stanoveny na podkladě statistiky rodinných účtů za rok 1968.

Pramen: Jan Mach, Revize indexů maloobchodních cen a indexů životních nákladů, čas. Statistika, 1972, č. 7—8.

značka nepoměrně zdraží. Statistika však musí setrvat na zvolených položkách — většinou tomu u velkého počtu položek bývá tak, že nadprůměrné a podprůměrné změny cen se navzájem přibližně vyrovnají, takže index nakonec zcela dobře plní svůj úkol ukazatele všeobecného vývoje.

K jakým podstatným změnám spotřebních zvyklostí za delší dobu dochází, i když se omezíme jen na potraviny a zcela zanedbáme výskyt nových technických přístrojů, ukázal mimo jiné francouzský futurolog Jean Fourastié. Zjistil, že ve Francii v 18. století pozůstávala výživa nejširších vrstev obyvatelstva ze tří čtvrtin z chleba, zbytek připadal na mléčné výrobky, trochu zeleniny, zatímco masité jídlo bylo jednou týdně. Průměrná roční spotřeba vína činila asi 20 litrů, roční spotřeba cukru asi 2 kg — a spotřeba masa nebyla za celý rok větší než 5 kg.

Indexy se však nevypočítávají a neuveřejňují vždy pravidelně. *Často se jich používá jen pro určitý účel.* Předpokládejme třeba, že by šlo o otázku, jak se vyvíjely mzdy (např. dělníků v mlékárnách) v poměru k čistým ziskům (např. v mlékárenském průmyslu). Vycházíme z předpokladu, že by to vypadalo takto:

Rok	Index mezd	Index zisku
1956 (= 100)	100	100
1957	103	109
1958	107	124
1959	120	123
1960	122	130
1961	127	132
1962	130	141
1963	134	152
1964	146	154
1965	149	151
1966	157	150
1967	158	157
1968	160	160

Posuzováno z hlediska základního roku 1956 je situace zcela jasná: mzdy i zisky vzrostly v letech 1956 až 1968 o 60 %.

Neexistuje však ani právní, ani matematický, ani logický důvod, který by nás nutil k tomu, aby rok 1956 byl zvolen za základ.

Dojde-li k jednání o nové kolektivní smlouvě, mohou zaměstnanci argumentovat např. takto: „V posledních třech letech vzrostly mzdy o necelé 2 %, ale zisky o téměř 7 %."

Je to pravda? Toto zdůvodnění vychází z volby roku 1966 jako základního roku a tabulka vypadá takto:

Rok	Index mezd	Index zisku
1966 (= 100)	100	100
1967	100,8	104,7
1968	101,9	106,6

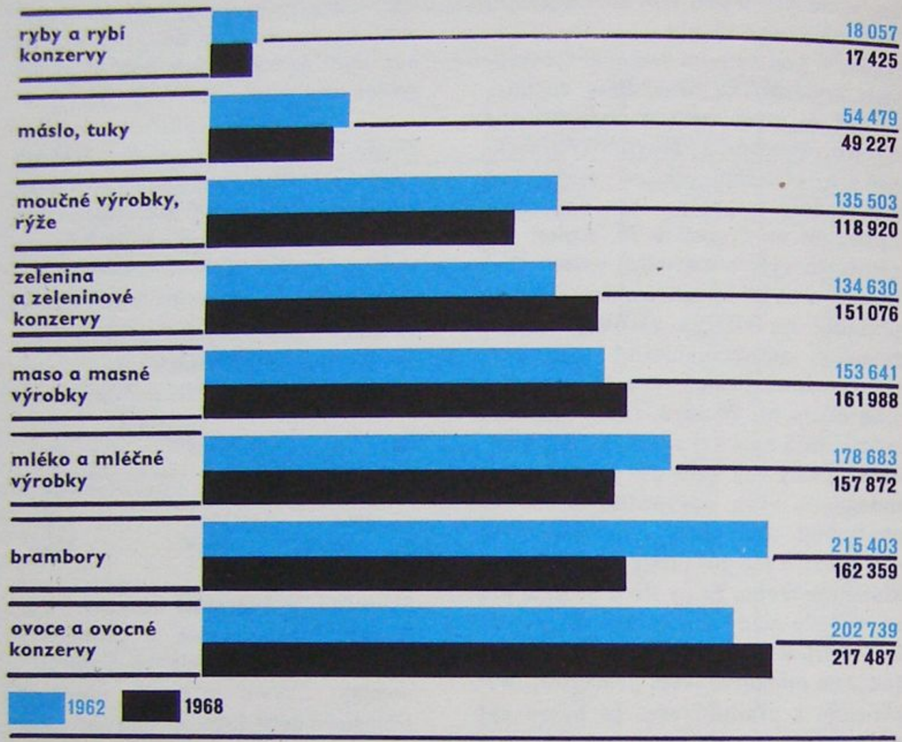
Skutečně, mlékárenští dělníci byli poškozeni ve srovnání s podnikateli. Opravdu? Podnikatelé to popírají a uvádějí: „Vývoj je nutno posuzovat z hlediska delší doby. Během posledních šesti let mzdy přímo vyskočily, zatímco zisky zaostaly."

A položí na stůl následující údaje:

Rok	Index mezd	Index zisku
1963 (= 100)	100	100
1964	109	101,2
1965	111	99,5
1966	117	98,7
1967	118	103,3
1968	119,5	105,2

Skutečně, zisky podnikatelů velmi zaostaly v poměru ke mzdám. Statistika znovu „dokázala" zcela opačné tvrzení. Je to však *jen rozdílná volba základního roku*, která vede k tomuto domnělému rozporu.

Jestliže se základní rok stanoví až ná-



I v průběhu několika málo let se může spotřební schéma podstatně změnit, jak ukazují tato čísla ze západního Berlína. Především je nápadné, že brambory v letech 1962—1968 musely postoupit první místo ovoci a ovocným konzervám. Také spotřeba mléčných a moučných výrobků poklesla, zeleniny a zeleninových konzerv a masa a masných výrobků se zvýšila.

sledně, je to vždy podezřelé. *Základní rok však může klamat* i jinak.

Srovnávají-li se indexy akcií půl tuctu zemí, dojde se — pokud není znám základní rok indexu — k úplně nesmyslným výsledkům. Tak např. koncem března 1971 dosáhl americký Dow-Jonesův index 913, švýcarský SKA-index 234, vídeňský CA-index 63, tokijský Dow-Jonesův index 2365; Frankfurt (Commerzbank) hlásil 726, Paříž naproti tomu 105 a Milán 58.

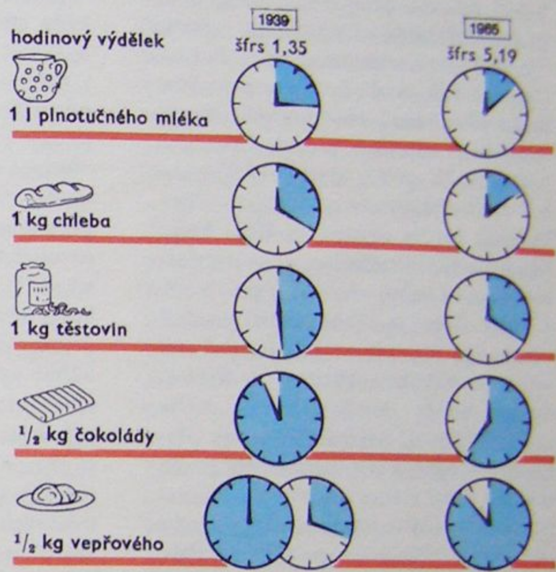
Mezinárodní porovnání indexů je však sotva možné i tam, kde byl zvolen stejný základní rok, neboť faktory podmiňující výpočet indexu jsou většinou navzájem tak rozdílné, že porovnání způsobuje víc zmatku než užitku. Když se do přepočtu ještě zahrne oficiální měnový kurs volně nesměnitelné měny, je chaos úplný.

Někdy existují pokusy provádět mezinárodní srovnání na základě propočtů, jaké množství mléka, chleba nebo cukru

Aby bylo možno provést mezinárodní porovnání životní úrovně, vztahuje se někdy kupní síla mzdy na určité spotřební zboží. Toto porovnání není na mezinárodní úrovni bez problémů. V jedné zemi, jako v tomto případě ve Švýcarsku, je naproti tomu dosti přesvědčivé.

Kupní síla mzdy v roce 1939 a 1965

K tomu, aby bylo možno koupit následující potraviny, musel dospělý dělník pracovat .... minut:



si může průměrný dělník koupit za průměrný hodinový výdělek. To může v jisté míře poskytnout orientační hodnoty, které však o životní úrovni vypovídají velmi málo.

Zpravidla lze říci, že *index ztrácí svou platnost na státní hranici* — a to platí i pro mnohé další statistické údaje, u kterých mezinárodní normování nedosáhlo dosud žádného povzbuzujícího pokroku.

Populárním druhem indexu je *doložka ceny chleba nebo zlatá doložka*, kterých se používá v některých letech (zvláště v poválečných a inflačních) legálně nebo ilegálně ve smlouvách. Oba případy nelze označit za šťastnou volbu, protože právě chléb a zlato jsou příkladem manipulovaného zboží, poskytujícího přirozeně ochranu proti cválající inflaci — pokud jejich použití není zakázáno, a tím také soudní ochrana bezpředmětná.

Modernější formou doložky než cena chleba je mzdová a důchodová dynamika, která je většinou vázána na index životních nákladů. (Zda tato dynamika nevyvolává také zvýšenou dynamiku mzdové a cenové spirály, tím se zde naštěstí nemusíme zabývat.)

Ještě na jednu okolnost je nutno u indexu poukázat. V tomto případě, stejně jako u všech slovních výkladů statistických údajů, není opatrnosti nikdy nazbyt. Vzpomeňte si na náš fiktivní příklad ze začátku 16. století: ve třech po sobě jdoucích letech stouply životní náklady ze 100 (základní rok) na 102,1 a 105,5. Z druhého na třetí rok výdaje tedy stouply o 3,3 %, ale o 3,4 body. Tento rozdíl je prakticky bezvýznamný. Co ale s tím, když index stoupne např. z 502 na 505? Stoupne o 0,6 % nebo o 3 body? *Nuže, chceme-li, aby vzestup vypadal větší, mluvíme raději o bodech.* (Jestliže laik myslí, že jde o procenta, je to jeho vina.) *Má-li vypadat menší, mluvíme o zlomcích procent.* Klesne-li index ze 70 na 63, klesl o 10 % nebo o 7 bodů atd.

To jsou hříčky se slovy a čísly, které někdy vznikají náhodou a neúmyslně, ale někdy i velmi vědomě a plánovitě. A ještě něco je třeba uvážit: každá měna — i legendární švýcarský frank — trpí plíživou inflací; jedna více, druhá méně. Jestliže chceme měřit *„reálné" životní náklady*, je nutno přihlédnout také *k poklesu hodnoty peněz* a s ním ruku v ruce jdoucímu *zvyšování příjmů*. Cenový index životních nákladů se např. v NSR v letech 1962 až 1969 zvýšil o 19,5, ale index průměrného hrubého týdenního výdělku průmyslového dělníka ve stejném období o 58 %. U všech skupin povolání není ovšem tak snadné zjistit hrubé a zejména čisté mzdy, jako je tomu u dělníků a zaměstnanců. Čisté příjmy např. samostatně činných osob je možno zjistit jen pomocí povinně přiznaného příjmu pro důchodovou daň. Značné zdroje příjmů některých povolání, jako je spropitné a příjmy z „melouchů", nelze odhadnout vůbec.

Jestliže pomyslíme na tyto potíže, z nichž všechny nebyly ani zdaleka uvedeny, pak je skoro zázrak, že indexy úřední statistiky jsou nakonec přece jen tak překvapivě blízko pravdě. Jak blízko, nelze ovšem vždy určit, neboť stoupne-li např. index z jednoho měsíce na druhý ze 108,6 na 108,8, neříká to nic víc než toto: pravděpodobnost, že hodnoty, které jsou základem propočtu, stouply, je nepatrně větší než pravděpodobnost, že se nezměnily nebo poklesly.

V následující kapitole budeme ještě blíže zkoumat otázku přesnosti. Zde bychom chtěli především zdůraznit, že i když předpokládáme směrodatnou odchylku přesnosti jen ve výši 3 ‰ (a to je při pečlivé práci nereálně málo), musíme říci: údaj 108,6, který jsme uvedli za předchozí měsíc, leží s 95% pravděpodobností mezi 108,0 a 109,2 (dvojnásobná směrodatná odchylka dává 95 % pravděpodobnosti, „pravidlo 2σ"); údaj 108,8, který byl určen za nový měsíc, leží mezi 108,2 a 109,4. Není tedy vůbec vyloučeno, že správný index za předchozí měsíc je 108,8 a za nový měsíc jen 108,5, nebo také že oba jsou si přesně rovny.

Jestliže však naproti tomu delší řada takových indexů vykazuje stejný vývoj, stává se správnost pozorování stále pravděpodobnější. Následují-li např. za 108,6 a 108,8 jako další čísla 109,1 a 109,5, můžeme právem — nikoliv však s absolutní jistotou — předpokládat, že vývoj indexu za dané 4 měsíce vyjadřuje skutečně existující vzestupný vývoj.

Žádný index není zcela přesný, a jestliže tomu tak je, pak jen náhodou. To však není argument proti indexu nebo proti jakémukoliv jinému statistickému šetření. Není-li možno získat žádnou dokonalou informaci, musíme se spokojit s pokud možno nejpřesnějšími odhady. A nejpřesnější odhad je stále jen odhad — ale přece nepoměrně cennější než nevědomost, prázdná domněnka nebo odhad „podle oka". Každý koš zboží je konec konců jen výběrový soubor a už v samé podstatě výběru je, že nemůže zprostředkovat absolutní jistotu o celém souboru.

Potřebujeme však vždy výpočty na zlomky procent? Naše myšlení je většinou příliš ovládáno utkvělou předšinou představou, že číslo vypočítané až na poslední jednotkové nebo desetinné místo je vrcholem přesnosti a pravdivosti. Ve skutečnosti je tomu často naopak. Jestliže při sčítání lidu v roce 1970 bylo v New Yorku zjištěno 7 771 730 obyvatel, pak přinejmenším poslední tři místa jsou absolutně bezvýznamná. Uvedené číslo ve skutečnosti znamená: „s největší pravděpodobností těsně 7,8 miliónu." Proti domnělým chybám při sčítání lidu protestovali nejen starostové a městští radní v Americe, ale bylo veřejným tajemstvím mezi statistiky, že zvláště ve věkové skupině 20—30letých černochů mužského pohlaví se „ztratilo" až 10 %, protože nebyli při sčítání lidu zastiženi.

Úplně groteskně vypadá, když se v nějaké příručce říká, že počet obyvatel Madrásu činil 1 833 504 a toto číslo se pak doplněním (+) vykazuje jako odhad! Zde se dá realisticky říci jen toto: „bezpochyby téměř necelé dva milióny." Avšak víra v magii přesných čísel je tak veliká, že jí proti svému lepšímu vědomí a svědomí platí daň dokonce i statistikové.

Všeobecně platí: *Žádný strach před zaokrouhlenými čísly*, neboť jsou málokdy méně nesprávná než čísla údajně přesná a nic nepředstírají. Jen zřídkakdy je možno na otázku „kolik je hodin?", odpovědět („gong oznámí") „15 hodin 32 minuty 40 vteřin". Stejně užitečná a méně lživá je odpověď: „půl čtvrté".

4.4 Kouzlo procent

Jak působivě zní, jestliže se dozvíme, že 14,28 % všech dotázaných ještě nikdy nepilo červené víno nebo nejedlo mandarínky! Každý laik však může získat bez zvláštních nákladů stejně přesný výsledek — stačí, aby se zeptal 7 osob, zda už někdy jedly mandarínky. Řekne-li jedna z nich „ne", je to 14,28 % dotázaných. (Při dvou „ne" je to 28,55 % — také impozantní výsledek.)

Pomocí procentuálních údajů se lze případně vyhnout trapné otázce, jak vlastně velká byla *základna výpočtu*, jak velký byl rozsah výběrového souboru: jeden ze sedmi je 14,28 % stejně jako 500 ze 3500 nebo 1998 ze 13 990; s tím malým rozdílem, že výsledek, který se zakládá na sedmi údajích měření nebo výsledcích dotazů, je většinou bezcenný.

Nemusí to vždy být zlý úmysl, který předvádí procenta místo absolutních čísel. Někdy to může být náhoda, hloupost, bezmyšlenkovitost nebo naivní radost z velkých procent. Tak v roce 1967 bylo možno číst v jednom americkém motoristickém časopise, že rozsah prodeje jednotlivých značek zahraničních motorových vozidel se v roce 1966 proti roku 1965 velmi značně změnil.

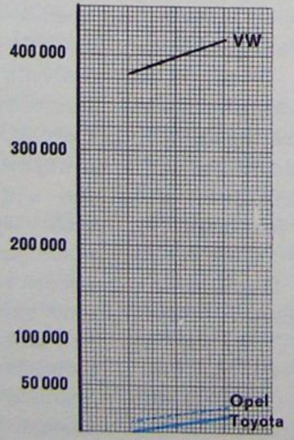
Volkswagen např. měl přírůstek 7 %, Opel 84 %, MG dovezl jen o 3 % více, a naproti tomu Toyota o 751 % více. Jen uvažte — zatímco Volkswagen mohl zvýšit svůj vývoz do Ameriky pouze o 7 %, zaznamenal japonský podnik ve stejné době zvýšení o 751 %, tedy sedmiapůlnásobek (= na osmiapůlnásobek).

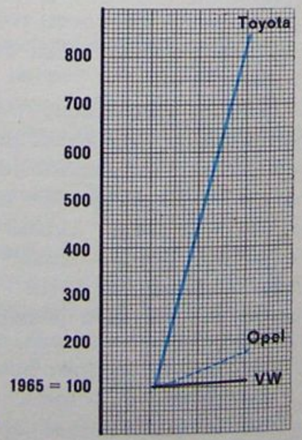
Naštěstí jsou však uvedeny absolutní údaje, a tak může i laik při bližším pohledu brzy pochopit zprvu nepochopitelné; statistik by se již na první pohled dohadoval, co se vlastně stalo. U Toyoty byli se senzačním rozmachem jistě velmi spokojeni. Ale je zcela dobře myslitelné, že ve Wolfsburgu (sídle západoněmecké firmy Volkswagen) byli také spokojeni: buď jak buď, export VW do USA stoupl v absolutních číslech o téměř 30 000 vozů, vývoz japonského podniku naproti tomu o sotva 18 000. Díváme-li se však na procentuální přírůstek, vznikne klamný dojem, jako by proti „raketovému" vzestupu japonského závodu se méně než skromně plazil Volkswagen.

Procenta vycházejí ze základny předchozího roku a tehdy to bylo tak, že VW prodal v roce 1965 na americkém trhu 380 000 vozů, zatímco Toyota teprve s prodejem začínala a zmohla se jen na prodej 2800 vozů. Je ovšem jasné, že podnik, který již tak hluboko pronikl na trh, může dosáhnout z roku na rok jen poměrně malých zvýšení, zatímco firma, která přijde na trh nově, může v následujícím roce prodat i pěti, deseti, ba dokonce padesátinásobek.

To lze zjistit i z údajů o počtu nově přihlášených osobních vozů ve Švýcarsku, kde např. Mazda v roce 1970 dosáhla padesátinásobku odbytu z roku 1967, zatímco naproti tomu Fiat jen 1,3násobku. V absolutních číslech se však prodej Mazdy zvýšil z 32 (!) o 1764 na 1796, Fiatu ze 17 070 o 5465 na 22 535. Kdo tedy chce získat přesný a přehledný obraz o situaci, *musí se vždy podívat nejen na procenta, ale také na absolutní údaje*. Jestliže se mu tato čísla neposkytnou, pak je nejvyšší čas, aby se stal nedůvěřivým. Kdo měří svůj růst jen v procentech, ten pravděpodobně začal z velmi malého základu a je malý ještě i dnes. To nemusí být vždy špatné, protože koneckonců každý začínal jako malý, ale schovávat se vytrvale za procenta není moc dobré znamení.

Přitom je třeba uvážit ještě další okolnost. Vezmeme-li si k ruce tabulku složitého úrokování, můžeme z ní mimo jiné zjistit, že 1000 DM zúročených na 3 % vzroste za 20 let na 1806 DM; jsou-li však zúročeny 6 %, zvětší se na





Vysokých procent přírůstku lze dosáhnout tím snadněji, čím nižší je výchozí základna. V letech 1965—1966 stouply v USA dovozy automobilů Toyota na osminásobek, u Volkswagenu naproti tomu ani ne o 10 %. Ale absolutní přírůstek obratu (vlevo) byl u VW přesto mnohem vyšší.

3207 DM, při 8 % dokonce na 4661 DM. Je-li tedy výchozí základna stejná (1000 DM), znamenají vyšší úroky a úroky z úroků stále rostoucí náskok. Ale co se stane, když někdo uloží 10 000 DM na 3 % a druhý 1000 DM na 6 %? Bohatší má dnes o 9000 DM více — ale jak tomu bude za 20 let? Bohatší bude mít 18 061 DM a chudší přes vyšší zúročení jen 3027 DM — rozdíl se tedy zvýší na téměř 15 000 DM. K tomu je nutno přihlížet vždy, když se navzájem srovnávají přírůstky.

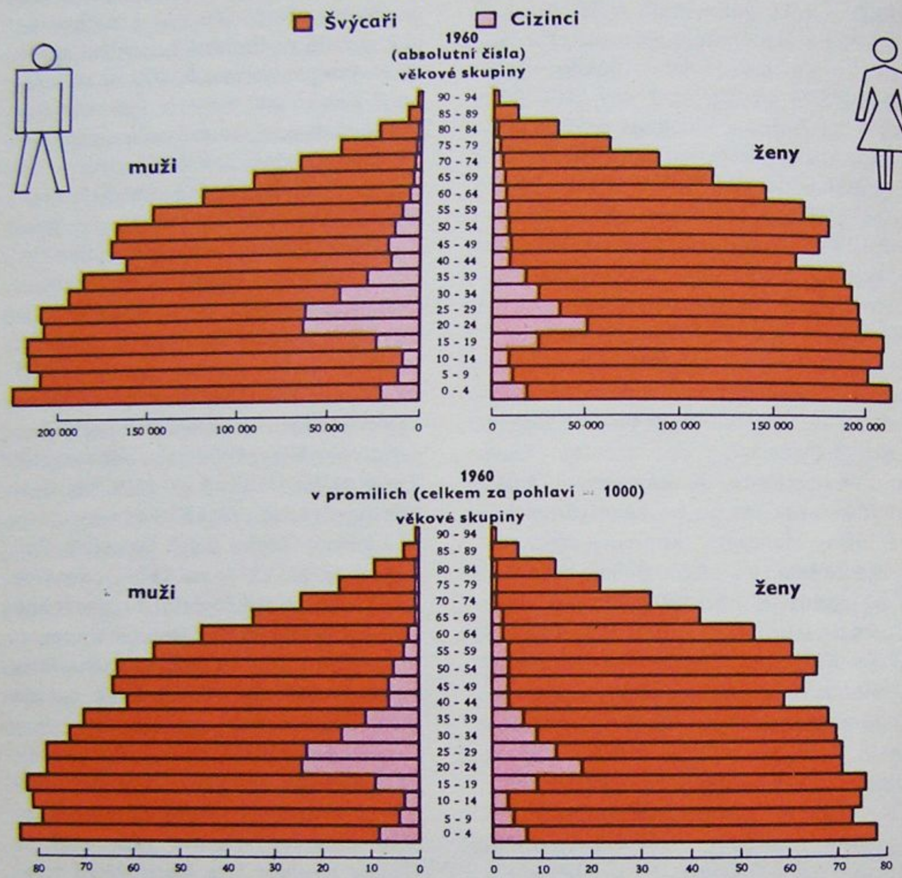
Zároveň však chceme již na tomto místě varovat před dalším zneužíváním statistiky, kterým se budeme později ještě více zabývat — před přehnaným „*extrapolováním*", promítáním čísel z minulosti do budoucnosti. Omezíme se tedy jen na to, že ještě jednou se vším důrazem konstatujeme, že *procenta sama o sobě říkají málo nebo nic*. Kdo zamlčuje absolutní čísla a dělá poplach s procenty, chce zřejmě něco skrýt — to se potvrzuje všude tam, kde se mohutně zdůrazňuje plnění a překračování plánovaných úkolů v procentech, aniž se cokoliv bližšího řekne o plánu a plánovaných úkolech samých. Slovo „skrýt" neznamená nezbytně, že se chtějí zastírat neúspěchy, nýbrž docela jednoduše to, že určité údaje mají zůstat utajeny — a pro utajení skutečností se procenta opravdu výborně hodí.

Tato hra se dá hrát i obráceně, a to tehdy, když je *snaha imponovat absolutními čísly, protože poměrná čísla vypadají příliš nepříznivě*. Vyvolávat dojem absolutními čísly není těžké potud, pokud ochotný čtenář statistiky pozoruje věci ze své osobní perspektivy. V takovém případě jsou pak přirozeně všechna miliónová a miliardová čísla náramně působivá. Za první světové války zavládlo jednou mezi hladovějícím obyvatelstvem velké vzrušení, když se proslechlo, že 50 000 tun brambor se prý zkazilo nedbalostí železniční správy — jen pomyslete: 50 000 tun brambor! Zemědělští experti vyslovili sice politování nad nesprávným disponováním, nicméně pokrčili rameny; při roční sklizni 50 miliónů tun a normálním „úbytku" 10 % neznamená uvedených 50 000 tun prakticky vůbec nic. Jiný „optický klam" s poměrnými a absolutními hodnotami je znám každému pozornému burziánovi: absolutní vzestup (nebo pokles) kursů nějaké akcie vede při zběžném pohledu velmi často k omylu o mnohem důležitějším relativním burzovním zisku nebo ztrátě. To je zvlášť důležité na těch burzách, kde jsou různě „těžké" hodnoty. Jestliže v New Yorku např. stoupne Pan American ze $13\,^1/_8$ na $15\,^7/_8$, rovná se tento vzestup o 2,75 dolaru kursovnímu zisku o zhruba 20 %. Stoupne-li naproti tomu IBM z 273 na 302, znamená kurs zvýšený o 29 dolarů přírůstek ani ne jedenáctiprocentní.

Dokonce i grafické znázornění ve stejném měřítku pro oba druhy akcií vyvolalo nesprávný dojem. *Chceme-li opticky porovnat relativní růst, hodí se k tomu mnohem lépe logaritmická stupnice*. Logaritmické, semilogaritmické a jiné nekonvenční stupnice jsou ovšem většinou poněkud podezřelé tehdy, když se předkládají laickému publiku, které je neumí správně číst.

Těmito, ale i jinými klamnými grafickými manévry se budeme zabývat ještě v dalším oddílu (viz str. 233). Nyní chceme ukázat na jiném příkladu manipulaci s absolutními a poměrnými čísly, která nemá demagogický záměr, ale je přesto značně matoucí. V roce 1966 sdělilo vídeňské letiště: „Mezi

VĚKOVÁ STRUKTURA USEDLÉHO ŠVÝCARSKÉHO OBYVATELSTVA

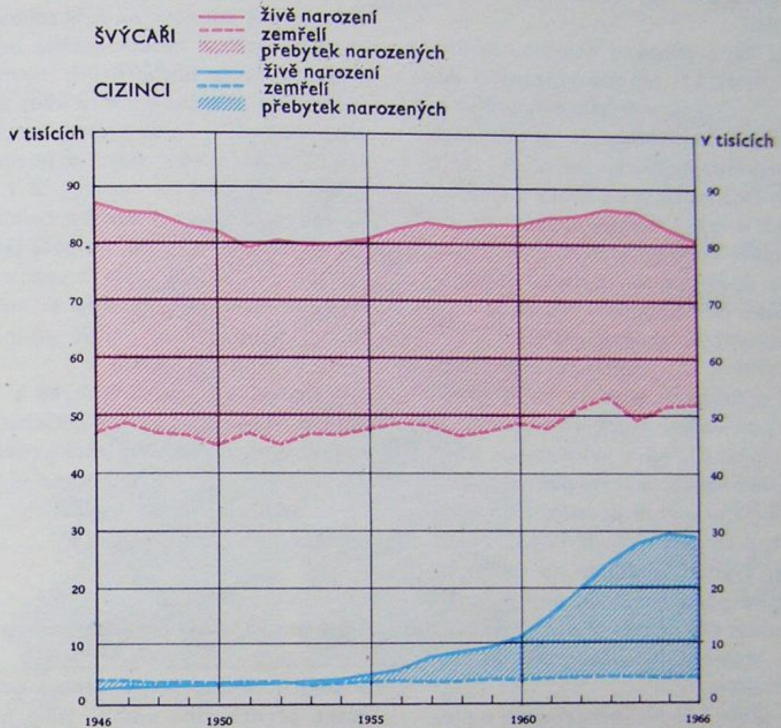


37 západoevropskými letišti se Vídeň řadí... sice ještě mezi menší letiště, pokud však jde o přírůstky dopravy, je Vídeň již na čtvrtém místě. V roce 1964 bylo při 22 818 startech odbaveno 725 049 cestujících... V nejsilnějších dnech je registrováno až 5000 cestujících."

Zde je v několika slovech obsaženo téměř všechno, na co je nutno brát zřetel, chceme-li se naučit zacházet se statistikami. Jádro výpovědi („řadí se mezi menší") je odsunuto stranou slovem „sice" a pak se vynáší trumf: již na čtvrtém místě v přírůstcích. Toto „již" je však zcela nemístné, protože přírůstky jsou vysoké téměř vždy, jestliže je výchozí základna malá.

Pak následují absolutní čísla určená pro laika, který nemá možnost porovnání: 22 818 letů, 725 049 cestujících — to je přece ohromné! Absolutně ano, relativně nikoliv. Letiště Rýn — Mohan odbavilo téměř 4 milióny cestujících,

Narození a úmrtí podle domovské příslušnosti 1946—1966



Věková struktura usedlého švýcarského obyvatelstva zřetelně ukazuje, proč jsou počty narozených cizinců relativně vyšší než švýcarských občanů: ve věkových třídách od 20 do 34 let je podíl cizinců zvlášť vysoký — právě tato věková skupina zahrnuje ženy v plodném věku a mladé manžele. (Graf: Švýcarský statistický úřad.)

západní Berlín 2,7 miliónu, Hamburk, Düsseldorf a Mnichov po 1,4 miliónu, nemluvě ani o amerických letištích, o Paříži a Londýně.

A nakonec ještě jako zvláštní pozoruhodnost poukaz na nejsilnější dny a v nich dosahované absolutní nejvyšší hodnoty („až"). Věcně je jistě úplně správné, že v jednom takovém „nejsilnějším dnu" bylo jednou registrováno až 5000 cestujících, protože se však současně neuvádí žádný denní průměr, utkví čtenáři v paměti představa: „denně 5000 cestujících". To tam není výslovně řečeno — ne, zcela určitě ne, ale už je to tak, že se čtení statistik nestalo ještě všeobecně ovládaným uměním. Statistikou vídeňského letiště se tak podrobně nezabýváme proto, že by byla obzvlášť rafinovaná, záludná nebo demagogická, nýbrž proto, že umožňuje zřetelně ukázat, čeho je nutno se při čtení nějaké statistiky vyvarovat. A tu platí za prvé, za druhé

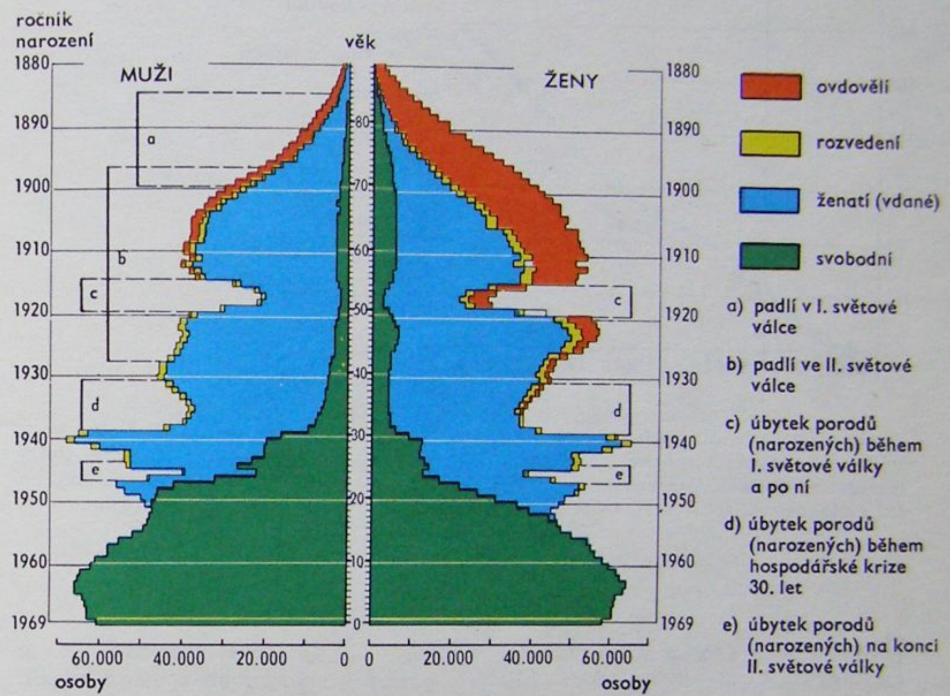
a za třetí: *vyčíst více než je uvedeno.* „V nejsilnější dny až..." neznamená „denně".
Pokud jde o poněkud nejasnou formulaci „mezi 37 západoevropskými letišti...", chceme v daném případě s důvěrou předpokládat, že je to těch 37 nejvýznamnějších. Je však zcela dobře možné, že se za touto formulací skrývá lstivý předchozí výběr: Stuttgart může být například co do velikosti docela dobře pátým z deseti západoevropských velkoměst. Postačí, aby těch deset bylo vhodně vybráno!
Absolutní čísla mohou klamat také jiným způsobem: když se např. předstírá porovnání, které ve skutečnosti není namístě. Tak v jedněch novinách bylo uvedeno, že „v první polovině roku 1966 činil přebytek porodů u Švýcarů 15 261, u cizinců však již 12 652, takže se my, Švýcaři, podílíme na celkovém přebytku porodnosti jen 54,7 %". To je strašné: cizí dělníci přivádějí na svět skoro stejně tolik dětí jako domácí obyvatelstvo.
Nebo přece ne? Ne, protože jde o *přebytek* porodů, tedy o rozdíl mezi narozenými a zemřelými. Pro přebytek porodů je však směrodatná věková struktura: čím větší je podíl mladých lidí (a v tom žen v plodném věku), čím menší je podíl starých lidí, tím více je porodů, tím méně úmrtí a tím vyšší pak přebytek porodů. Cizí dělníci jsou však zpravidla mladí a zdraví, poměrně řídká úmrtí vyplývají spíše z pracovních úrazů než ze stařecké slabosti.
V daném případě to pisatel také správně poznal — jedna z chvályhodných výjimek — neboť nakonec lapidárně konstatoval: „Je zřejmé, že naprosto chybějí předpoklady k porovnání údajů o narozených mezi cizinci a Švýcary. Tato čísla lze však zneužít."

Téměř všechna čísla — a proto i všechny statistiky — je možno zneužít. Kdo nechce padnout za oběť takovému zneužití, kdo se nechce nechat od demagogů nebo přehorlivých novinářů vehnat do úzkých, bude se vždy s pochybností ptát: *Co se s čím porovnává?* Má toto porovnání smysl a je oprávněné? A především: netvrdí se *v původním textu více, než dovolují čísla sama poznat?* A konečně nikdy nemůže škodit, jestliže se zeptáme jako u soudu: *Cui bono?* Komu to slouží? Kdo se pomocí těchto čísel jeví ve vzlášť příznivém světle?
Demagogické zneužití i chyby z nedbalého výkladu jsou tak důležité, že se jimi budeme později ještě podrobně zabývat (viz str. 229). Prozatím však tento oddíl ukončíme doplňujícím přehledem o poměrných číslech.

4.5 Míry, podíly, poměrná čísla

V tomto oddíle, stejně jako i na začátku předchozího oddílu, jsme několikrát mluvili o relativní četnosti, procentech a indexech, a proto teď chceme jako shrnutí a doplnění uvést několik slov o *poměrných (relativních) číslech.* Statistika rozlišuje *podíly, míry a poměrná čísla* a toto rozlišení je víc než pouhá hra se slovy, protože každý druh poměrných čísel má svá vlastní omezení a podléhá vlastním zákonitostem.
Podíly jsou ještě relativně nejméně problematické. Při nich jde o to, že *se měří díly nějakého celku:* 12 % všech vysokoškoláků je krátkozrakých, 30 % evropské produkce piva pochází z NSR, 63 % osobních příjmů ve Velké Británii připadá na mzdy a platy atd. Celková čísla vysokoškoláků, evropské produkce piva a britských osobních příjmů

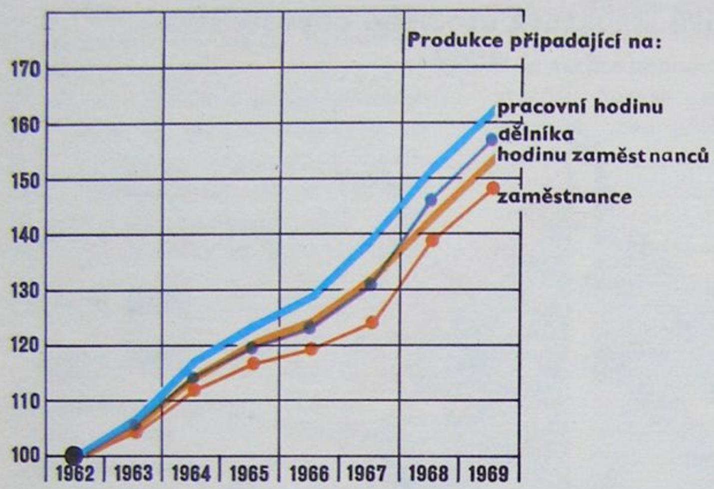
Věková struktura usedlého obyvatelstva 1969



Anatomie stromu života: kombinované zobrazení stavu obyvatelstva a jednotlivých ročníků narození. (Pramen: Statistisches Handbuch für die Republik Österreich, 1970.)

se kladou do poměru jako odlišitelná složka z celkového souboru. Na začátku třetí kapitoly jsme dali do vzájemného poměru veškeré v Mnichově bydlící obyvatelstvo v roce 1960 s podskupinou „cizinci a osoby bez státní příslušnosti v základním souboru mnichovského bydlícího obyvatelstva" a v této souvislosti jsme mluvili také o tom, že kruhové výseče (dortové řezy) se zvlášť dobře hodí pro grafické zobrazení takových skutečností.

U měr, které jsou velmi příbuzné s podíly, se veličiny *stejného druhu vztahují k jedné z nich* (nebo k průměru všech) jako k základu. Produkce piva v NSR činila v roce 1967 71,3 miliónu hektolitrů, ve Francii 20,6 miliónu hektolitrů. Zde máme k výběru dvě míry: a) produkce piva v NSR činí 345 % produkce Francie; b) produkce Francie činí 29 % produkce NSR.
Celkové množství lze členit na dílčí množství, které může být dále členěno.

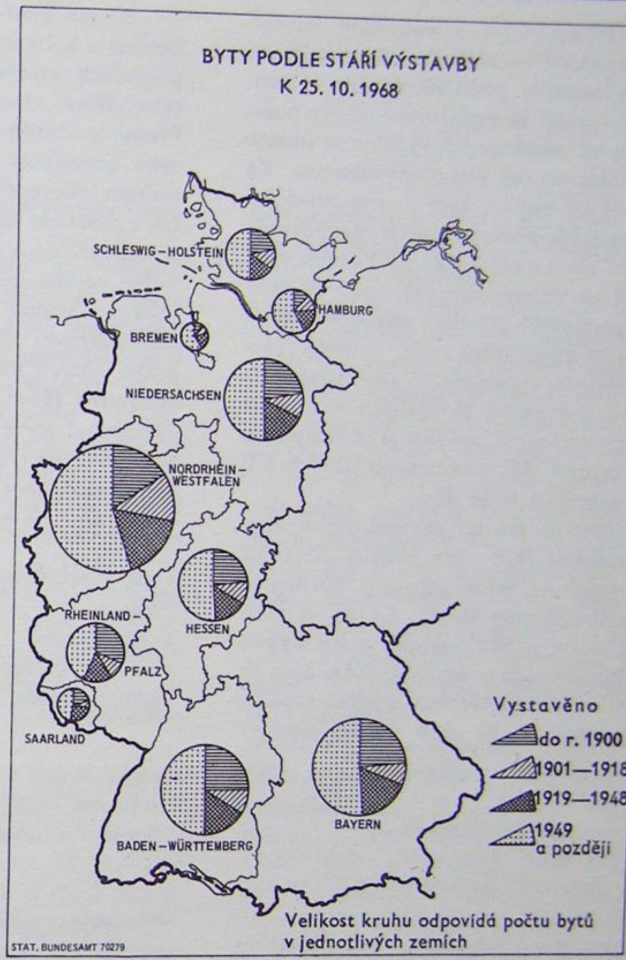


Téměř každou situaci, každou skutečnost je možno zobrazit různým způsobem. Nejen proto se zdá, že statistiky tak často „lžou". Na tomto grafu Spolkového statistického úřadu (NSR) je možno zjistit, že se „produktivita" od roku 1962 do roku 1969 zvýšila ze 100 na 148, na 152, na 157 nebo na 161 — podle toho, na které osoby nebo výkony se tento index vztáhne.

Díváme-li se na dílčí množství nezávisle na jejich absolutní velikosti jako na stoprocentní dílčí soubory, je porovnání struktury zvlášť zřetelné. Jeden švýcarský statistik to velmi pěkně dokumentuje na příkladu řeči usedlého švýcarského obyvatelstva: asi 75 % švýcarských občanů mluví německy, asi 20 % francouzsky, zbývajících 5 % připadá převážně na italštinu, kromě malého počtu mluvících rétorománsky. Z cizinců však mluví téměř 55 % italsky, jen asi čtvrtina německy a 10 % jinými jazyky, převážně asi španělsky. Míry někdy porovnávají v jisté míře jen své vlastní hodnoty. Je tomu tak v případě, zvolíme-li určitý rok jako základní a zároveň vztažný bod časové řady. Základní rok se zpravidla stanoví jako 100, takže všechny následující hodnoty vyjadřují dosažené procento. To je postup, který již dobře známe z indexů. Index se však liší od jednoduché míry především tím, že váženým výběrem více veličin charakterizuje celou kategorii četností, událostí nebo hodnot, zatímco míra se vztahuje na jednotlivé nebo nečetné objekty.

Poměrná čísla je třeba používat s větší opatrností než míry nebo podíly. Vyjadřují vztahy, které statistik teprve hledá, a jejich oprávněnost může být tudíž sporná. U podílů je samozřejmé, že celkové množství a dílčí množství jsou v přímém a těsném vztahu; také u měr tato souvislost téměř vždy existuje. Poměrná čísla závisejí však na dobré vůli a rozumu toho, kdo je vypočítává. Zde se totiž vytváří vztah mezi zcela rozdílnými veličinami, o kterých

Kruhové výseče („dortové řezy") udávají sice jen relativní četnosti (podíly na souboru), mohou se však použít také k porovnání veličin jako v tomto grafu Statistického spolkového úřadu (NSR): celkový stav bytů v jednotlivých zemích se vyjadřuje velikostí kruhové plochy.



se předpokládá (nebo tvrdí, nebo se lze domnívat), že jsou v určité souvislosti. Nejznámější případ poměrných čísel jsou čísla „na jednoho obyvatele", o nichž jsme již mluvili v předchozím oddíle. V daném případě můžeme dospět ke „kojenci, který kouří doutníky". Chceme-li se něčemu podobnému vyhnout, vzniká otázka, od kterého roku věku určit „schopnost kouření". Patnáctiletí by neměli kouřit, ale mnozí kouří — kde má vést věková hranice? Jak je to s údaji o sňatcích a úmrtích? V západním Berlíně jsou sňatky podstatně vzácnější a úmrtnost o polovinu vyšší než v NSR, jestliže je vztáhneme na 100 000 obyvatel. Ano, úmrtnost v západním Berlíně je téměř třikrát větší než ve Venezuele. Proč? Věková struktura obyvatelstva je jiná.

Rozvojové země s populační explozí „explodují“ v základně stromu života obyvatelstva, podíl dětí je příliš vysoký, tak vysoký, že v důsledku velkého podílu osob neschopných výdělku se snižuje životní úroveň všeho obyvatelstva. Za dnešních hygienických a zdravotnických poměrů je úmrtnost dětí a mladých ročníků malá (i když stále ještě o něco vyšší než ve střední Evropě); vymírají tedy v prvé řadě přestárlí obyvatelé, kteří žili již na přelomu století v tehdy řídce obydlené Venezuele. V západním Berlíně (a v jiných evropských velkoměstech v mírnější podobě) je obyvatelstvo přestárlé, žije tam mnoho starých lidí a poměrně málo dětí.

Porovnání jen na základě počtu obyvatelstva říká tedy vlastně až příliš mnoho — neboť zahrnuje porovnání celého stromu života. Kdyby se však měla na tomto základě zjistit hygienická situace a lékařská péče, bylo by asi nutno porovnávat úmrtnost podle generací. Kolik osob věkové skupiny 50—60 let umírá ročně v západním Berlíně a kolik ve Venezuele — v obou případech vztaženo na 100 000 osob této věkové skupiny?!

Přesné statistiky neporovnávají proto také porodnost obvykle s celkovým počtem obyvatelstva, nýbrž s počtem žen v plodném věku. Stejně jako kojenci nekouří doutníky, muži nerodí děti. Volba vztahu není vždy zcela snadná. Jaké je „správné“ porovnávací měřítko pro dopravní nehody? Počet vozidel, počet ujetých kilometrů, počet osobokilometrů (který je možno přirozeně odhadnout jen z výběrových souborů), počet obyvatel, hustota provozu? Každou alternativu je možno zdůvodnit i zamítnout ve prospěch jiné.

Tolik k poměrným číslům. Teď si však chceme všimnout hesla, které právě padlo a které je ústředním problémem moderní statistiky. Toto heslo zní: *Výběrový soubor*.

5 Výběry a dotazníky

5.1 Sběr dat, šetření a zpracování

Dosud jsme se pohybovali v pohádkové zemi blahobytu statistiky: když jsme potřebovali jakákoliv data, pozorování, měření, sčítání a jiné výpočetní podklady, měli jsme je po ruce, ať již šlo o průměrnou roční teplotu v Quitu, o příjem obyvatel Zbohatlíkova, o ceny olivového oleje v době hraběte Carliho nebo o oběti úderů kopyty v pruské armádě. Nyní se však musíme zamyslet nad tím, jak se statistik dostane ke všem těmto údajům a číslům, z nichž pak sestavuje tabulky, které klade do vzájemného vztahu, zkoumá z hlediska hodnoty výpovědi a přetváří více nebo méně složitými matematickými postupy.

Při statistické práci se většinou rozlišují tři stupně: za prvé *šetření*, za druhé *zpracování*, za třetí *vyhodnocení*. Daleko nejdůležitější částí práce se zdá být vyhodnocení (tím se zpravidla zabývají všechny učebnice statistiky nejpodrobněji), nesmíme však přitom zapomenout na elementární pravdu: *žádná statistika není lepší než její surovina*. Tak jako nemůže být správný úsudek, jsou-li nesprávné předpoklady, stejně tak nejsou k ničemu nejobtížnější početní operace, když číselný materiál je hned od počátku nesprávný nebo nedostačující. A co je ještě horší — početní chyby lze opravit, nevhodné metody zpracování mohou být nahrazeny lepšími, avšak je-li prvotní záznam údajů chybný, nezbývá než uvést jako moto tato Schillerova slova: „To právě je kletbou zlého činu, že množením se musí stále rodit zlo.“

Co tedy může statistik učinit, aby této kletbě předešel, aby na počátek své práce nepostavil „zlý“, nýbrž dobrý čin? První otázkou je, *zda ví, co chce*. Není tomu tak vždy, neboť mnohé statistické údaje úřady shromažďují v očekávání, že později mohou posloužit jako cenné podklady, aniž se v daném okamžiku dá přesně říci, co se vlastně hledá. Tak pruské armádní sbory rok co rok poctivě zaznamenávaly příčiny úmrtí svých vojáků — a jednoho dne z nich statistik vylovil zabité úderem kopyta a dokázal použitelnost Poissonova rozdělení.

Mnohem horší však je, když sice víme, čeho se chce dosáhnout, *ale zvolíme k tomu nevhodnou cestu*. Tak se např. jednou zkoušela odolnost gumových člunů. Byly podrobeny zatěžkávací zkoušce, a když obstály, pokládaly se za použitelné. Později se ukázalo, že právě tyto zkoušené čluny v provozu poměrně často praskaly. Proč? Tvrdou zatěžkávací zkouškou byly totiž i dobré čluny tak narušeny, že část z nich nebylo již možno označit jako nezávadné. Wallis a Roberts, z jejichž knihy „*Statistické metody*“ jsme tento příklad převzali, dělají následující závěr: „Údaje se tedy vztahují na počet výrobků, které obstály v určité zkoušce, nikoliv však na to, že tyto výrobky vydrží skutečné zatížení.“

Někdy jsou údaje potřebné pro statistické zkoumání po ruce a je pouze třeba uspořádat je z jiného hlediska. Je tomu tak v případě, kdy je možno použít bohatého materiálu úředních statistik. V tomto případě mluvíme o *sestavení druhotné statistiky*. Protikladem k tomu je *zhotovení prvotní statistiky*, kde musíme nejprve prvotní údaje získat zjišťováním a provést jejich členění a kdy se nepoužívá údajů, které jsou k dispozici.

Pokud jde o rozsah zjišťování, je třeba rozlišovat mezi *výběrem a vyčerpávajícím zjišťováním*, při němž se zkoumá, měří, počítá nebo jiným způsobem zjišťuje celý v úvahu přicházející základní soubor (statistický soubor, populace).

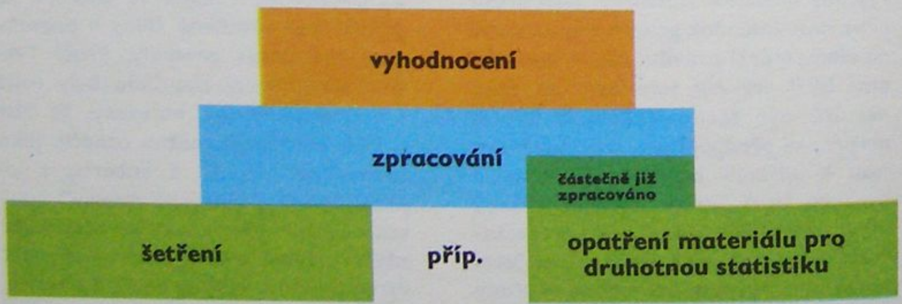
Vyčerpávající zjišťování se provádí zpravidla jen tehdy, když základní soubor je poměrně malý. Obsáhlé zjišťování a dotazování veškerého obyvatelstva země si může dovolit jen stát — a i ten jen jednou za deset let.

Avšak ať již jde o výběr nebo vyčerpávající zjišťování, postupy se omezují v obou případech na několik málo základních metod. Nejjednodušší metodou je *pozorování:* děj se pozorně sleduje a — pokud možno — číselně zachycuje. Tak např. pohraniční orgány počítají osoby odjíždějící ze země a současně zaznamenávají, jaké předkládají pasy. Nebo na výpadové silnici stojí dva lidé, kteří pozorují provoz, a za každý vůz jedoucí do města nebo z města si udělají čárku, anebo pověřený pracovník hlásí denně vedení závodu, kolik vagónů černého uhlí bylo předchozího dne naloženo apod.

Taková pozorování se týkají „pravých“ dějů; při experimentech se naproti tomu děje a situace plánovitě aranžují, často podle přesného *„plánu vzorků“*. Tak např. z výrobního pásu se vezme každá pětistá žárovka a zkouší se, jaké vydrží zatížení a jakou má dobu životnosti. Prostředky na ochranu rostlin se zkoušejí na pokusných polích a jejich účinek se měří a srovnává podle sklizně.

Jestliže se nejedná ani tak o věci jako o lidi, o jejich názory, úmysly, přání a reakce, má dotazování nejdůležitější úlohu, jíž se proto teď budeme zabývat podrobněji.

Nejdříve je však třeba říci několik slov k druhému kroku statistické práce —



Tři stupně popisné statistiky: materiál se nejdříve opatří šetřením nebo získá z prvotních statistik, které jsou již k dispozici. Potom se zpracuje (člení, uspořádá) a nakonec se vyhodnocuje a interpretuje.

ke zpracování. Zde jde o shrnutí a uspořádání prvotních údajů (jeden příklad jsme již poznali, když jsme popisovali zavádění třídních intervalů a zobrazení četností — viz str. 53). Toto stadium statistické práce charakterizuje především sestavování tabulek. V nejjednodušší formě zaznamenává tabulka jen rozdělení četností *jednoho* znaku — např. příjmové poměry abiturientů (viz str. 38). Mnohem častěji však vzniká matice z řádků a sloupců, které mohou být dále členěny. Členění řádků ukazuje např. výňatek z demografické ročenky otištěné ve „Statistické příručce pro Rakouskou republiku“ (viz tabulku dole).

Členění je provedeno nejen podle zemí, ale v jejich rámci i za dvě šetřená období, kde jsou zase zvlášť uvedeni „muži“ a „ženy“. Členění sloupců ukazuje jiná tabulka z téže příručky (str. 136).

Střední očekávaná délka života

Státy	Úmrtnostní tabulky z let	Pohlaví	Střední očekávaná délka života osob ve věku								
			0	10	20	30	40	50	60	70	80
Evropské státy											
Belgie	1891—1900	m	45,4	50,3	41,8	34,2	26,7	19,7	13,4	8,1	4,6
		ž	48,8	52,8	44,4	37,0	29,5	21,9	14,8	8,9	4,9
	1959—1963	m	67,7	59,9	50,3	40,9	31,7	22,9	15,5	9,7	5,3
		ž	73,5	65,3	55,5	45,9	36,3	27,2	18,7	11,4	6,1
Dánsko	1901—1905	m	52,9	54,0	45,4	37,4	29,4	21,8	15,0	9,2	5,0
		ž	56,2	55,8	47,5	39,6	31,7	23,8	16,3	10,0	5,5
	1965—1966	m	70,1	62,1	52,5	42,9	33,5	24,5	16,6	10,3	5,6
		ž	74,7	66,2	56,4	46,6	37,1	28,0	19,4	11,9	6,3
Německo	1901—1910	m	44,8	51,2	42,6	34,6	26,6	19,4	13,1	8,0	4,4
		ž	48,3	53,4	44,8	36,9	29,2	21,4	14,2	8,5	4,7
	1965—1967	m	67,6	59,9	50,4	41,1	31,9	23,1	15,4	9,6	5,4
		ž	73,6	65,5	55,7	46,1	36,6	27,5	18,9	11,5	6,1
Anglie a Wales	1910—1912	m	51,5	53,1	44,2	35,8	27,7	20,3	13,8	8,5	4,9
		ž	55,4	55,9	47,1	38,5	30,3	22,5	15,5	9,6	5,5
	1965—1967	m	68,7	60,5	50,9	41,4	31,9	23,0	15,3	9,5	5,5
		ž	74,9	66,4	56,6	46,9	37,3	28,2	19,8	12,4	6,9
Finsko	1901—1910	m	45,3	49,9	42,2	34,9	27,4	20,0	13,6	8,3	4,6
		ž	48,1	51,7	44,5	37,3	29,9	22,3	15,1	9,0	4,9
	1961—1965	m	65,4	57,3	47,8	38,5	29,5	21,2	14,3	8,9	4,9
		ž	72,6	64,2	54,4	44,7	35,2	26,0	17,5	10,3	5,3
Francie	1898—1903	m	45,3	49,3	41,0	33,9	26,7	19,8	13,3	7,9	4,4
		ž	48,7	51,5	43,6	36,4	29,1	21,6	14,6	8,7	4,9
	1966	m	68,2	60,0	50,4	41,1	32,0	23,5	16,1	10,2	5,8
		ž	75,4	66,9	57,1	47,5	38,0	29,0	20,5	12,8	7,0

Tato statistická tabulka ukazuje mnohonásobné členění. V řádcích je provedeno členění podle zemí a v jejich rámci za dvě časová období a ještě zvlášť vždy podle pohlaví. Ve sloupcích je členění podle věku.

Obyvatelstvo podle pohlaví, věku a povolání

Země	ze 100 obyvatel										
	v roce	ženy	bylo ve věku...					v roce	byli činní		
			méně než 15	15 až 30	30 až 45	45 až 65	65 a více		muži	ženy	celkem
Evropské země:											
Belgie	1966	51	24	20	21	23	12	1968	56	24	40
Bulharsko	1967	50	23	23	23	22	9	1965	58	46	52
Dánsko	1966	50	24	23	18	23	12	1960	64	28	46
NSR	1966	52	23	22	20	23	12	1967	60	30	45
NDR	1967	54	23	21	18	23	15	1964	60	40	49
Finsko	1965	52	27	25	19	21	8	1960	58	35	46
Francie	1966	51	24	22	20	21	13	1967	57	26	41
Řecko	1967	51	25	23	22	21	9	1961	60	28	43
Velká Británie	1968	51	23	21	18	25	13	1966	63	33	47
Irsko	1966	50	31	21	16	21	11	1966	57	20	39
Itálie	1966	51	24	23	22	21	10	1968	56	20	37
Jugoslávie	1966	51	29	24	23	17	7	1961	60	31	45
Lucembursko	1967	51	23	20	21	24	12	1966	58	21	39
Holandsko	1967	50	28	24	18	20	10	1960	57	16	36
Norsko	1966	50	25	21	18	24	12	1968	58	19	38
Rakousko	1969	53	24	21	18	23	14	1961	61	36	48
Polsko	1967	51	29	24	21	19	7	1960	55	40	47
Portugalsko	1967	52	29	22	19	21	9	1960	66	13	39
Rumunsko	1967	51	26	23	23	20	8	1966	61	46	54
Švédsko	1967	50	21	22	18	26	13	1965	59	30	44
Švýcarsko	1967	51	23	25	19	22	11	1965	63	26	46
Španělsko	1960	52	27	23	21	20	9	1968	60	17	38
Československo	1967	51	25	23	20	22	10	1963	52	39	45
Maďarsko	1967	52	22	23	21	23	11	1963	63	33	48

Mnohonásobné procentní členění obyvatelstva podle tří hledisek (která jsou ještě dále členěna) ve sloupcích a v řádcích podle zemí. Protože jde o korektní statistiku (z mezinárodní části rakouské statistické ročenky), uvádějí se vždy i léta, ke kterým se údaje vztahují, s cílem zabránit mylným porovnáním.

V záhlaví tabulky, které uvádí obsah sloupců, je mimo jiné procento zaměstnaných rozděleno ještě na „muže" a „ženy".

V těchto dvou případech jde vlastně o tzv. *kombinační tabulky*, které jsou pro použití již přehledně členěny. Tabulky, v nichž jsou prvotní údaje členěny do všech podrobností, se naproti tomu jmenují *„prosté tabulky"*. Tabulka operující s procenty (podílové rozdělení) bude vždy „výslednou tabulkou", protože „prosté tabulky" obsahují vždy jen absolutní četnosti.

Tečka (.) v tabulce znamená, že pro tento údaj čísla nejsou nebo jsou nepoužitelná. „Nula" (0) neznamená vždy „absolutní nulu", nýbrž „nepatrný podíl" anebo „nepatrné číslo" (menší číslo než polovina v tabulce ještě uvedených číselných řádů): používá-li se v tabulce např. jako jednotka tuna a uvádí se až do prvního desetinného místa, znamená 0 „méně než 50 kg". Hodnotu „přesně nula" označujeme spíše pomlčkou (—), někdy však rovněž pomocí nuly.

Vysvětlení značek používaných ve Statistické ročence ČSSR

— znamená, že jev se nevyskytoval

0, 0.0, 0.00 značí více než nulu, ale méně než nejmenší jednotku vyjádřitelnou v tabulce

údaj není k dispozici nebo je nespolehlivý

× zápis není možný z logických důvodů

\+ (před číslicí) — údaj je odhad

* (před číslicí) — předběžný údaj

[] údaj se vztahuje na předválečné území příslušného státu

Není-li v hlavičce tabulky nebo v její legendě uveden místní údaj (země, kraj), týkají se data celého území Československé socialistické republiky (v datech za předválečná léta bez Podkarpatské Rusi, není-li uvedeno jinak).

Ukazatele vyjádřené v peněžních jednotkách za období 1948—1953 byly kromě zahraničního obchodu přepočteny v poměru 1 : 5, v zahraničním obchodě v poměru 1 : 6,94.

5.2 Dotazníky a dotazování

Pozorováním, sčítáním a měřením mohou statistikové získat jen zlomek číselného materiálu. Stále znovu se ukazuje jako nezbytné použít anketního šetření. Pro rozlehlé oblasti průzkumu trhu a veřejného mínění je to samozřejmé, neboť trh se skládá z více nebo méně koupěchtivých lidí, jejichž ochota kupovat se má prozkoumat, a mínění je zcela subjektivní představa, poněkud měřitelná jen tehdy, když se projeví určitou akcí (např. odevzdáním hlasu při volbách).

Staré vtipné úsloví říká, že člověku byl dán dar řeči proto, aby mohl skrýt své myšlenky — a při dotazování se toto rčení potvrzuje více, než je zúčastněným osobám milé. Nejdříve svou zvědavost maskují tazatelé pokud možno bezelstně znějícími otázkami a pak někdy na své rafinovaně vychytralé formulace dostávají nejasné, nesprávné nebo vědomě nepravdivé odpovědi. K tomu vedou rozmanité pohnutky, kterými se na tomto místě nemůžeme podrobně zabývat. Chceme jen naznačit typické reakce: je to nedůvěřivá obrana jednak proti oficiálním úředním dotazům: „z toho určitě nekouká nic dobrého", jednak vůči soukromým průzkumům mínění: „co je jim do toho?" Je proto až dojemně nesmyslné, když všechny statistické úřady s oprávněným přesvědčením ujišťují, že prý údaje nebudou nikomu přístupné: rád bych viděl *toho* obchodníka, který do statistického dotazníku uvede vyšší obrat nebo zisk, než jaký uvedl v daňovém

Tazatel s dotazníkem je stále ještě překvapující, ne však vždy vítanou návštěvou. Ale odpovědi, které přinese pořadatelům průzkumu, jsou většinou přece jen důležitou pomocí a — přes nízký počet dotazovaných — většinou velmi charakteristické pro základní soubor, z něhož se vytvořil výběrový soubor dotazovaných.

přiznání. Pro občana je úřad jako úřad a ani žádný statistický úřad není výjimkou. Ovšem — úřadu je *nutno* odpovědět, třeba i nesprávně.

Odpovídat se *nemusí* při *soukromých* průzkumech — a přece se většinou odpovídá. A někdy se šidí z tisíce a jednoho důvodu a jindy zcela bezdůvodně. O tom pěkně napsala jedna švýcarská novinářka: „Předevčírem jsem překvapila naši nejmladší, když vyplňovala dotazník o laku na vlasy. Zuzce je 12 let... má velmi pěkné vlasy, které vůbec nepotřebují zpevnit pomocí laku... Zuzana uvedla, že prý je mladá dívka na zkušené v cizí rodině a potřebuje lak na vlasy především večer, když jde za zábavou... Musela jsem se smát, neboť jsem si vzpomněla na testy, kterých jsem se zúčastnila jako mladá hospodyně: »Kupujete punčochy za 2—3 franky, za 3—4 atd.?« Pro slavnostní příležitosti jsem kupovala punčochy za 5,90 Fr. To jsem také uvedla do dotazníku. Že jsem kupovala na všední den punčochy za 1,90 Fr, jsem stydlivě zamlčela, protože jsem se domnívala, že se to pro ženu mého postavení nehodí."

Velký počet dotazovaných má sklon odpovídat *podle domnělé* normy, jiní se *uchylují do extrémů*. Jeden důchodce si libuje v líčení své chudoby ve zvlášť křiklavém světle, jiný se ji snaží zastřít. Jeden obchodník vystupuje skromně, jiný, se stejným příjmem, fantazíruje o svém bohatství atd.

Tím se také stává, že třeba údaje získané o výdajích na péči o tělo a hygienu daleko přesahují skutečnost: toaletní mýdlo se spotřebovává v takovém množství, o kterém výrobci mýdla mohou jen snít. V mnoha případech jsou odpovědi nepravdivé nejen

proto, že dotázaný se chce ukázat v lepším světle, ale často také z nevědomosti spojené s upřímnou ochotou k spolupráci: mladý, příjemný muž tu stojí s tužkou a dotazníkem a klade jednoduchou otázku — v takovém případě se přece musí něco odpovědět, nechce-li člověk vypadat hloupě nebo nezdvořile. Tak se např. ptá: „Kolik kousků mýdla kupujete měsíčně?" — nebo: „Jak často chodíte v průměru do divadla?" — nebo: „Jste pro pomoc rozvojovým zemím?" — nebo: „Jste spokojen(a) s výsledky svých dětí ve škole?"

U mýdla přece není možno odpovědět „nevím". Tak se řekne dva, ne raději čtyři: ano, tak čtyři kousky. Do divadla? Posledně jsme byli — kdy jen to bylo? Nahlížíte do dotazníku, kde je též rubrika „jednou měsíčně". Ano, jednou měsíčně. Pro pomoc rozvojovým zemím? Na to se ovšem nedá tak lehce odpovědět, ale zásadně není přece možno nechat ty chudáky umřít hladem: ano, jsem pro. Výsledky dětí ve škole? Copak to znamená „spokojen"? Trošku pilnější by být mohly, ale pak to bude vypadat, jako by děti byly hloupé. Tedy: ano, spokojen.

Spokojeně odchází také tazatel s nesprávnými, neúplnými nebo mylnými odpověďmi, které pak dostane trpělivý zpracovatel údajů. Někdy záleží přímo na tazateli, jaké odpovědi dostane, i když mu dá jeho ústav příkaz neodchylovat se od určitého znění otázek a vyhnout se každému nápadnému zdůrazňování nebo reagování. Tak během druhé světové války bylo provedeno mezi americkými černochy šetření, zda se domnívají, že vítězství Japonců nebo Němců by zhoršilo postavení amerických černochů. Jestliže tuto otázku kladli bílí tazatelé, byli téměř všichni černoši přesvědčeni, že by vítězství mocností Osy bylo velkým neštěstím. Byl-li však tazatelem černoch, podstatně větší podíl barevných odpověděl: „nezměněn", „lepší" nebo „nevím".

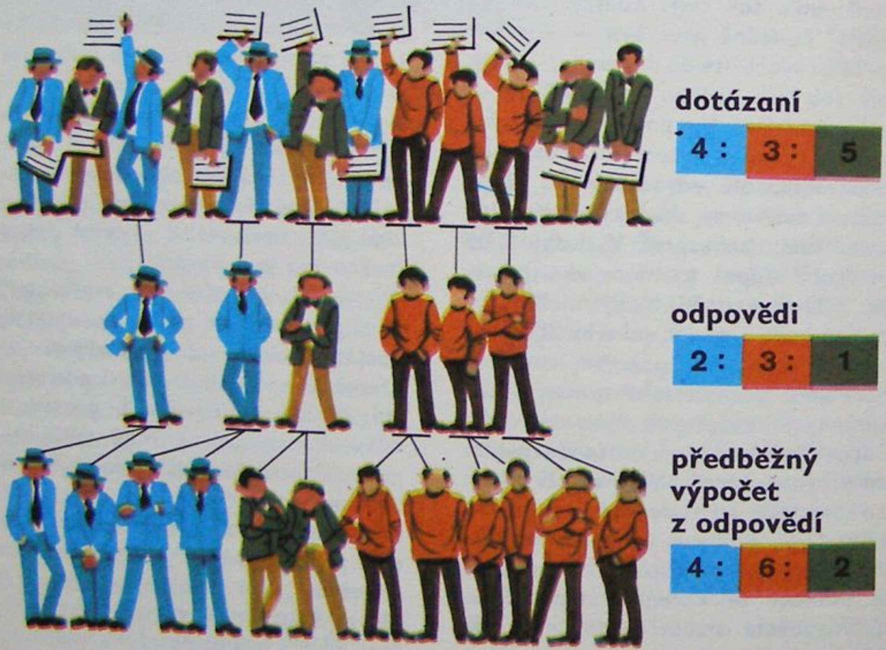
Osobní vliv tazatele se zpravidla projeví tak, že se dotazovaný cítí volněji a odpovídá ochotněji, případně je nevrlý a chce mít rozhovor co nejdříve za sebou. Tazatel je jinak relativně bezvýznamný prostředník, vlastně „lidský dotazník". O hodnotě nebo bezcennosti šetření se proto zpravidla rozhoduje většinou již při *vypracovávání dotazníku:* složité, špatně pochopitelné nebo nedůvěru budící otázky, dotěrné znění nebo rozvláčné formulace mohou šetření již předem odsoudit k neúspěchu. Avšak i příliš přímé kladení otázek může vést k chybným závěrům, zejména jde-li o téma, kterým se většina dotazovaných doposud pravděpodobně nezabývala. V takovém případě se doporučuje dané téma vhodnými předchozími otázkami „uvést" — avšak pokud možno bez ovlivňování.

Toto „*pokud možno bez ovlivňování*" platí přirozeně jen tehdy, sleduje-li se dosažení pokud možná vědecky co nejpřesnějšího výsledku. Jsou-li potřebné jen číselné podklady „k provedení důkazu", existují nesčetné možnosti, jak dosáhnout žádoucího výsledku. Pro ilustraci malý příklad, který rozhodně nepatří mezi extrémy a lze jej v jisté míře pochopit a omluvit. Velká letecká společnost uveřejnila ve své reklamě větu: „Tyto četné výhody vysvětlují, proč 97,5 % našich pasažérů vyslovuje svou spokojenost."

Nutno však s uspokojením zdůraznit, že k tomuto konstatování byla připojena tato poznámka: „Na základě šetření, které provedl profesor X (jméno je

uvedeno), prohlásilo 97,5 % z 1500 cestujících, že jsou spokojeni nebo velmi spokojeni; jen 2,5 % projevilo nespokojenost."
Nuže, jestliže jste přistáli na místě určení jen poněkud přesně, stěží řeknete, že jste nespokojen, nýbrž rozhodnete se pro střední z nabízených možností (velmi spokojen — spokojen — nespokojen), ledaže by vám letuška vylila kávu na šaty. Ostatně by se mohlo také říci: „V posledních týdnech si nikoli méně než 375 cestujících stěžovalo na servis."

Studie o technice zjišťování a psychologii dotazování mohou zaplnit celé svazky (a také už mnoho svazků bylo o tom popsáno). Kdo se mimořádně zajímá o tento speciální obor, může použít bohatou odbornou literaturu.
Na závěr chceme jen uvést, že dotazování je spojeno s určitým prvkem nejistoty, který nelze přesně zachytit ani pomocí počtu pravděpodobnosti, ani jinými matematickými metodami. Jak velké může být rozpětí, ukazuje mimo jiné příklad uváděný Elizabeth Noellovou-Neumanovou, která jistě nemá

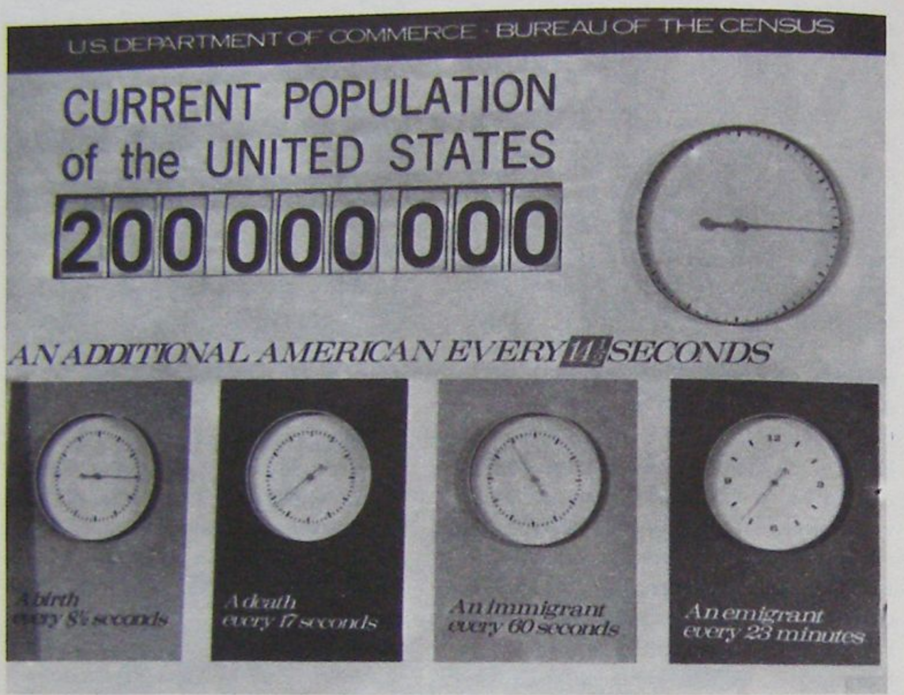


Dotazníková akce může poskytnout nesprávné informace zvláště tehdy, jestliže určitá skupina dotazovaných odpovídá neochotně, zatímco druhá skupina odpovídá velmi ochotně. V našem zjednodušeném schématu odpovídají všichni tři muži s roláky, ale jen jeden z pěti mužů s motýlkem. Jestliže se tento výsledek dotazníkové akce propočítá na základní soubor, vznikne obraz, který vůbec neodpovídá skutečnosti.

zájem na tom, aby průzkum mínění byl postaven do nesprávného světla.[1] Na otázku, zda by všichni dělníci v podniku měli být v odborech, odpovědělo „ne" jednou 20 %, zatímco v druhém případě 70 %; obě skupiny byly stejnorodé, avšak pojetí otázky bylo různé. V prvním případě otázka zněla: „Myslíte si, že všichni dělníci v podniku by měli být v odborech?" V druhém případě byla však rozšířena takto: „...nebo se musí každému ponechat k rozhodnutí, zda chce být v odborech nebo ne?" Druhou otázkou se téměř sugestivně vyvolává problém osobní volnosti, který v kratším znění otázky vůbec nezaznívá, a tak došlo k uvedenému výraznému rozdílu (tento příklad je charakteristický i pro „*zkušební dotazování*", jež se provádí při přípravě dotazníků, aby se našla pokud možno jasná a věcná formulace).
Pokud jde o dotazníky, slouží při šetření pomocí tazatelů především jako „texty", z nichž se může klidně a plynně předčítat. Jinou metodou dotazování je rozesílání dotazníků s prosbou o jejich vrácení (poštovné platí příjemce). Proti tomuto způsobu dotazování se útočí pádným argumentem, že zkušenosti ukazují, že většina těchto dotazníků buď z nepozornosti, nebo lhostejnosti putuje do koše na papír. Odpoví-li však jen zlomek dotázaných, není již výběrový soubor dostatečně reprezentativní. Pak není možno dělat závěry pro základní soubor nebo dokonce pro celé obyvatelstvo.
Přesto může být akce s dotazníky ve skutečnosti užitečná. Dá se totiž předpokládat, že lidé, kteří si dali práci s vyplněním a vrácením dotazníku, se

[1] E. Noellová, Výzkum veřejného mínění, Svoboda, Praha 1968.

o dané téma opravdově zajímají. Dostali otázky, jimiž se už zřejmě sami zabývali a ke kterým chtějí říci svůj názor právě proto, že je živě zajímají nebo že jsou dobře informováni. Dovíme se tedy sice málo nebo téměř nic od „mlčenlivé většiny", avšak velmi mnoho o mínění „opinion leaders", neformálních tvůrců veřejného mínění.
Z podobných důvodů ztrácí také na závažnosti argument, že při vyplňování dotazníku spolupůsobí radou celá rodina, čímž nevzniká hledané mínění jednotlivce, nýbrž shodný anebo kompromisní názor celé rodiny. Právě tím se však odpověď koneckonců dodatečně přibližuje skutečnosti: kdo se o dotazníku radí s rodinou, udělá to pravděpodobně i v případě, bude-li muset stejnou otázku řešit v denním životě. Přesto je nutno s veškerým důrazem opakovat: opravdu reprezentativní šetření pomocí rozeslaných dotazníků se skoro nikdy nepodaří.
Zmiňme se ještě o posledním způsobu zjišťování údajů, o tzv. *běžném zjišťování*. Spočívá v tom, že hromadná šetření, např. sčítání lidu, jsou stále doplňována běžným hlášením. Tak je stálý soubor „obyvatelstvo" doplňován „běžnými zápisy", a to tak, že se běžně zjišťují údaje pohyblivých souborů „narození" a „úmrtí". Během času jsou však „běžené záznamy" stále méně přesné (zejména zůstala-li v nich skryta systematická chyba), přesto však po značnou dobu poskytují spolehlivé odhadové hodnoty. Mechanismus „běžných zjišťování" byl dokonce povýšen na pamětihodnost: ve Washingtonu se na prosvícené tabuli promítají ve dne v noci „běžné záznamy" o počtu obyvatel USA tak, že za určitý počet vteřin se počet obyvatelstva zvyšuje o jedničku. Vypadá to velmi pěkně, avšak dříve nebo



Plynulé „běžné záznamy“ obyvatelstva ukazují (kalendářní, digitální) hodiny v „cenzovním úřadu“ ve Washingtonu. Každé 14 $^1/_2$ sekundy poskočí o 1 a zaznamenají tím přírůstek obyvatelstva. Obrázek zachycuje „historický okamžik“, kdy bylo zaznamenáno 200 miliónů obyvatel a je ovšem, jako každé jiné uváděné číslo, konec konců jen optickým gagem. Ani v den sčítání lidu nelze udat počet obyvatelstva s přesností na 10 000, nemluvě už vůbec o „běžných záznamech“, založených na odhadech a vývojových tendencích minulosti.

později jsou údaje nesprávné řádově o desetitisíce — a každých deset let po hromadném sčítání lidu se musí tyto hodiny s „běžnými záznamy“ obyvatelstva postrčit nazpět nebo dopředu.

5.21 Z dějin sčítání lidu

Při slově „statistika“ se člověku může vybavit všelicos — samozřejmě čísla, vzorce, normální křivka, rčení o statistice jako nejhorší formě lži, průměry, index, spotřební koš nebo dotazníky.

Na jedno se však jistě nevzpomene, ačkoliv by to bylo velmi oprávněné: na vánoce.

Zalistujeme-li v *Lukášově Evangeliu*, můžeme se dočíst, jak se stalo, že se Kristus narodil v Betlémě ve stáji: „I stalo se v těch dnech, vyšlo poručení od císaře Augusta, aby byl popsán všechen svět... I šli všichni, aby zapsáni byli, jeden každý do města svého. Vstoupil pak i Josef od Galilee... do města Davidova, kteréž slove Betlém (proto že byl z domu a z čeledi Davidovy), aby zapsán byl s Marií, zasnoubenou sobě manželkou, těhotnou. I stalo se, když tam byli, naplnili se dnové, aby porodila.“ (2, 2—6.)

Nařízení o sčítání lidu pro celou říši římskou za císaře Augusta, při němž měl být každý sčítán ve svém domovském místě, nebylo však ani první v říši římské, ani první, o kterém je zmínka v bibli — a už vůbec ne první v dějinách lidstva. Již v 6. stol. př. n. l. pamatovala *ústava krále Servia Tullia* na periodické sčítání všech občanů pro jejich zařazení do příslušných daňových skupin. Také v *Helladě* bylo členění do *cenzovních tříd* zcela obvyklé. (Postupně došlo k posunutí významu slova cenzus, které dnes, zejména v angloamerické oblasti, označuje „sčítání lidu“ a jeho původní fiskální akcent není již patrný.)

Sčítání lidu má v bibli tak významnou úlohu, že čtvrtá kniha Mojžíšova se jmenuje „Numeri“, protože se v ní od samého začátku podrobně mluví o „přehlídce národa“. Jsou v ní dokonce jména prvních cenzovních úředníků: Elisur, Salamiel, Názon atd., po jednom příslušníku dvanácti kmenů izraelských pod vrchním vedením Mojžíše a Arona: „A shromáždili všecko množství prvního dne měsíce druhého, kteříž přiznávali se k rodům svým, po čeledech svých, po domích otců svých, a vedle počtu jmen, od dvadcítiletých a výše po osobách svých.“ (IV. Mojžíš, 1, 18.)

Jestliže bylo toto sčítání lidu právem nařízeno od samého Boha, nechal se král David později svést satanem, aby nařídil z vlastní pravomoci sčítání lidu, jehož vylíčení najdeme ve 2. Knize Samuelově a v 1. Knize Paralipomenon. Když se ovšem porovnají tyto údaje navzájem, ukáže se již zřetelně, že na tomto sčítání byla kletba (která někdy vrhá svůj stín až do naší přítomnosti) — čísla navzájem nesouhlasí. Jestliže se v druhé Knize Samuelově říká: „I dal Joáb počet sečteného lidu králi. Bylo pak lidu Izraelského osmkrát sto tisíc mužů silných, kteříž mohli vytrhnouti meč k bitvě. Mužů také Judských bylo pětkrát sto tisíc“, v první Knize Paralipomenon se uvádí: „A bylo všeho lidu Izraelského jedenáctekrát sto tisíc mužů bojovných, lidu pak Judského čtyřikrát sto tisíc, a sedmdesáte tisíc mužů bojovných.“

Tak ale na tomto sčítání ležela kletba a Bůh dal Davidovi na vybranou buď sedm let hladomoru, nebo tři měsíce útěku před nepřáteli, anebo tři dny moru. David se rozhodl pro mor: „... a padlo jich z Izraele sedmdesáte tisíc mužů.“ (I. Paralipomenon 21, 14.)

Toto biblické varování jako by určilo poměr ke statistickému zjišťování až do nové doby. Byla to nepochybně trestuhodná zvědavost a všetečná opovážlivost pokoušet se nahlížet pomocí sčítání lidu nebo dokonce soustavným pozorováním narození, nemocí a úmrtí do nevyzpytatelných záměrů božích.

Ve středověku se proto šetření omezují téměř výlučně jen na soupisy majetku. Tak Karel Veliký (nebo Ludvík Němec) v „*Capitulare de villis*“ nařídil přesné zjištění stavu královských domén a velkých soukromých statků a o 250 let později dal Vilém Dobyvatel sestavit „*Domesday Book*“, kterou je možno označit za první pozemkový katastr v západních zemích. Její jméno, odvozené od anglosaského „domes daeg“, znamená „soudný den“ nebo také „den posledního soudu“ (v tomto významu používá angličtina ještě dnes slovo „doomsday“). Výklad se různí. První význam by ukazoval na nepodplatitelnou přesnost knihy při projednávání hraničních sporů, druhý na ničím nepřekonaný smysl pro spravedlnost, který z něho vyplývá. A je to opět Anglie, kde se na počátku 16. století začíná na příkaz lorda kancléře Thomase Cromwella (vzdáleného příbuzného Olivera Cromwella) soustavně zaznamenávat narození a úmrtí v církevních matrikách — a opět vzniká pokušení pomyslit na působení staré kletby, protože Thomas Cromwell byl obviněn z velezrady a popraven. (Protože však Jindřich VIII. nechával s oblibou popravovat své ministry a své manželky, není toto úmrtí nijak nápadné.)

V 17. století začal konečně anglický obchodník se suknem John Graunt jako první porovnávat matriky narozených a zemřelých. Počet těchto matrik byl již tehdy velký, i když mnohé byly značně neúplné. V roce 1622 vyšla jeho práce „*Natural and Political Observations upon the Bills of Mortality*" (Přirozená a politická pozorování o seznamech zemřelých v městě Londýně), první dílo o statistice obyvatelstva.

Věk začínajícího merkantilismu a osvíceného absolutismu urychlil další zkoumání o struktuře obyvatelstva proto, že rychlý růst obyvatelstva byl pokládán za záruku vzkvétajícího hospodářství. Toto spojení teorie o státě, dříve hospodářské vědy a matematiky, je charakteristické pro počátky dějin statistiky. Vysvětluje i označení „*politická aritmetika*", které se od konce 17. století vždy znovu používá — poprvé je v roce 1687 použil anglický národohospodář sir William Petty v práci „*Five Essays on Political Arithmetic*" (Pět esejů o politické aritmetice).

Mezitím byly také v kontinentální Evropě shromážděny podklady pro statistiku obyvatelstva. Poprvé je vyhodnotil rovněž Angličan, proslulý astronom Halley, přítel de Moivra. Halley napsal v roce 1693 knihu „*An Estimate of the Degrees of Mortality of Mankind, drawn from curious Tables of the Births and Funerals at the City of Breslau, with an Attempt to ascertain the Price of Annuities upon Lives*" (Zjišťování stupňů pravděpodobnosti úmrtí lidí, sestavené podle tabulek narození a úmrtí města Vratislavi, spojené s pokusem zjistit výši pojistného pro životní pojištění). Tímto matematicko-pojistným dílčím hlediskem počátků statistiky se budeme ještě zabývat na jiném místě.

V německy mluvící oblasti to byl konečně pruský polní kazatel Johann Peter Süssmilch, který pomohl statistice obyvatelstva k vítězství a zároveň našel cestu k odstranění biblické kletby tímto výkladem: „Takové údaje dovolují zvlášť zřetelně poznat moudrý pořádek, kterým Bůh uspořádal pozemské záležitosti." Proto také jeho kniha — první dílo o statistice obyvatelstva v němčině — nese název „*Die göttliche Ordnung in den Veränderungen des menschlichen Geschlechts aus der Geburt, dem Tode und der Fortpflanzung desselben erwiesen*" (Božský pořádek o proměnách lidského rodu prokázaný z jeho narození, smrti a rozmnožování) (z roku 1741, později často znovu vydávaná).

Také Süssmilchovo dílo je založeno na údajích církevních matrik, které mu pastoři asi tisíce kuronských obcí dali k dispozici. Ale již za několik let později začala své vítězné tažení i úřední statistika. V roce 1741 byla ve Švédsku zřízena „*tabulková komise*", jejímž úkolem bylo sestavovat podklady pro statistiku obyvatelstva. Biblická kletba byla překonána, zvědavosti státu nestálo již nic v cestě.

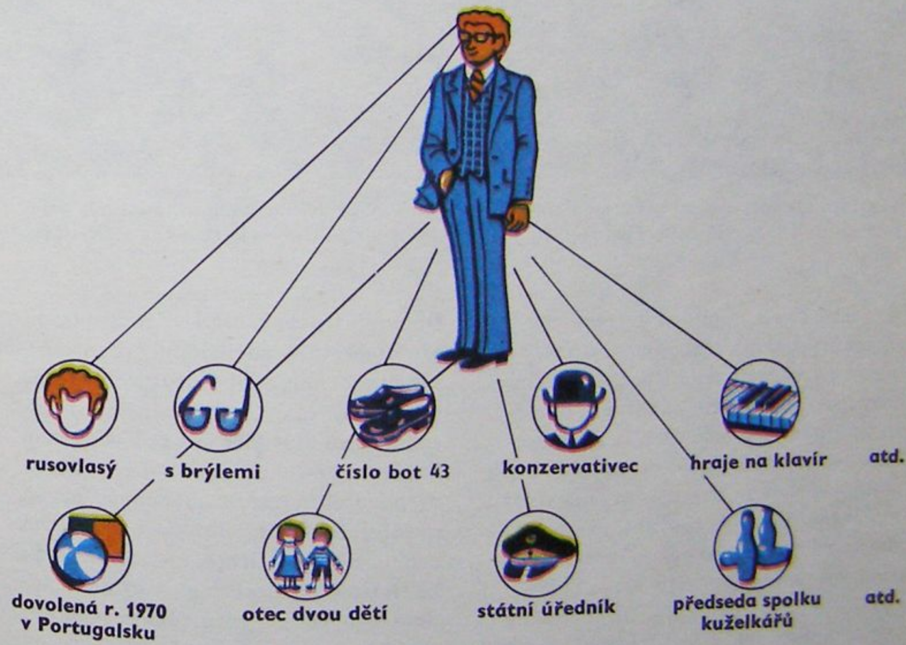
Spojené státy provádějí jako první

země na světě již od svého založení pravidelná sčítání lidu. Již v roce 1790, rok po zvolení Washingtona prvním prezidentem mladého státu, se konalo první z těchto sčítání. Byly zjištěny sotva 4 milióny obyvatel v tehdy 17 státech Unie. Rytmus sčítání v zaokrouhlených desetiletích se prosadil i mezinárodně. Naposledy se konalo hromadné sčítání obyvatelstva v roce 1970.

5.3 Reprezentativní průřez obyvatelstvem

Chceme-li vědět, jak chutná víno uložené v hektolitrovém sudu, nemusíme hned vypít celý sud — již malý doušek stačí k jeho posouzení. Chceme-li zjistit, zda náklad ořechů v nákladním autě je z velké části zkažený, postačí, když vezmeme pár ořechů z různých míst nákladu a rozloupneme je. Na základě určitého rozsahu vzorku je možno dělat použitelné závěry o jakosti celého nákladu.

Ale jak tomu je, když mají být dotazováni, testováni a zařazováni lidé? Nedělí se přece jako ořechy jen na dobré a špatné: lidé zabydlují svět jako čtyři miliardy individuí, navzájem lehce rozlišitelných, většinou již zevnějškem (pravých dvojníků je málo), ale daleko více diferencovaných vnitřně a způsobem myšlení. Cožpak je možné odvážit se převést je na společného jmenovatele



Každý člověk má mnoho „úloh", je nositelem mnoha znaků. Statistika může pomocí malých výběrových souborů vypovídat o četnosti znaků a jejich kombinaci, ačkoliv všechny ostatní znaky nejsou pro toto šetření vůbec potřebné. Abychom se dověděli, kolik státních úředníků nosí brýle, nemusíme se jich dotazovat na barvu vlasů, na cíl dovolené nebo na číslo bot.

I když je „základní soubor“ ještě mnohem větší než tato masa lidí na stadiónu, musí mít každý jednotlivec stejnou možnost být zahrnut do vzorku, neboť jen tak se může skutečně počítat s náhodou.

a vtěsnat do společného schématu? Z takových úvah o osobnosti a individualitě vycházejí mnohá pochybovačná tvrzení o statistice všeobecně a o průzkumu mínění zvlášť. Sčítání obyvatelstva — budiž, ale není to domýšlivé a pošetilé strkat desetitisíce individuí do jednoho pytle, spočítat je pomocí *jednoho* dotazníku? A pak se jeden nahodile vytáhne a pokládá se za „mluvčího“, který zastupuje stovky nebo tisíce jiných lidí, s nimiž možná nemá společného nic jiného než lidskou postavu! Jak může být zjištěno *mé mínění, mé chování* dotazováním úplně cizího člověka?

Domnělý rozpor mezi růzností lidské přirozenosti a radikálními rovnostářskými machinacemi statistiky se začne rozplývat, jakmile porovnáme např. známky na vysvědčeních se zjišťováním kvality ořechů. Žádný učitel, žádný vychovatel a ani rodiče nepochybují o tom, že každé dítě je samostatná osobnost — a přece nikomu nevadí, že chování, znalost angličtiny nebo nadání pro ruční práce těchto jednotlivých individuí se paušálně zařazují do stupňů „výborný“ až „nedostatečný“.

Podobně se mohou vyskytnout ženy se svérázným přirozeným odstínem

vlasů, celkem však nikdo nebude mít námitky proti tomu, jestliže vytvoříme čtyři nebo pět skupin: blondýny, brunety, černovlasé, rusovlasé a případně šedovlasé.

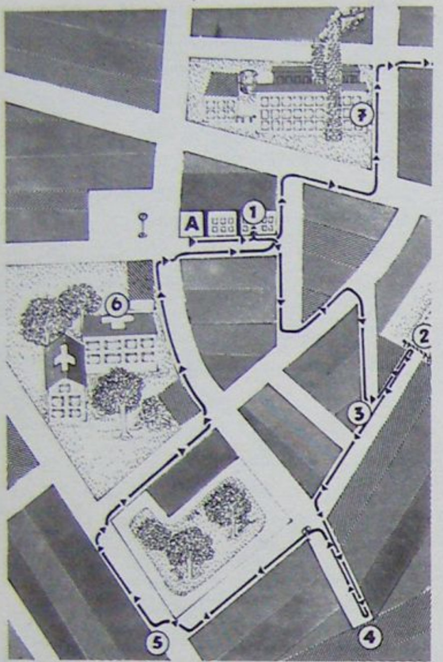
Sociologie pokládá člověka za nositele řady rolí — má např. úlohu „otce“, „příslušníka určitého národa“, „úředníka“, „sportovního rybáře“, „katolíka“ atd. Jejich vyjmenováním se vytvoří celková osobnost.

Pro potřeby statistiky však tento „rozkladný proces“ pokračuje dál — dokonce tak daleko, že člověk zcela zmizí: objevuje se jen v jednotlivých případech jako nositel znaků: jako příjemce 25 000 až 30 000 Kčs ročního příjmu, jako návštěvník sportovních stadiónů, jako televizní divák, jako stoupenec trestu smrti, jako rekreant v Jugoslávii, jako oběť nakažlivé nemoci. Vše ostatní, čím ještě je, není zajímavé, jeho anonymita zůstává zachována nejen z ohleduplnosti, ale hlavně z důvodů účelnosti a nutnosti. Statistická šetření jsou složitá, časově náročná a dosti nákladná, i když se soustředí na nejdůležitější. Proč tedy ještě přidávat statisíce činitelů, které ve svém celku tvoří individuum?

Právě proto na tom není nic ponižujícího ani metodicky nepřípustného, jestliže se jednotlivé osobnosti přece jen hodí do jednoho pytle — totiž jako nositelé několika málo z těch 10 000 a více znaků, které mají. K tomu, abychom se dověděli o způsobech pití vína nebo zvyklostech o dovolené, o oblíbených jídlech nebo koníčcích celého národa, zpravidla postačí, jestliže se zeptáme tří až čtyř tisíc lidí. Tyto osoby musí ovšem vyhovovat jediné podmínce: musí být určeny zcela náhodně nebo přinejmenším za rozhodujícího spolupůsobení náhody.

„Náhoda“ v tomto případě znamená: *každá jednotka základního souboru musí mít stejnou naději dostat se do výběrového souboru.* Má-li být dotazováno bydlící obyvatelstvo Paříže, pak musí mít domovnice ve staré postranní uličce v Belleville stejnou naději, že jí budou kladeny otázky, jako student z Madagaskaru ve svém pokojíku za Pantheonem nebo jako penzionovaný bankovní úředník na kraji boulońského lesa. Jestliže všichni němečtí vinaři tvoří základní soubor, který má být dotazován, nestačí zajet do údolí Mosely a tam se zeptat deseti vinařů, které náhodou zastihnu doma. Jde-li o mínění švýcarských dělníků, nestačí počkat u vrat továrny v Basileji a po skončení pracovní doby pohovořit s hrstkou lidí, kteří jsou ochotni odpovídat. Stejně tak mínění rakouského obyvatelstva nezjistím na schodišti vídeňského Burgtheatru.

K plnému uplatnění náhody by bylo nutné, aby existoval úplný seznam souboru, který má být dotazován a z kterého by se pak nějakým nahodilým způsobem „vylosovalo“ určité procento nebo promile. Takto vybrané osoby (a jen tyto) by musely být dotázány — a to bez výjimky. V praxi je takový postup zpravidla neproveditelný, protože nejsou k dispozici uvedené seznamy a také proto, že náklady takové dotazníkové akce by byly příliš vysoké. Používají se tedy smíšené náhodné systémy, např. *dvoustupňový výběr.* V daném případě se „nelosují“ jednotlivé osoby, ale skupiny osob: do vzorku se nezahrnuje např. 2000 švýcarských dělníků jako jednotlivců, nýbrž 100 švýcarských továren a závodů, ve kterých je pak — opět náhodným výběrem — dotazován potřebný celkový počet 2000 dělníků. Vzorek je možno dál



Plánek přiložený k pokynům pro tazatele, určující postup oblastního výběru. Ukazuje, jak použít při postupu od výchozího bodu A základního pravidla pro výběr domů („zahnout střídavě vpravo a jít po pravé straně ulice a pak zahnout vlevo a jít po levé straně ulice“) v případě, že zelené plochy, slepé uličky atd. vyvolávají potíže. Druhé pravidlo tohoto postupu zní: „navštívit každou třetí domácnost na této trase“; „počítat od přízemí nahoru až do nejvyššího obývaného poschodí a — je-li v tomtéž poschodí více nájemníků — zleva doprava“. (Z pokynů pro tazatele Ústavu pro demoskopii v Allenbachu.)

„rozvrstvit“, abychom se ubezpečili, že jsou zastoupeny všechny dílčí jednotky. Jestliže např. máme zjišťovat investiční záměry „podniků“, měly by velké koncerny, na jejichž investičních záměrech záleží především, při čistě náhodném výběru pouze stejnou naději na vylosování jako každá malá obuvnická dílna nebo zmrzlinářský stánek. Utvoříme tedy tři nebo více „vrstev“: a) *velké podniky*, b) *střední podniky*, c) *malé podniky*, a na tomto základě zajistíme, aby byli dotazováni reprezentanti všech skupin. Přitom bude dotázáno, bude-li to třeba, všech 20 velkých podniků, 50 ze středních podniků a 100 ze 400 malých podniků. Jak budou jejich odpovědi „váženy“, závisí už na pojetí otázky.

Při průzkumu veřejného mínění se používá tzv. *kvótní výběr (quota sampling)*, který má s výše uvedeným systémem mnoho společného, i když nejsou vůbec totožné. V daném případě se obyvatelstvo „rozvrstvuje“ podle několika standardních znaků a tazatel dostane např. tento příkaz: „Dotazujte se 20 lidí starších 16 let, z toho 11 žen a 9 mužů, pěti mladších 30 let, deseti ve věku 30—60 let a pěti starších. Čtyři osoby mají být samostatně činné, ostatní z poloviny zaměstnanci, případně mimo výdělečnou činnost.“ Tyto „rozvrstvovací“ příkazy se mohou měnit v závislosti na tématu dané otázky.

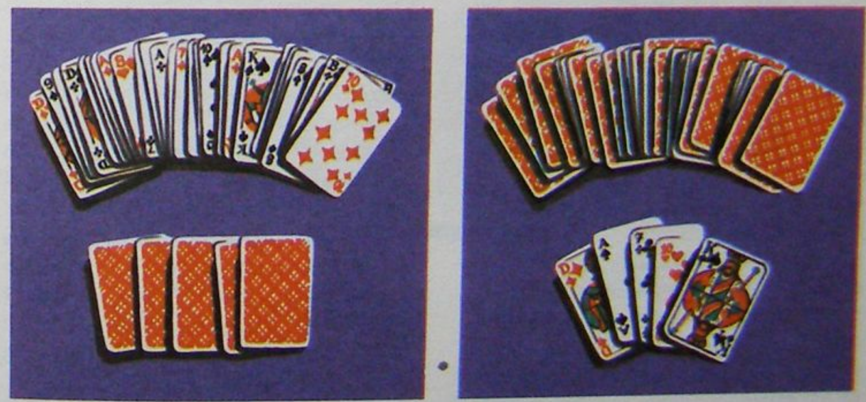
Téměř čistý *náhodný výběr (random sampling)* spočívá v tom, že se náhodně určí domy ve dvou tuctech míst, kam se vyšlou tazatelé s předepsaným směrem cesty, který vypadá jako itinerář rallye „lov na lišku“: „Pravá strana ulice, druhý dům, první patro vlevo, dotázat se dospělé ženy; potom čtyři domy dál, druhé patro, třetí dveře vpravo, dotázat se hlavy domácnosti, potom nejbližší příčná ulice vlevo...“ atd.

Možnými zdroji chyb při použití těchto metod se nebudeme dále zabývat. *S výjimkou rozvrstveného náhodného vzorku jsou všechny méně spolehlivé než čistý náhodný výběr*, podle zkušeností však stále ještě dostačují k tomu, aby se získaly použitelné výsledky.

Proč se vůbec přistupuje k vrstvení a tak namáhavě se konstruuje náhodný výběr, když ještě na počátku této kapitoly jsme tvrdili, že jde jen o jednotlivé znaky, a nikoliv o individua? Proč se přesto provádí rozdělení na muže a ženy, na mladé a staré, na samostatně činné a zaměstnance, a je-li třeba dokonce na katolíky a protestanty? Ještě jednou připomínáme: Teoreticky nejlepší výběrovou metodou je čistý náhodný výběr, který toto rozdělení nezná. Rozdělení se používá proto, aby další dobré metody mohly proběhnout v jisté míře pod vlivem náhody, aby náhoda byla jaksi synteticky vytvořena. Bez uvedených pokynů by bylo tazateli ponecháno na vůli, *zda* se přece jen nepostaví k vratům továrny nebo na stanici pouliční dráhy a tam se bude dotazovat tak dlouho, dokud nesplní předepsaný kontingent. Všichni penzisté v parku by unisono volali: „Ne, beat nemáme rádi.“ A tazatel by po pravdě poznamenal: „Všichni dotazovaní beat rozhodně odmítají.“

Proč ale vůbec takové křečovité úsilí o skutečnou nebo alespoň podle pravdy věrně napodobenou náhodu? Protože tehdy a jen tehdy platí zákony počtu pravděpodobnosti, které jsme v nejhrubších obrysech poznali.

Je jen otázkou počtů, aby se zjistilo, jaké bude uspořádání, když se vytáhne 5 karet ze hry o 52 kartách, což bylo vysvětleno pomocí příkladů. Je možno lehce určit, jaká je mezi těmito tahy pravděpodobnost vytažení např. kombinace „tři piky a dvě srdce“, neboť

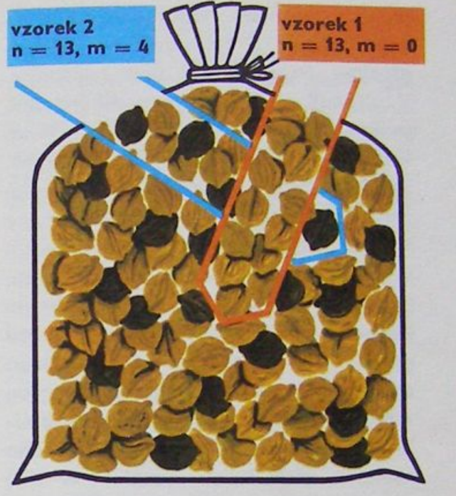


Teorie pravděpodobnosti ukazuje, jak lze ze známého základního souboru odhadnout složení neznámého vzorku: např. jaká je pravděpodobnost, že mezi pěti kartami jsou přesně dvě černé (obrázek vlevo)? Teorie výběrových šetření ukazuje na základě známého vzorku na neznámý základní soubor: např. jaká je pravděpodobnost, že v této karetní hře neznámého složení je třicet dva černých, ale jen dvacet červených karet (obrázek vpravo)?

známe skutečné rozdělení základního souboru.
V praxi statistiky a výběrových souborů se však musí postupovat obráceně: pět karet (případně vícekrát tažených) má podat zprávu o složení „balíčku karet", jehož skutečné složení neznáme a který může sestávat z 20 piků, 15 srdcí, 10 křížů a 7 kár. Mají-li v těchto případech platit zákony o pravděpodobnosti poměrů výběrových souborů, musí — jako při tažení z uvedeného balíčku karet — rozhodovat jedině náhoda. Kdybych naproti tomu nechtěl mít ve výběrovém souboru žádnou kartu s ohnutými, případně nalomenými rohy, již to by mohlo — snad — vytvořit zkreslený obraz.
V určitých případech může být účelnější zeptat se na mínění 5 odborníků než 500 laiků. Vše záleží na tom, co chci vlastně zjistit. Jestliže chci vědět, zda je Mars obydlen, zeptám se pěti odborníků na kosmický prostor a astronomů. Chci-li vědět, zda obyvatelstvo věří, že na Marsu žijí lidé, musím vytvořit náhodný výběrový soubor z celého obyvatelstva. Jestliže chci vědět, zda chlapci ve věku mezi 10 a 15 lety věří, že Mars je obydlen, musím vytvořit náhodný výběrový soubor z této věkové skupiny a bylo by zde zcela nesprávné ptát se dětí tří známých a tuctu dětí ze sousedství — výsledek by mohl být zkreslen úrovní vzdělání mých známých, prostředím mého okolí a řadou dalších faktorů. Každý osobní výběr vnáší subjektivní prvek, který si mnohdy ani neuvědomujeme, který však přesto více nebo méně vyřadí nebo přehluší zákonitosti náhody, na nichž se mají zakládat další výpočty.

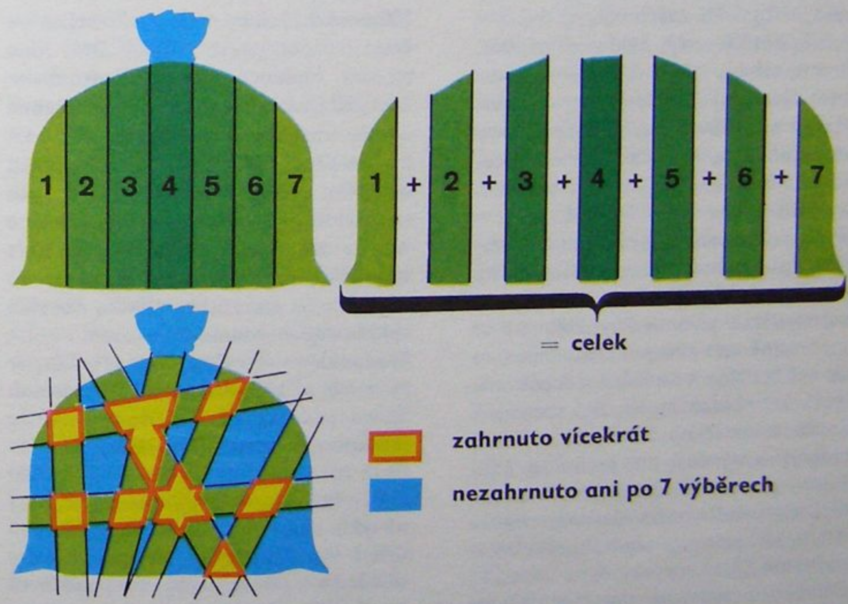
5.4 Rozsah výběrových souborů

V jedné staré anekdotě se vypráví, že jeden Angličan cestoval kdysi kočárem z Holandska do Belgie a rozhodl se přenocovat v Cáchách. Jeho věrný sluha mu našel nocleh v hostinci, a když onen Angličan ráno hodlal pokračovat v cestě, pomáhal místní zrzavý podomek nosit jeho zavazadla do kočáru. Angličan mu pak vtiskl do ruky minci a vlídně se ho zeptal na jméno. „Bedřich," řekl podomek a uklonil se. Angličan kývl hlavou, nasedl do kočáru, zatáhl záclonky a jel dál směrem na Lutych. Do svého zápisníku si však poznamenal: „Němci se jmenují Bedřich a mají zrzavé vousy."
Něco podobného se může stát, jestliže se z velmi malého výběrového souboru ukvapeně usuzuje o základním souboru. I když uvedený malý příběh působí bizarně, život často ukazuje tento všední případ: přede mnou stojí pytlík ořechů; sáhnu dovnitř, vytáhnu ořech, rozlousknu jej a je prázdný. Co z toho mohu usoudit?
Optimista řekne: „No tohle, jediný prázdný ořech, a právě ten musím vytáhnout — no budiž, aspoň jsme se ho zbavili." Pesimista řekne: „Toho jsem se obával, v pytlíku je plno zkažených ořechů." A co řekne statistik? Nejdříve přizná, že pravdu by mohl mít jak optimista, tak pesimista. Prokázáno je totiž, že v pytlíku byl jeden prázdný ořech. Mohl být jediný, mohl být jeden z několika, jeden z mnoha, jeden ze všech. (Viz obr. na str. 151.)
K získání poměrně rozumného úsudku vytvoří statistik větší vzorek — řekněme dvacet ořechů z asi tisíce, které jsou podle odhadu v pytlíku. Ořechy předtím dobře promíchá a jeden vezme shora, jeden zespodu, jeden z pravé



Vlevo:
V tomto pytli jsou velmi dobře zamíchány dobré (světlé) a špatné (tmavé) ořechy. Výběrové soubory přesto poskytují značně rozdílné výsledky.

Dole:
Jestliže se odebrané výběry nevracejí nazpátek, pak při dalším postupném odběru se základní soubor pomalu vyčerpá a výběr se nakonec stane vyčerpávajícím šetřením (nahoře). Zpravidla se však utvoří ze základního souboru více náhodných vzorků. Je bezvýznamné, že některé jednotky jsou zachyceny vícekrát a jiné vůbec ne. Průměr 7 vzorků z našeho příkladu bude základní soubor charakterizovat dostatečně výstižně a bude méně nákladný než vyčerpávající šetření.



strany, jeden z levé a jeden z prostředka, až má před sebou dvacet ořechů. Z celkového souboru o N jednotkách vzal vzorek n; v našem případě je $N = 1000$, $n = 20$.

Nyní výběr prozkoumá a zjistí, že 15 ořechů je dobrých a 5 špatných. Soustředím-li se na znak „špatný" a nositele znaku ve výběru označím m, dostanu $m = 5$, proto $n - m$ (= ne špatné ořechy) = 15, a $m/n = {}^5/_{20} = {}^1/_4$. To je můj první odhad. Ze vzorku jsem zjistil podíl špatných ořechů, a to může platit jako první odhadnutá veličina za předpokladu, že p_M, totiž dílčí pravděpodobnost špatných ořechů v základním souboru (M), by také mohla být přibližně ${}^1/_4$. (Viz obr. na str. 151 dole.)

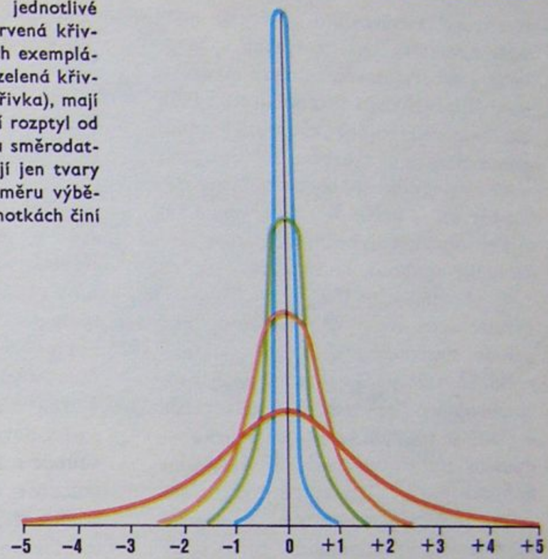
Vzorek o rozsahu $n = 200$ by snad ukázal $m = 42$, a tím odhad pro $p_M = 42/200 = 0{,}21$, který se opírá o větší rozsah výběru. Rozsah vzorku by bylo možno dál zvětšovat, až by konečně obsáhl celý základní soubor. Teprve tehdy — a ne dříve! — bychom měli absolutní jistotu o stavu všech ořechů v pytlíku. Jinak je možno podíl jen odhadovat se stále rostoucí přesností. Musíme však zavčas přestat, protože přece nemůže být úkolem zkoumání vzorků prozkoumat celý základní soubor. Míru dostatečného rozsahu výběru určuje koneckonců naše snaha o přesnost. Spokojíme-li se s obstojně spolehlivým odhadem „ne víc než čtvrtina a ne méně než pětina", může být vzorek menší, než chceme-li docílit s téměř absolutní spolehlivostí rozptyl prognózy na pouze $\pm 2\,\%$. A ještě jednou: absolutní jistotu o rozdělení, o střední hodnotě nebo o směrodatné odchylce základního souboru neposkytne *žádný* vzorek.

Dříve než začneme zkoumat, jak se navzájem chovají rozsah výběru, přesnost odhadu a základní soubor, podívejme se ještě na druhý příklad. V případě ořechů šlo o podíl jednoho (nebo více) nositelů znaků v jednom vzorku, případně základním souboru. Věda v tomto případě mluví o *„stejnoznačném kladení otázek"*. Vedle toho známe *„různoznačné kladení otázek"*, které se zabývá výběrovým souborem pro zjištění střední hodnoty a rozptylu kvantitativních znaků. Příklad: obvod prsou Quételetových vojáků. Nebo chceme odhadnout, jak vysoké jsou v průměru úspory samostatných rolníků v Bavorsku. Z našeho výběrového souboru (řekněme: rozsah $n = 500$) nevycházejí procenta, nýbrž 500 údajů o výši úspor — a protože se zde nezabýváme teorií a kritikou zjišťování, předpokládáme, že údaje jsou správné. (I kdyby nebyly, je to náš jediný výchozí materiál.)

500 dotázaných by mohlo jako celkovou částku úspor uvést 240 000 DM. Jako střední hodnotu vzorku dostaneme: $\bar{x} = 480$ DM. Mimoto získáme potřebné údaje, abychom mohli vypočítat rozptyl, a předpokládáme, že směrodatná odchylka výběru by činila 60 DM. Tuto směrodatnou odchylku výběrového vzorku označíme s, podobně jako jsme označili střední hodnotu výběrového vzorku $\bar{x}$; μ a σ zůstávají jako obvykle vyhrazeny základnímu souboru.

Předpokládejme dále, že výše úspor by měla přibližně normální rozdělení. Tento předpoklad je ovšem — a to musíme dodat — velmi smělý. Jsou to totiž zpravidla úspory (a v ještě větší míře šekové knížky), které nemají obvykle normální rozdělení: jednotlivé velmi vysoké účty zkreslují rozdělení téměř tak, jak tomu bylo u boháče ve Zbohatlíkově. Kromě toho mnoho do-

Vyjmeme-li z velkého souboru jednotlivé exempláře, mají velký rozptyl (červená křivka). Vytvoří-li se vzorky po čtyřech exemplářích (oranžová křivka), po devíti (zelená křivka) nebo po pětadvaceti (modrá křivka), mají průměry těchto výběrů vždy menší rozptyl od skutečného průměru. (Čísla nejsou směrodatnými odchylkami, nýbrž znázorňují jen tvary křivek. Směrodatná odchylka průměru výběrového souboru o pětadvaceti jednotkách činí jen $1/\sqrt{25} = {}^1/_5$ dílčího rozptylu.)



tazovaných rolníků odpoví asi „nula" nebo „nic", takže rozdělení bude „šikmé" (rozdělení L krásných věcí, vzpomínáte-li si).

Protože nám však jde především o ukázku početního postupu, zůstaneme u tohoto příkladu a uvedených předpokladů. Předpokládejme ještě základní soubor „samostatní rolníci v Bavorsku" $N = 250\,000$, a tím máme všechny údaje potřebné pro výpočet. Vzniká otázka — a teď již můžeme mluvit docela všeobecně a bez ohledu na náš poněkud „šikmý" obraz — jak dalece střední hodnota $\bar{x} = 480$ ($s = 60$), vypočítaná ze vzorku o rozsahu $n = 500$, odpovídá skutečné střední hodnotě. Nutno totiž uvážit, že i nepatrné rozdíly v náhodném výběru by také mohly docela dobře vést ke střední hodnotě 463, 485 nebo i 502. Protože neznáme složení základního souboru (na rozdíl od balíčku karet, ze kterého jsme táhli 5 karet), nebudeme nikdy s to zjistit skutečné μ. Můžeme však vypočítat, s jakou pravděpodobností se nachází uvnitř relativně úzké míry rozptylu kolem $\bar{x}$. (Stejně je tomu u ořechů, kde P se musí odhadnout z p.)

Střední hodnoty výběrových souborů mají však normální rozdělení i tehdy, jestliže základní soubor normální rozdělení nemá. To by bylo možno matematicky dokázat v teoreticko-pravděpodobnostní exkurzi, ale zavedlo by nás to příliš daleko do oblastí, kam si netroufáme. Mimoděk jsme se totiž dotkli *„centrální limitní věty"*, která má pro matematickou statistiku základní teoretický význam, jehož nastínění se však můžeme s dobrým svědomím zříci.

Je jen nutno ještě jednou zopakovat a zdůraznit rozhodující tezi této *centrální limitní věty:* ať je rozdělení základního souboru jakékoliv, rozdělení střední hodnoty výběrového souboru (i rozptyl výběrového souboru) bude vždy normální, jestliže rozsah výběrového souboru dosáhne alespoň určité minimální velikosti. (Viz obr. na str. 153). Za nejmenší rozsah výběru — jde-li pouze o zjištění skutečné střední hodnoty základního souboru, můžeme pokládat asi $n = 30$. K tomu, abychom získali možnost vypočítat i rozptyl základního souboru, musí rozsah výběru činit přinejmenším 100.

Jestliže tyto úvahy přeneseme na náš případ, dostaneme tuto ilustraci centrální limitní věty: kdybych chtěl šetření mezi rolníky opakovat např. třicetkrát — vždy se stejným rozsahem vzorku — dalo by těchto 30 vzorků jistě střední hodnoty $\bar{x}_1, \bar{x}_2 \ldots \bar{x}_{30}$, které jsou kolem skutečné střední hodnoty μ „normálně rozděleny". Předpokládali jsme $\bar{x}_1 = 480$; dalších 29 výsledků by leželo téměř bez výjimky rovněž mezi 475 a 485.

Představa, že by se měl výběr o rozsahu $n = 500$ opakovat třicetkrát, je již teoreticky velmi podivná a v praxi úplně absurdní. Nevyžaduje příliš mnoho fantazie představit si úvodní vzorek o rozsahu $n = 500$ jako množství vzorků. Každý z 500 výsledků měření je možno pokládat za vzorek o rozsahu 1 (anebo, trochu opovážlivěji, za střední hodnotu fiktivního vzorku o rozsahu 2 nebo více).

Tím se nemění nic na principu centrální limitní věty: střední hodnoty výběrových souborů (a tím i výsledky jednotlivých výběrů) mají normální rozdělení kolem skutečné střední hodnoty základního souboru, pokud je rozsah výběrového souboru dostatečně velký.

Přesnost výpovědi o této skutečné střední hodnotě závisí na dvou faktorech: předně na rozptylu celkového souboru a za druhé na rozsahu výběru.

Matematicky se vyjadřuje tato tzv. směrodatná chyba vzorcem

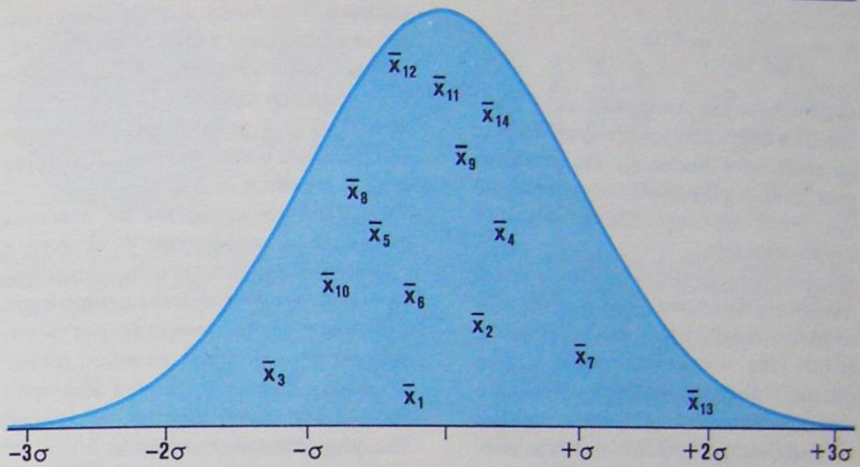
$$\boxed{\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}}$$

Jsou v něm obsaženy obě rozhodující hodnoty: směrodatná odchylka základního souboru a počet jednotek zahrnutých do výběrového souboru. Teď sice známe rozsah výběrového souboru, nikoliv však σ, směrodatnou odchylku u základního souboru, nýbrž jen s, směrodatnou odchylku u výběrového souboru. Je-li však výběrový soubor dostatečně velký, a to znamená např. n větší než 30, je možno s dobrým svědomím σ nahradit s jako „odhad" za σ. Měli jsme $n = 500$, můžeme proto klidně použít s místo σ a dostaneme směrodatnou chybu:

$$\sigma_{\bar{x}} = \frac{60}{\sqrt{500}} = 2{,}7.$$

Ve shodě se zákonitostmi normálního rozdělení jde při této hodnotě směrodatné chyby o směrodatnou odchylku. To znamená: ve dvou třetinách (68,2 %) všech případů bude střední hodnota výběrového souboru v rozmezí $\pm 2{,}7$ skutečné střední hodnoty. Protože však A je od B přesně stejně daleko jako B od A, mohu říci obráceně: ve dvou třetinách všech případů nebude naše střední hodnota výběrového souboru (v našem případě to bylo $\bar{x}_1 = 480$) více než $\pm$ 2,7 vzdálena od skutečné střední hodnoty.

Dvě třetiny případů je však neuspokojivě málo. Budeme se snažit z důvodů účelnosti dosáhnout asi 95% jistoty,



Průměry výběrových souborů, i když pocházejí z nenormálního rozdělení, jsou normálně rozděleny. Čtrnáct průměrů výběrových souborů od x_1 do x_{14} se v prostoru pod normální křivkou pohodlně rozdělí. Tento úkaz je matematicky vyjádřen v centrální limitní větě.

a proto vypočítáme dvojitou směrodatnou odchylku: dvakrát $\pm 2{,}7$ dává však $\pm 5{,}4$. Usuzujeme proto z našeho $\bar{x} = 480$, že skutečná střední hodnota je někde mezi 474,6 a 485,4.

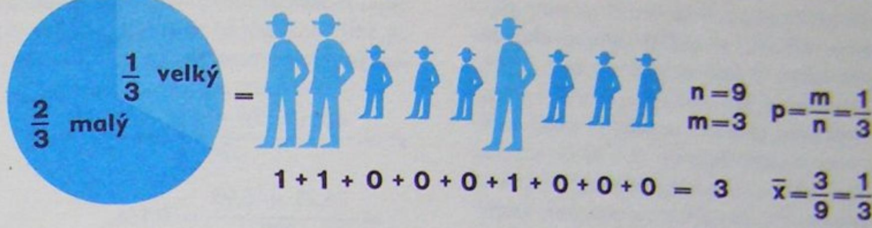
Tento postup je však nepoužitelný při malých výběrových souborech (pod 30), protože rozptyl s výběrového souboru je zde příliš nespolehlivý a protože existuje — při pěti nebo deseti výsledcích z velkého základního souboru — příliš velká pravděpodobnost, že dostaneme náhodou necharakteristické výsledky, a tím i příliš nesprávnou střední hodnotu. Takové malé výběrové soubory vyžadují speciální matematické postupy, s nimiž se ještě seznámíme (str. 196).

Nejdříve se však ještě jednou zblízka podíváme na vzorec pro výpočet směrodatné chyby — a tu nás napadne něco překvapivého. Nacházíme v něm sice n, rozsah výběrového souboru, nikoliv však N, rozsah základního souboru. To ale znamená, že vzorec pro přesnost měření závisí pouze na rozsahu výběrového souboru, nikoli však na jeho poměru k základnímu souboru! Je to možné? Zdravý selský rozum řekne rozhodně „ne", a zase jednou nemá principiálně pravdu. Matematik si prohlédne vzorec podrobněji a pak řekne „ano" — a přirozeně má zase jednou pravdu. Matematik ovšem k tomuto tvrzení připojí omezení: „*Ano*, když je základní soubor nekonečně velký."

Našich 250 000 rolníků je sice hodně veliký, nikoliv však nekonečně veliký základní soubor. Podívejme se, co se stane, vezme-li se hypotetický případ, že bychom výběrový soubor neustále rozšiřovali, až by nakonec obsáhl celé obyvatelstvo. Pak dostaneme — jak protismyslné! — stále ještě směrodatnou chybu střední hodnoty:

$$\sigma_{\bar{x}} = \frac{60}{\sqrt{250\,000}} = 0{,}12.$$

Ovšem že je jen velmi malá, 0,12 v poměru ku 480 (o kterých předpokládáme, že tvoří také hodnotu), ale přece to není nula, jak by tomu muselo vlastně být, když už není žádná náhodná chyba myslitelná.
Je ovšem pošetilé chtít vypočítat směrodatnou chybu střední hodnoty tam, kde už žádná nemůže být. V daném případě, stejně jako i v jiných, nesmí vzorec nahradit logické myšlení. Náhodná chyba 0,12, která je bezpodmínečně příliš velká, by vznikla, kdybychom se sice nedotazovali všech 250 000, ale jen 249 990 osob. Musíme tedy použít *opravného koeficientu*, který redukuje 0,12 téměř na nulu. Později se k němu ještě vrátíme (viz str. 160). Již teď je ale zřejmé, že při velkém rozsahu základního souboru *závisí* přesnost výběru skutečně jen na *rozsahu výběrového souboru n a nikoliv na N.*
To však neznamená nic více a nic méně, než že rozsah výběrového souboru může být prakticky stejný, ať už se základní soubor skládá ze 60 miliónů obyvatel NSR, 10 miliónů obyvatel Bavorska, z půldruhého miliónů Mnichovanů nebo z 500 000 Mnichovanek. Přitom samozřejmě platí, jako při každém náhodném výběru, že každá jednotka základního souboru musí mít stejnou naději být zahrnuta do výběru. V jednom případě budeme tedy muset hledat našich řekněme 3000 osob vzorku na celém území NSR, v jiném případě jen na území Mnichova. To proto, že výběrový soubor z Mnichovanek nejen neřekne nic o mužích z Hamburku, ale ani nic o „Bavorácích", výběrový soubor z Bavorska pak nic o „občanech NSR".
Jak je to v případě těch výběrových souborů, kde nejde o střední hodnotu, nýbrž o rozdělení znaků („kladení otázek stejného významu")? Zkoumejme situaci našich ořechů. Tam jsme měli $N = 1000$ a v prvním malém výběrovém souboru o rozsahu $n = 20$ hodnoty $m = 5$, a tedy $p = {}^5/_{20} = {}^1/_4$.
Zde je nutno upozornit na klamnou dvojznačnost uvedeného p. Známe p (nebo také P) jako zkratku „*pravděpodobnosti*" (lat. *probabilitas*) a víme např. z příkladu házení mincí, že pravděpodobnost pro „dvakrát za sebou hlava" se udává jako $p = {}^1/_4$. Stejně je $p = {}^1/_4$ za: „ze hry karet táhnout jeden pik". Tím přenášíme nám známou úměrnost piků v celém balíčku karet zcela logicky na pravděpodobnost vytáhnutí jednoho piku: je-li v balíčku ${}^1/_4$ pikových karet, je pravděpodobnost vytažení jedné z nich také právě ${}^1/_4$. Poměr P určuje tedy pravděpodobnost p (nebo P).
Mluvíme-li však o výběrových souborech, vzniká *jistý posun významu*. Vždyť známe jen proporci uvnitř výběrového souboru (označujeme ji p) a používáme ji jako odhad pro domnělou proporci v souboru, kterou označujeme P. Když jsme tedy mluvili o tom, že v našem výběrovém souboru bylo $p = {}^1/_4$, znamená to pouze, že poměr souboru k nositeli znaku ($n : m$) byl ve výběrovém souboru 1 : 4. V daném případě však nemáme co činit s výpovědí o pravděpodobnosti. V dalším výkladu brzy uvidíme, že skutečná hodnota (pravděpodobnost správnosti domněnky) se zjišťuje docela jinak a nemá s velikostí podílu znaků téměř nic společného. Chceme-li se vyhnout nedorozuměním, může jako východisko posloužit to, že se napíše místo proporce výběrového souboru p raději m/n, místo proporce P v celkovém souboru raději M/N.
Ve druhém, větším výběrovém souboru



Podílovou četnost lze vyjádřit jako aritmetický průměr. V binomickém rozdělení velký — malý je možno třetinový podíl velkých na základním souboru vyjádřit také tím, že se jednomu znaku přidělí charakteristická veličina 0, druhému 1.

jsme měli $n = 200$, $m = 42$, $p = 0{,}21$. Protože jak N, tak n jsou dostatečně velké, aby se mohlo místo binomického rozdělení použít rozdělení normálního (a protože p je pro Poissonovo rozdělení příliš veliké), budeme v našich dalších úvahách a propočtech používat normálního rozdělení.
Normální rozdělení je určeno svým aritmetickým průměrem a rozptylem σ^2, případně *směrodatnou odchylkou* σ. Průměr však odpovídá v případě četnosti znaků skutečnému podílu M na N, je tedy $M/N = P$. Rozptyl ale dává hodnotu: $P\,(1 - P)$.
Tomu, jak je to zde napsáno, nutno věřit. Bylo by však možné dokázat to matematicky přesně. Zvolíme střední cestu a pokusíme se o důkaz poněkud jednodušší. Kvalitativní rozdělení znaků (dobrý — špatný, velký — malý, světlý — tmavý atd.) totiž můžeme změnit na kvantitativní tím, že např. stanovíme: všechny špatné ořechy mají charakteristickou hodnotu 1, všechny dobré 0. Nestanovíme tedy už „podíl", poměr, nýbrž sčítáme: $1 + 0 + 0 + {}$ $+ 0 + 1 + \ldots$ V případě našeho většího výběrového souboru (s rozsahem $n = 200$ a $m = 42$) dostáváme celkem 42, a tím aritmetický průměr $\bar{x} = 0{,}21$, což je přesně poměr m/n, proporce výběrového souboru. A tento průměr $\bar{x}$ se má k průměru stejně jako poměr výběrového souboru m/n (nebo p) k proporci základního souboru M/N (nebo P).
Chybu relativní četnosti můžeme proto pokládat za zvláštní případ směrodatné chyby (nebo také obráceně) a použít stejného vzorce jako dříve — jen místo $\sigma_{\bar{x}}$ píšeme nyní σ_p:

$$\boxed{\sigma_r = \frac{\sigma}{\sqrt{n}}}$$

Tím ovšem opět vzniká otázka, *odkud máme vzít* σ? Přinejmenším je můžeme vypočítat jako rozptyl výběrového souboru s, stejným způsobem jako kdysi ve Zbohatlíkově, když jsme se seznamovali se směrodatnou odchylkou. Tehdy musela být každá odchylka od průměru $\bar{x}$ nejdříve umocněna na druhou, abychom dostali rozptyl, jehož odmocnina byla směrodatnou odchylkou.
V našem nynějším případě je to zvlášť snadné potud, že nemusíme umocňovat bezpočet (resp. ani tucet) různých odchylek. Vždyť máme jen dvě jednoduchá čísla, charakteristické hodnoty 0 a 1 — a samozřejmě průměr $\bar{x} = 0{,}21$. Jsou tedy jen dvě možné odchylky od tohoto průměru. Buď máme před sebou cha-

rakteristickou hodnotu 1 a pak platí $(x - \bar{x})^2 = (1 - 0{,}21)^2$, anebo charakteristickou hodnotu 0 a pak platí: $(x - \bar{x})^2 = (0 - 0{,}21)^2$.
V prvním případě je tedy $(1 - p)$ v závorce a umocňuje se.
V druhém případě, protože záporné znaménka se při umocňování stejně ztrácejí, je výraz rovný p^2.
Charakteristická hodnota 1 se vyskytuje ve výběrovém souboru *m*krát ($m = 42$), charakteristická 0 podle toho $(n - m)$ krát $(n - m) = 158$. Souhrn čtvercových odchylek, jak jej potřebujeme pro výpočet rozptylu (srovnej oddíl 2.6), činí tedy: 42krát $(0{,}79)^2$ + 158krát $(0{,}21)^2$. Abychom tento rozptyl obdrželi, dělíme uvedený výběrový soubor celkovým počtem měření a dostáváme:

$$s^2 = \frac{42 \cdot (0{,}79)^2 + 158 \cdot (0{,}21)^2}{200} =$$

$$= \frac{26{,}25 + 6{,}95}{200} = 0{,}166,$$

z toho kořen $s = \sqrt{0{,}166} = 0{,}41$.

Stejné hodnoty však můžeme získat jednodušším způsobem, když dosadíme: $s^2 = p(1 - p)$ a $s = \sqrt{p(1 - p)}$. Spočítáme, že $s^2 = 0{,}166$.

k \ n	20	50	100	150	200	300	500
0	0—145	0—71	0—36	0—24	0—18	0—12	0—7
1	5—223	1—107	1—54	0—37	0—28	0—18	0—11
2	25—293	7—137	4—70	2—47	2—36	1—24	1—14
3	53—356	17—166	8—85	5—57	4—43	3—29	2—17
5	116—474	40—218	20—113	13—76	10—57	7—38	4—23
10	289—711	100—338	49—176	31—119	24—90	16—60	10—36
20	855—1000	264—548	126—292	83—199	62—151	41—101	25—61
50		929—1000	398—602	259—415	191—316	126—215	75—130
100			964—1000	585—741	429—571	280—395	165—238
150				976—1000	684—809	443—557	259—341
200					982—1000	605—720	355—446
250						785—874	456—544

Jak jsou intervaly spolehlivosti vzdáleny ještě u výběrů středního rozsahu, lze poznat z této tabulky, ve které je nahoře uveden rozsah výběru *n*, vlevo četnost znaku *k*, nalezená v tomto výběru. Jestliže se např. ve výběru o rozsahu $n = 150$ hledaný znak najde pětkrát (tedy $k = 5$), činí interval spolehlivosti (uvádí se v promilích) 13 až 76. Protože tabulka vychází z 95% pravděpodobnosti jistoty, znamená to, že na základě tohoto výsledku mohu předpokládat, že v tomto — velmi velikém až nekonečném — základním souboru činí četnost znaku nejméně 13 ‰. Takto mohu odpovědět s 95% jistotou. Zůstává pak ještě pravděpodobnost omylu ve výši 5 %, tzn. že četnost je menší než 13 ‰ nebo větší než 76 ‰.

Směrodatnou odchylku získanou z výběrového souboru dosadíme do již známého vzorce pro směrodatnou chybu relativní četnosti a dostáváme:

$$\sigma_p = \frac{\sqrt{p \cdot q}}{\sqrt{n}} = \sqrt{\frac{p \cdot q}{n}}.$$

S hodnotami z našeho příkladu vypočítáme náhodnou chybu:

$$\sigma_p = \frac{0{,}41}{14{,}2} = 0{,}029.$$

Chyba relativní četnosti činí v tomto výběrovém souboru $\pm$ 2,9 %.
To však znamená za prvé: ze všech teoreticky možných výběrových souborů tohoto rozsahu ($n = 200$) dávají dvě třetiny podílové hodnoty mezi 18,1 (= 21 — 2,9) a 23,9 (= 21 + 2,9). Dvě třetiny proto, že jde o prostou směrodatnou odchylku. To pak za druhé znamená, že dostaneme-li takový výsledek výběrového souboru, můžeme s pravděpodobností (přesně) 68,3 % soudit, že skutečný podíl celkového množství je v těchto mezích.
Přejeme-li si, jak je většinou obvyklé, 95 % jistoty, použije se dvojité směrodatné odchylky: jestliže výběrový soubor o rozsahu $n = 200$ dává četnost podílu 21 %, je skutečná hodnota podílu s 95% jistotou mezi 15,2 a 26,8 %.
Znovu se vynořuje otázka, zda skutečně *na velikosti základního souboru vůbec nezáleží*. Protože jde v podstatě o stejné závěry, které jsme již rozvíjeli, dojdeme i v tomto případě ke stejnému výsledku. Pro nekonečně velké základní soubory vzorec souhlasí, pro velmi veliké základní soubory souhlasí do té míry, že může být bez připomínek použit.
Jiná situace vznikne, jestliže základní soubor není příliš velký, avšak výběr ve srovnání s ním velký je. Přesně tento případ jsme měli v našem druhém příkladu výběrového souboru ořechů: $N = 1000$ není moc veliký základní soubor, $n = 200$ je v poměru k tomu veliký výběrový soubor. V těchto případech se ke vzorci připojuje opravný koeficient, který přihlíží k tomu, že základní soubor je výběrovým souborem do značné míry „spotřebován".
Tento doplňkový opravný koeficient — konečnostní násobitel — je stejný jak pro náhodnou chybu četnosti, tak pro náhodnou chybu střední hodnoty:

$$\sqrt{\frac{N - n}{N - 1}}\text{, nebo jednoduše a skoro stejně:}$$

$$\sqrt{1 - \frac{n}{N}}.$$

Pro příklad rolníků ($N = 250\,000$, $n = 500$) dostaneme tento opravný koeficient:

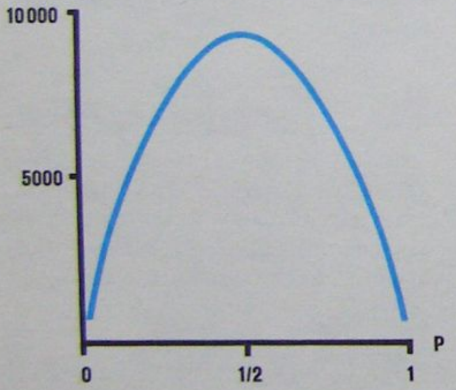
$$\sqrt{1 - \frac{500}{250\,000}} = \sqrt{\frac{499}{500}} = 0{,}999.$$

Ten se téměř rovná 1 a neovlivňuje to tedy náš původní výsledek. Jinak je tomu v příkladu s ořechy, kde $N = 1000$ a $n = 200$. Zde již dostaneme:

$$\sqrt{1 - \frac{200}{1000}} = \sqrt{0{,}8} = 0{,}894.$$

Náš výsledek byl $\sigma_p = 0{,}029$ a správně bychom jej měli ještě násobit tímto faktorem. Pak dostáváme: $\sigma_p = (0{,}029) \times (0{,}894) = 0{,}026$.
Náš odhad je proto dokonce přesnější než původní předpoklad, což je pochopitelné vzhledem k tomu, že jsme

vzorkem zachytili pětinu základního souboru. Nyní můžeme naši původní výpověď upřesnit tak, že řekneme: s jistotou 95 % je podíl prázdných ořechů mezi 15,8 a 26,2 %.
Nyní uvedeme právě rozvinuté myšlenkové pochody o spolehlivosti výběrů a o náhodných chybách průměrů a četnosti v jiné souvislosti. Dosud jsme vycházeli z toho, že máme určitý výběr o rozsahu $n = 200$ nebo $n = 500$ anebo $n = 20$ a zkoumali jsme, jaké závěry a s jakou pravděpodobností je z nich možno udělat pro základní soubor. Nyní úvahu obrátíme a ptáme se: Jak velký výběr musím zvolit za předpokladu, že chci dosáhnout 95% spolehlivosti? Tato otázka stojí vždy znovu před statistiky a těmi, kdo zkoumají veřejné mínění. Abychom mohli nalézt odpověď, podívejme se ještě jednou na



Nejmenší rozsah výběrového souboru závisí na podílu znaku. Toto vyobrazení — podle náčrtku u Pfanzagla — ukazuje žádaný rozsah výběrového souboru při požadované četnosti ± 1 % a při pravděpodobnosti 95 %. Činí-li četnost znaků např. $^1/_2$ ($p = 0{,}5$), musí být rozsah výběrového souboru bezmála 10 000. Při mnohem nižší nebo mnohem vyšší četnosti znaků může klesnout pod 1000.

oba již uvedené vzorce. Především je to směrodatná chyba průměru:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}\,.$$

Pak je to směrodatná chyba četnosti:

$$\sigma_p = \sqrt{\frac{pq}{n}}\,.$$

Tuto směrodatnou chybu mohu stanovit libovolně. Vyplývá ze spolehlivosti, které chci dosáhnout: chci dosáhnout 95 %, 99 % nebo 999 ‰? Čím má být odhad přesnější, tím větší musí být výběrový soubor. Na první pohled je však jasné, že v důsledku úzkého spojení směrodatné chyby s *odmocninou* rozsahu výběrového souboru se *dvojnásobná* přesnost vykupuje *čtyřnásobným* rozsahem výběrového souboru. Nemá proto smysl žádat větší přesnost, než je nezbytně nutné.
Kromě toho potřebujeme p a q ($= 1 - p$), resp. v případě „kladení různorodých otázek" směrodatnou odchylku σ. Jestliže pro ně mám směrné hodnoty z dřívějších šetření, z literatury nebo z odhadů expertů, mohu je dosadit. Nemám-li vůbec nic, musím z opatrnosti předpokládat to nejméně příznivé — při rozptylu σ nezná můj pesimismus téměř hranic. Při rozdělení četnosti je rovnoměrná pravděpodobnost $p = q$, jak hned uvidíme, tím nejméně příznivým případem.
Vezměme znovu náš pytel ořechů ($N = 1000$) a ptejme se: jak velký musí být vzorek, mám-li na jeho základě s 95% jistotou vědět, že podíl prázdných ořechů se neodchyluje o více než ± 4 % od četnosti zvoleného výběrového souboru?
Hledáme tedy n (rozsah výběrového souboru). Nejdříve upravíme pro naše

účely rovnici vztahující se na náhodnou chybu četnosti a dostaneme:

$$n = \frac{p \cdot q}{(\sigma_p)^2}\,.$$

Dvojitá chyba četnosti má činit nejvýše ± 4 %, jednoduchá tedy ± 2 %. Tím je dán jmenovatel zlomku. O domnělém rozdělení nevím zatím vůbec nic, možná že je zkaženo půl pytle. Tedy $p = q = 0{,}5$. Protože však chyba četnosti je vyjádřena v procentech, musíme je upravit na 0,02 (nebo: $p = q = 50$ %, $\sigma_p = 2$ %. Dostaneme tedy:

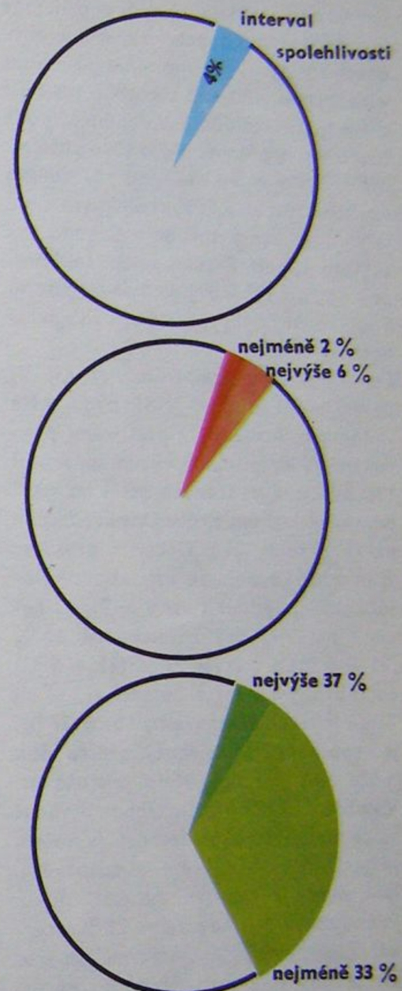
$$n = \frac{0{,}5 \cdot 0{,}5}{(0{,}02)^2} = \frac{0{,}25}{0{,}0004} = 625.$$

Museli bychom testovat více než půl pytle, abychom dosáhli takové přesnosti! Je účelné spokojit se s menší přesností: ± 6 %, ale nadále při 95% pravděpodobnosti.

$$n = \frac{0{,}25}{(0{,}03)^2} = \frac{0{,}25}{0{,}0009} = 278.$$

Stále ještě příliš mnoho. Několik ořechů by jistě odpadlo, kdybychom použili opravného koeficientu pro malé základní soubory — pohodlnější by však bylo, kdyby v čitateli byla menší hodnota. Jak by tomu ale bylo v případě této úvahy: pytel ořechů nám darovali dobří přátelé, kteří tam jistě podle svých možností dali dobré ořechy, neboť koneckonců znají své ořešáky. Je-li 20 % špatných, je to už jistě mnoho. Teď tedy $p = 0{,}2$, $q = 0{,}8$, přesně na 6 % a 95 % spolehlivosti:

$$n = \frac{(0{,}2) \cdot (0{,}8)}{(0{,}03)^2} = \frac{0{,}16}{0{,}0009} = 178.$$



Interval spolehlivosti ve výši 4 % má různou váhu u malých a velkých podílů. Jestliže podíl vzorku činí např. 4 %, sahá interval spolehlivosti od 2 do 6 %, tedy až do trojnásobku. Činí-li podíl vzorku naproti tomu 35 %, je rozpětí „mezi 33 a 37 %" dostatečně informativní a přesné; 37 % je totiž jen asi o desetinu větší než 33 %.

I to se může stále ještě zdát nepříjemně mnoho k tomu, abychom zjistili kvalitu ořechů v pytli, ale nesmíme pominout všeobecnou platnost těchto a také již dříve zjištěných údajů. V daném případě nejde jen o ořechy, ale o toto zjištění: má-li být z velmi velikého základního souboru zjištěn s 95% spolehlivostí určitý znak, který se podle odhadu vyskytuje v pětině základního souboru, pak postačí výběrový soubor o rozsahu $n = 178$, aby byl vypočítán podíl s přesností $\pm$ 6 %.

Jestliže se tedy domníváme, že asi $^1/_4$ až $^1/_6$ všech občanů NSR bydlí ještě v domech, které byly postaveny před rokem 1900, postačí, abych se zeptal 178 náhodně vybraných osob na stáří domu, ve kterém bydlí. Odpověď bude na 95 % spolehlivá. Docílená přesnost však zdaleka nebude uspokojivá. Dosáhneme-li výsledku asi $p = 21$ %, nevíme nic víc, než že nejméně 15 % (21 — 6 %) a nejvýše 27 % (21 + 6 %) osob bydlí ve starých domech.

Tento interval spolehlivosti, 15 až 27 %, je nepochybně příliš veliký, přesto však může pro některé účely postačovat. V každém případě je ale třeba všimnout si, že náhodná chyba četnosti (v našem případě $\sigma_p = 3$ %) je tím závažnější, čím menší je relativní četnost znaků. Právě jsme viděli, že při $p = 21$ % interval spolehlivosti dvojnásobné směrodatné odchylky (pravděpodobnost jistoty 95 ze 100) činí vždy v obou směrech dvojnásobek a že proto docházíme k širokému intervalu 15 až 27 %. Jestliže je určitý znak velmi četný, nemusí se 12% rozpětí jevit jako přehnané: $p = 52$ % dává pak interval spolehlivosti právě od 46 do 58 %. Je-li však podíl malý, dostaneme dokonce zcela nesmyslné výsledky. Předpokládejme například, že bychom dostali $p = 7$ %. Tento výsledek by se musel vykládat takto: „s největší pravděpodobností mezi 1 % a 13 %" — nepřihlížíme-li navíc k tomu, že intervaly spolehlivosti směrem k nule nejsou již symetrické.

Kdyby tedy např. volební prognóza měla náhodnou chybu četnosti jen ve výši $\pm$ 2 %, znamenalo by to v případě velkých politických stran, které mají zhruba stejný počet hlasů, velkou nejistotu, a naopak u malých stran může rozpětí 4 % znamenat: „nejméně 3 a nejvýše 7 %" (ať žije malý rozdíl!).

Ukázali jsme již, že rozsah výběrového souboru musí být největší tehdy, když $p = q$, tzn. při domněle rovnoměrném rozložení znaku a neznaku (po 50 %). To se však vztahovalo jen na *absolutní* rozpětí intervalu spolehlivosti. Musím prohlédnout 278 ořechů, abych mohl s $\pm$ 6 % vymezit okruh 50 %, ale postačí 178 pro $\pm$ 6 % v oblasti kolem 20 %. V oblasti kolem 10 % by stačilo ještě méně ořechů. *Relativní* přesnost při 50 % je přirozeně mnohem větší (interval spolehlivosti 44—56 %) než při 20 % nebo dokonce 10 % (interval spolehlivosti je 4—16 %).

Podívejte se nyní, jak velký musí být rozsah výběrového souboru, aby se četnost odhadovaná jen na 4 % zachytila s přesností $\pm$ 1 %, což stále ještě znamená relativně široký interval spolehlivosti — „s největší pravděpodobností mezi 3 až 5 %".

Protože se snažíme docílit pravděpodobnosti ve výši 95 %, odpovídá přípustnému rozsahu kolísání ve výši $\pm$ 1 % σ_p v poloviční výši. Obdržíme tedy:

$$n = \frac{(0{,}03)\,.\,(0{,}97)}{(0{,}005)^2} = \frac{0{,}029}{0{,}000025} = 1160.$$

Ano, rozsah výběrového souboru $n = 1160$ by snad mohl být přijatelný.

Může nám však zaručit stejnou přesnost i při silně rozloženém znaku? Předpokládejme, že se má hledat znak s odhadovanou četností $p = 0{,}4$ (tj. asi 40 % souboru). Vypočítáme střední chybu četnosti pro tento případ a s nepříjemným překvapením zjistíme, že

$$\sigma_p = \sqrt{\frac{(0{,}4)\,.\,(0{,}6)}{1160}} = \sqrt{0{,}00206} =$$

$$= 0{,}0143.$$

Již jednoduchá směrodatná odchylka dává v tomto případě $\pm$ 1,43, požadovaná 95% pravděpodobnost dokonce dvojnásobek, tedy $\pm$ 2,86.

Jestliže však zůstaneme při našem původním požadavku $\pm$ 1 % z 95% pravděpodobnosti, musíme rozsah výběrového souboru podstatně zvětšit, neboť:

$$n = \frac{(0{,}4)\,.\,(0{,}6)}{(0{,}005)^2} = 9600.$$

Museli bychom se tedy dotázat alespoň 9600 osob, chceme-li pracovat s uvedenou přesností a spolehlivostí. To je však takový rozsah výběrového souboru, který se při dotazování téměř nikdy nepoužívá. Zpravidla je rozsah výběrového souboru mezi 2000 a 3000 dotazovaných, velmi zřídka větší, často ještě menší.

Jakou přesnost bychom mohli na naši poslední otázku (odhadovaná četnost znaků $p = 0{,}4$) očekávat od výběrového souboru o rozsahu $n = 2500$ osob, jestliže se opět snažíme docílit 95% pravděpodobnosti?

Dosadíme-li $n = 2500$ a $p = 0{,}4$ do známého vzorce, dostaneme:

$$\sigma_p = \sqrt{\frac{(0{,}4)\,.\,(0{,}6)}{2500}} = \sqrt{0{,}00096} =$$

$$= 0{,}0098,$$

nebo pohodlněji počítáno s procenty:

$$\sigma_p = \sqrt{\frac{40\,.\,60}{2500}} = \sqrt{0{,}96} = 0{,}98.$$

Dvojnásobek 0,98 dává 1,96 nebo zaokrouhleně 2 % a na tomto základě můžeme říci, že výběrový soubor o rozsahu $n = 2500$ může určit znak obsažený v základním souboru asi čtyřmi desetinami s přesností na ± 2 %, a to s 95% pravděpodobností. Nic více nelze z výběrového souboru vyčíst. Přitom vycházíme vždy z přesvědčení, že před sebou máme ideálně náhodný výběr, tzn. výběrový soubor nezkreslený systematickou chybou a v důsledku toho ještě daleko méně přesný! Tak je také třeba číst výsledek šetření: „41,3 % dotázaných" ve skutečnosti znamená: „s největší pravděpodobností ne méně než 39,3 % a ne více než 43,3 %".

5.5 Spolehlivost výběrových souborů — intervaly spolehlivosti

Pravděpodobnost jistoty, spolehlivost, vypovídací schopnost, přesnost, interval spolehlivosti — je to vše totožné? Pokusíme se ještě zřetelněji ukázat souvislosti a rozdíly a především se nad těmito pojmy poněkud zamyslíme.

Co je například *spolehlivost*? Zpravidla jsme tím mysleli onu spolehlivost, která říká, že výpověď je z 90 nebo 95 % správná. To je pravděpodobnost jistoty, která vymezuje také interval spolehlivosti a kterou se budeme ještě zabývat podrobněji.

„Spolehlivost" se však může vztahovat také na všeobecnou kvalitu statistické výpovědi. V tomto smyslu např. systematická chyba narušuje spolehlivost.

Další otázkou je, zda vytváření výběrových souborů je vůbec spolehlivou vědeckou metodou. Kromě toho experiment musí být zásadně opakovatelný. Specifikem statistických experimentů však je, že nelze téměř nikdy docílit přesně stejných výsledků, nýbrž jen velmi podobných. Zkoumání výběrových souborů má být opakovatelné, jeho výsledky však mohou být pouze navzájem porovnatelné, nikoli totožné. Další požadavek, který se klade na všechny vědecké pokusy, je *požadavek významnosti, vypovídací schopnosti a důležitosti*: musí se skutečně dokázat to, co má být dokázáno, nikoliv jen něco podobného.

Vraťme se však zpět k 95% *pravděpodobnosti* jistoty, ke stupni jistoty. Co si pod tím máme představit? Můžeme opět pomyslit na rozdělení četnosti pod normální křivkou a předpokládat soubor všech možných výběrových souborů v rozsahu n jako normálně rozdělený. Pak 95% pravděpodobnost znamená, že výběrový soubor, který jsme právě náhodně vytvořili, je právě v toleranci 95 %, které dávají střední hodnotu $\bar{x}$, ležící maximálně 2σ od skutečného průměru μ.

Zda postačí 95 % jistoty nebo nikoliv, nelze říci všeobecně. O tom je třeba rozhodnout v každém jednotlivém případě samostatně. Jestliže se nemá vydat moc peněz a postačí-li přibližný přehled (jako např. u mnohých otázek průzkumu trhu), je postačující 90% nebo ještě nižší jistota. Jde-li o zvlášť významné, velmi závažné rozhodnutí, bude snaha dosáhnout 99% nebo ještě vyšší pravděpodobnosti. Za postačující se však zpravidla pokládá 95% pravděpodobnost.

Jestliže chceme dosáhnout vyšší spolehlivosti, je nutno zkoumat větší výběrové soubory. Není-li to z nějakých důvodů možné, pak také nelze dosáhnout žádaného stupně jistoty — a dost! Nepomohou žádné početní triky či okliky a ani jiné komplikované kejkle neposkytnou východisko. Rozsah výběrového souboru a spolehlivost výsledku výběrového souboru jsou také v nejužší logické a matematické souvislosti s úkazem, který jsme pozorovali při prvním výkladu binomického rozdělení a Pascalova trojúhelníku: v protikladu k naivní představě o „zákonu velkých čísel" neexistuje sice přesné ustálení výkyvů na očekávané hodnotě, přesto však při rostoucím rozsahu výběrového souboru se budou proporce výběrového souboru relativně (ne absolutně!) méně odchylovat od skutečné proporce souboru základního.

V ruletě to například znamená: je zcela dobře myslitelné, že ve „výběrovém souboru" dvaceti po sobě jdoucích her je 16 červených a jen čtyři černé, to znamená odchylka o 6 od očekávané hodnoty 10. Naproti tomu je zcela vyloučeno, aby z 2000 her bylo 1600 červených a jen 400 černých. I poměr 1100 ku 900 by byl velmi nepravděpodobný. S rostoucím rozsahem výběrového souboru bude tedy stále přesněji zaměřován skutečný průměr základního souboru (který v případě rulety známe). Při rozsahu výběrového souboru $n = 20$ se mohu velice zklamat (např. jako u neuvěřitelného výsledku 16 červených, 4 černé), při $n = 200$ se však mohu spolehnout, že neskončím příliš daleko od pravdy. To je „zákon velkých čísel" v teorii a praxi výběrových souborů.

Pravděpodobnost výběrového souboru mohu také definovat ze záporného hlediska a obdržím „*riziko chyb*"; 95% pravděpodobnost znamená 5% riziko

chyb. Máme-li pro „pozitivní" a „negativní" dva různé výrazy, stane se situace ještě zmatenější, když se začne mluvit o „hladině významnosti". Není to zvlášť šťastně zvolený výraz, ale přesto je velmi rozšířen. „*Významnost*" je prakticky totéž co pravděpodobnost: výsledek je „statisticky významný na úrovni 5 %", protože by jen čirou náhodou nenastal v 95 % případů. Celá rozsáhlá oblast testování hypotéz (viz kap. 6), která nás bude ještě velmi zaměstnávat, se zakládá na takovýchto úvahách a výpočtech: je nebo není dosažený výsledek slučitelný s tou nebo onou hypotézou?

„Významnost" není tedy nic jiného než dohoda mezi těmi, kdo statistické metody používají. A zcela stejně jako u pravděpodobnosti lze různě posuzovat úroveň významnosti v závislosti na kladení otázek. *Zpravidla se všeobecně pokládá úroveň 5 %* (nazývaná také často úroveň 95 %) *za „významné", úroveň 1 % (99 %) za „vysoce významné"*.

Pravděpodobností jistoty je také zároveň určen *interval spolehlivosti*. Tento interval se měří jako směrodatná odchylka normovaného normálního rozdělení nebo také i v absolutních hodnotách. Měli jsme již příklady pro oba způsoby vyjádření: Úspory mají rozptyl $s = 60$ kolem průměru $\bar{x} = 480$, interval spolehlivosti (95 %) sahá od 360 do 600 DM, totiž od $(480 - 2\sigma)$ až do $(480 + 2\sigma)$. Mám 95 % jistoty, že náhodně vybraný rolník má více než 360 a méně než 600 DM úspor.

Poznámka: Pozor na záměnu *pravděpodobnost jevu a pravděpodobnost statistického rozboru*. Pravděpodobnost četnosti nějakého znaku zjištěná ve vzorku (proporce výběrového souboru p) je nezávislá na pravděpodobnosti jistoty tohoto vzorku. Mohu vytvořit tři vzorky o rozsahu $n_1 = 20$, $n_2 = 100$ a $n_3 = 500$, které (náhodou) všechny mají četnost výběrového souboru $p = 0{,}7$ (tedy 70 %) znaku. Avšak kvalita, vypovídací schopnost, pravděpodobnost jistoty malého výběrového souboru budou podstatně nižší než u většího souboru.

Nyní představíme zcela krátce ještě dva zvláštní případy praxe výběrových souborů — výběrové soubory poskytující přesnější výsledky než vyčerpávající šetření, a to výběrové soubory, na kterých je založen „předběžný výpočet".

5.6 Systematické chyby — „předběžný výpočet"

V článku německého odborného časopisu „*Wirtschaft und Statistik*" bylo se zřetelem na sčítání lidu z května 1970 řečeno mimo jiné: „Teprve při sčítání lidu v roce 1970 se důsledně sledovalo omezení vyčerpávajícího sčítání na ony znaky, pro které jsou potřebné pokud možno přesné výsledky regionálně vymezených malých jednotek; všechny ostatní znaky je však třeba zjistit jen za reprezentativní rozsah 10 % domácností, protože tyto výsledky... reprezentují skutečnost dostatečně přesně a... systematická chyba je jistě menší než při jejich začlenění do vyčerpávajícího zjišťování."

Poslední část věty je nutno pozorně přečíst ještě jednou, abychom zcela pochopili, co se vlastně říká: Celá řada otázek byla úmyslně dána jen 10 % domácností NSR, a nikoliv všem domácnostem, protože „systematická chyba je jistě menší než při jejich začlenění (takových znaků) do vyčerpávajícího zjišťování". Výběr je přesnější než vyčerpávající zjišťování, část dává lepší informace o celku než celek sám?! Zbláznil se někdo? Neodporuje to zcela tomu, co jsme se dověděli o pravděpodobnosti výběrů?

Domnělý rozpor se rozplyne, jakmile upozorníme na výraz *„systematická“*. Náhodná chyba výběru je vždy větší než náhodná chyba základního souboru, neboť u celkového souboru není žádná náhodná chyba; náhoda je nahrazena jistotou, odhad úplnou a celou skutečností.

U systematické chyby je tomu jinak. Vezměme jednu z otázek, které byly kladeny jen při výběru. Otázka 27: „Jste — mistr v průmyslu, dílenský mistr nebo podobně? (Vedoucí čety, předák nebo podobně? Vedoucí, obchodní vedoucí, člen představenstva nebo podobně? Činný ve vedení nebo v dozoru?)“ Nad těmito otázkami se ještě uvádí: „Pro osoby ve vedoucím nebo dozorčím postavení (bez samostatně činných).“

Předpokládejme nyní, že tento znak nějakého vedoucího postavení je vhodný pro 14 % základního souboru i výběru. Mnozí z dotazovaných si však myslí, že musí k této otázce něco uvést, anebo snad chtějí imponovat sami sobě či statistickému úřadu a vyplní „předák nebo podobně“. Jestliže z těch 86 % ve skutečnosti „ne vedoucích“ se jen malý zlomek tomuto impulsu poddá, posune se procento údajně „vedoucích“ ze 14 na, řekněme, 17 %.

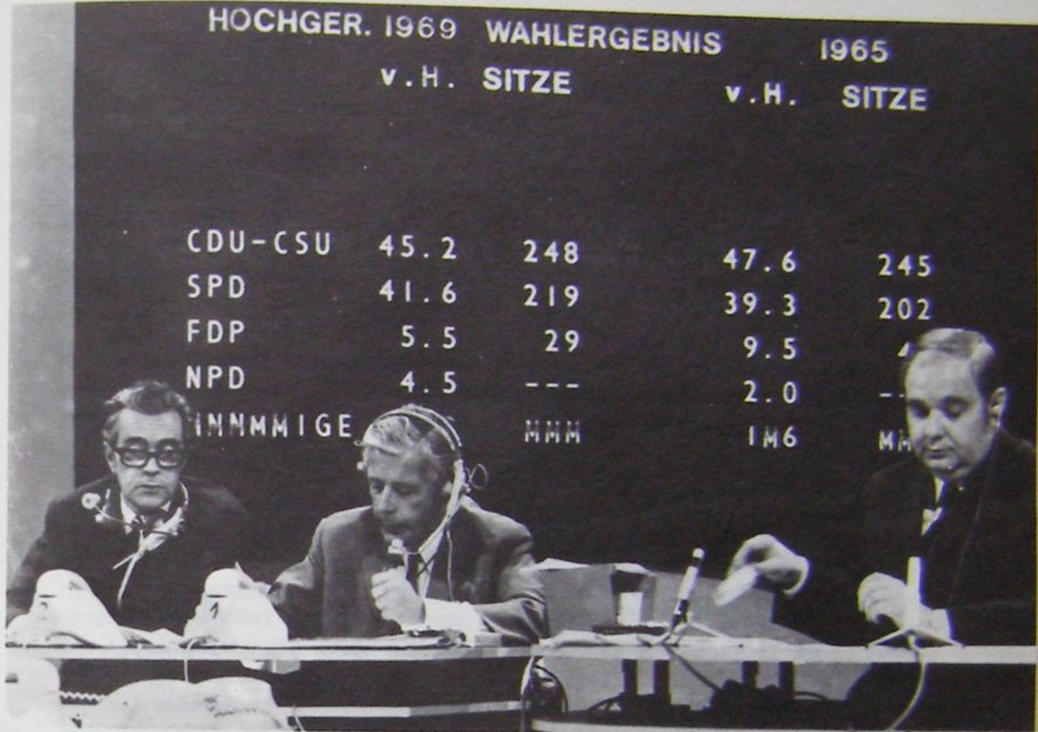
A teď uvažujme. Jestliže se výběr zakládá na přesnosti ± 3 % (při jedné desetině všech domácností je ovšem ještě podstatně přesnější), je výsledek tento: „Podíl osob ve vedoucím postavení je s 95% jistotou mezi 14 a 20 %.“ Přes značnou systematickou chybu vzniklou nesprávným zodpovězením otázky se přece jen objeví správný podíl 14 %, tvořící krajní mez intervalu spolehlivosti.

Dostaneme-li 17 % na základě vyčerpávajícího zjišťování, pak v tomto výsledku již není náhodná chyba a výpověď musí znít: „Podíl osob ve vedoucím postavení činí 17 %.“ Tato výpověď je méně správná a více matoucí než výsledek výběru, přesto nebo právě proto, že byl zahrnut celý základní soubor.

To, co vidíme v tomto zvláštním úkazu, není nic jiného než plíživý průnik nepoznané systematické chyby, která je ve výběrovém souboru skryta pod širokým pláštěm velkorysé náhodné chyby, potom však neustále sílí, až nakonec zcela pohltí náhodu, avšak systematická chyba sama zůstane nezlomena, nezakryta a také nepoznána.

Nakonec několik slov o módním výrazu *„předběžný výpočet“ (Hochrechnung)*, který přešel z oboru zpracování dat výpočetní technikou do slovníku všedního dne. Zásadou „předběžného výpočtu“ je *úsudek z části na celek:* Mám zlomkovité výsledky a vypočítám z nich domnělý celkový výsledek. To je v podstatě stejný postup jako při vyhodnocování výběrových souborů a někdy se ve výrazu „předběžný výpočet“ pro tento postup opravdu používá. Např. Haseloff uvádí ve své *„Malé učebnici statistiky“* tuto definici: „»Předběžný výpočet« označuje transformaci výběrového souboru do velikostních dimenzí základního souboru.“

Nejpopulárnější případ „předběžného výpočtu“ poskytuje předpověď celkového volebního výsledku po obdržení prvních dílčích výsledků, což již několik let provádějí rozhlasové a televizní stanice krátce po uzavření volebních místností, a to pomocí počítačů s odpovídajícím programem.



Předběžný výpočet (Hochrechnung) volebního výsledku z dílčích výsledků je velmi podobný rozboru výběrových souborů, nikoliv však totožný.

Někdy se takto vypočítávají i docela jiné údaje, a proto je nutno upozornit na skryté významové rozdíly. Jestliže se např. „předběžný výpočet“ *výsledku šetření* provádí tři týdny před volbami, kdy nejsou k dispozici žádné skutečně odevzdané hlasy voličů, žádné spočítané volební lístky, ale existuje jen průzkum mínění, napadnutelný již v samých svých základech, dovoluje to pouze tuto výpověď: „Na základě výsledků šetření se může předpokládat, že v den šetření (!) by se voliči rozhodli tak a tak. Tento odhad je s 95% pravděpodobností přesný na ± 2 %.“

Jiným extrémem „předběžného výpočtu“ je, když se např. na základě měrných údajů rakety letící na Mars pomocí počítače „předběžně vypočítává“ celková doba letu. V daném případě se dávají velmi rychle do vzájemných vztahů matematické a fyzikální skutečnosti a nejde již o odhad nebo pravděpodobnost, ale o přísně logické početní závěry: „Jsou-li dány A, B a C a tisíc a jedna jiných měrných hodnot, trvá let přesně 4 měsíce 3 dny 8 hodin a (budiž) asi 10 vteřin až do přistání na Marsu.“ Takový „předběžný výpočet“ není v zásadě nic jiného než závěr tohoto typu: když 3 — 2 = 1, pak 3000 — 2000 = 1000.

Proto nebudeme z opatrnosti používat mnohoznačný výraz „předběžný výpočet", který ani nepochází ze statistiky. Navíc jej vůbec nepotřebujeme. To nás ovšem vede (na základě zásady „z opatrnosti nepoužívat") k jinému slovu, které se v moderní statistice velmi často vyskytuje, a to „*parametr*". Kdybychom porovnali definice, kterých se používá, vyšlo by najevo, že nejsou zcela prosty rozporů. Všeobecně však jde o něco takového jako „*charakteristické hodnoty*": například normální rozdělení vyznačuje parametr *střední hodnoty* a *rozptylu*. Také *medián* a *šikmost rozdělení* jsou takovými parametry.

William B. Bean, šéfredaktor jednoho amerického časopisu, napsal správně v článku proti přehánění při módním používání cizích slov: „Jsou zvučně znějící slova, jako např. »parametr«, avšak z padesáti lidí, kteří toto slovo používají, jeho skutečný význam možná nezná ani jediný." Slyšíte-li tedy nebo čtete slovo parametr, nenechte se jím zastrašit a nehledejte v něm bůhvíco. Pomyslete na charakteristickou hodnotu, to postačí. V této knize se obejdeme bez parametrů. I tak to jde.

Jen tenkrát je nutno při slově „parametr" napnout pozornost, a to když se mluví o „*neparametrickém testu*". Jsou to testy zabývající se testováním výběru bez ohledu na rozdělení základního souboru — mimo jiné proto, že o tomto rozdělení není nic známo nebo že je důvod k domněnce, že není ani normální, ani binomické. Některé testy se proto také nazývají „*testy nezávislé na rozdělení*"; jsou to především tzv. „testy pořadí", o kterých budeme ještě mluvit v desáté kapitole (viz str. 335). Výraz „neparametrický" se vztahuje na dohodu, že se „parametr" ponechá jen pro základní soubor: např. μ a σ jsou parametry, $\bar{x}$ nebo s však nikoliv.

Jiné dnes velmi často používané slovo je „*stochastický*" — a také v tomto případě je člověk v pokušení vzpomenout na Beanovu kousavou poznámku o módě cizích slov. Toto slovo pochází z rané doby počtu pravděpodobnosti (razil je už jeden z bratří Bernoulliů), znovu je objevil na počátku našeho století Bortkiewicz z příkladu o úderech kopyt, ale teprve po druhé světové válce se stalo módním vědeckým výrazem. Neznamená koneckonců nic jiného než „náhodný", „podřízený pouze zákonům náhody". Náhodná rozdělení jako rozdělení normální, Poissonovo a binomické jsou příkladem stochastických rozdělení. Günter Menges uvádí ve svém díle „*Základy statistiky*" lapidárně toto: „Všechno, co je vybudováno na počtu pravděpodobnosti nebo s ním má cokoli společného, se označuje jako »stochastika«."

6 O zacházení s hypotézami

6.1 Pod kontrolou

O pytli s tisícem ořechů, jímž jsme se zabývali již několikrát, jsme předpokládali, že je to dar dobrých přátel. Vlastně jsme z něho neměli ani brát výběrové soubory, protože darovanému koni na zuby nehleď. (To „nehleď na zuby" je také jakýsi druh výběrového souboru: podle stavu zubů se odhaduje stav a stáří „základního souboru" — koně.) Mnohem častěji však půjde o to, abychom se přesvědčili o kvalitě zboží, za které musíme zaplatit hotově. Tím se stává *namátkový* výběrový soubor přesně *dojednanou přejímací kontrolou.* Kupující upozorňuje předem prodávajícího: převezmu zboží jen tehdy, jestliže ve výběru určitého rozsahu nepřekročí počet špatných kusů určitý počet.

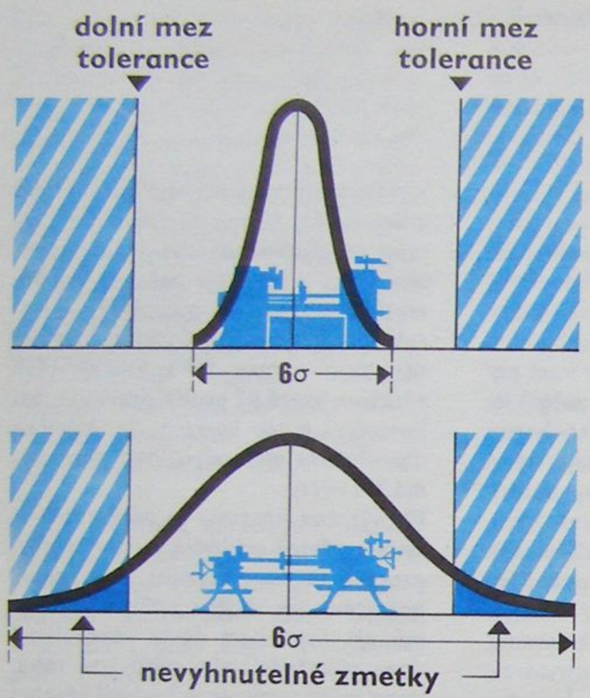
Prodávající by naproti tomu měl předem vědět, na jaký druh přejímací kontroly může přistoupit a v kterém případě se má lákavé objednávky raději vzdát: měl by totiž na základě běžné výrobní kontroly být schopen posoudit, do jaké míry asi jeho zboží odpovídá požadavkům kupujícího.

Přejímací kontrola — výrobní kontrola: na první pohled se zdá, že jde o totéž. Vytvoří se výběrový soubor odpovídajícího rozsahu a na tomto základě se stanoví, jaké množství z daného souboru je bezpochyby příliš lehké nebo příliš malé, příliš těžké nebo příliš velké, příliš křehké nebo příliš nečisté. Přesto jde o dva sice podobné, nicméně zároveň zcela odlišné postupy.

Výrobní kontrola se provádí běžně a má umožnit včas poznat a odstranit vyskytující se výrobní závady — např. seřízením určitého stroje nebo vyřazením zvlášť nepozorného pracovníka z výrobního procesu. *Odběratelská kontrola* se naproti tomu týká kvality těch výrobků, které již prošly sítem výrobní kontroly a o kterých se výrobce domnívá, že plně odpovídají jeho normám kvality.

Při výrobní kontrole je dále nutno si všímat přesně odchylek na *obě* strany, protože zatímco kupující může např. pozastavit jen zboží s nižší váhou, nemá výrobce zájem dávat zákazníkovi dary prostřednictvím zboží nad váhu. Ve výrobní kontrole je heslem: výrobní proces je „*pod kontrolou*". To nemusí bezpodmínečně znamenat, že každý kupující bude spokojen a každá přejímací kontrola uspokojivá. Proces je totiž „pod kontrolou" tehdy, jestliže nevykazuje žádné abnormální výkyvy, ale může se stát, že pro některé zájemce jsou normální, v podniku obvyklé výkyvy již příliš velké. Pak je sice výrobní proces statisticky „pod kontrolou", ale při přejímací kontrole vznikají neustále potíže.

Podívejme se třeba na tento příklad: Stroj vyrábí kovové rámy, které mají mít světlou výšku nejméně 82,5 a nejvýše 83,5 cm. Z první výrobní série se odebere 100 kusů, přesně se změří a zjistí se, že průměr výběrového souboru je $\bar{x} = 83{,}02$ a směrodatná odchylka $s =$ $= 0{,}11$. Protože jde o normální rozdělení, vyskytnou se odchylky za troj-



Jestliže stroj pracuje s menším rozptylem, než jaké jsou meze tolerance odběratele, nemůže ani extrémní odchylka narušit strojní výrobu. Jestliže výrobky daného stroje mají již od počátku tak velký rozptyl, že se meze tolerance stále porušují, může snad být výrobní proces tohoto stroje statisticky „pod kontrolou", ale jakost výrobků není vyhovující.

násobkem směrodatné odchylky jen dvakrát až třikrát při tisíci pozorováních. Bude-li proces probíhat dále tak jako na začátku, můžeme předpokládat s 99% jistotou, že se nevyskytne žádný výsledek měření nad $\pm$ 0,33 cm. Jestliže se však při pozdějších výběrových souborech ukáže, že se podobné výsledky vyskytnou častěji, není již výrobní proces statisticky „pod kontrolou".

A co se má stát teď? Především je třeba klidně rozvážit, co to znamená v praxi podniku. Je katastrofa, že každý stý rám měří např. méně než 82,69 cm (tzn. 83,02 — 0,33)? Jistě ne, protože dokonce ještě do 82,5 cm plní svůj úkol. Vyrábí se proto klidně dál jako dosud. Nanejvýš se budou brát výběrové soubory poněkud častěji než dosud, aby se zjistilo, zda stroj „nejede podle svého" a nevyrábí ze dne na den stále menší (nebo prostě nepravidelnější) rámy.

Mnohem horší je druhý případ: nechť je pro daný účel použití i nadále trpěna tolerance $\pm$ 0,5 cm, ale stroj dává produkci o průměru $\bar{x} = \pm$ 83,2 a $s = 0{,}25$. To však znamená, že horní hranice tolerance bylo již dosaženo při 1,2 σ a pohledem do tabulky normálního rozdělení zjistíme, že celých 11,5 % výroby je nepoužitelných. (Při dolní hranici tolerance není situace tak zlá: od 83,2 do 82,5 cm jsou téměř 3 σ.)

V daném případě je tedy výroba zcela „pod kontrolou", tzn. má normální rozptyl kolem střední hodnoty — a přesto musí být téměř 12 % produkce vyhozeno, sešrotováno nebo roztaveno. Snad by bylo možno pomocí velmi přesné úpravy dojít k žádoucímu $\mu = 83{,}0$, ale i potom by část výroby byla mimo hranice tolerance 82,5 až 83,5 cm: v tomto nejlepším případě by bylo asi 4,5 % zmetků. Stroj již zřejmě není pro požadovanou přesnost dostatečně dobrý, neboť rozptyl je příliš velký. Musí se použít jiný, lepší stroj, neboť stroj, který nepracuje dostatečně přesně, nelze okouzlit ani nejzvučnějšími statistickými formulemi.

6.11 Přejímací kontrola

Odběrateli je zcela lhostejné, co se děje při výrobní kontrole. Pro něho je rozhodující přejímací zkouška, i když někdy se může dohodnout o předání záznamů výrobní kontroly, aby se zjednodušila přejímací kontrola. Je třeba, aby se na podkladě vzorku nalezlo měřítko pro složení celé partie zboží.

Na první pohled se zdá, že ideální je případ, kdy průměrná výrobní kvalita výrobce A přesně odpovídá požadavkům odběratele B. Předpokládejme, že výroba v průměru vykazuje 1 % vadných výrobků a odběratel je ochoten toto 1 % trpět. Má být dodáno 1000 kusů a dohodne se rozsah výběrového souboru $n = 100$. Odběratel je ochoten převzít zásilku i tehdy, jestliže se v tomto výběrovém souboru najde jeden vadný kus. Přesto by byl tento postup pro obě strany nanejvýš nepříjemný. Co se stane? Poissonovo rozdělení udává pro výskyt „řídkého jevu: zmetek" tyto pravděpodobnosti:

žádný zmetek ve výběru:

$$p_0 = \frac{1}{e} = 0{,}37$$

jeden zmetek: $p_1 = \frac{1}{e} \cdot 1 = 0{,}37$

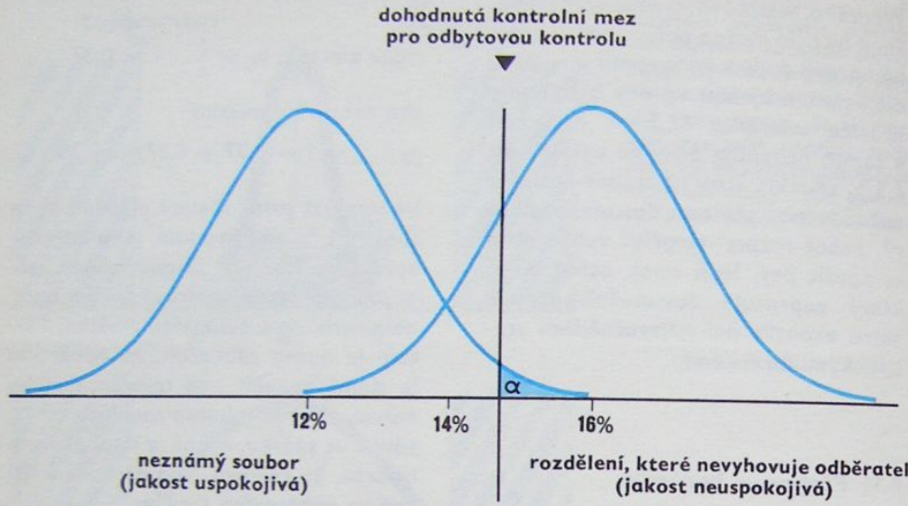
dva nebo více zmetků:

$$p_{2\,\text{a více}} = 1 - (0.37 + 0{,}37) = 0{,}26.$$

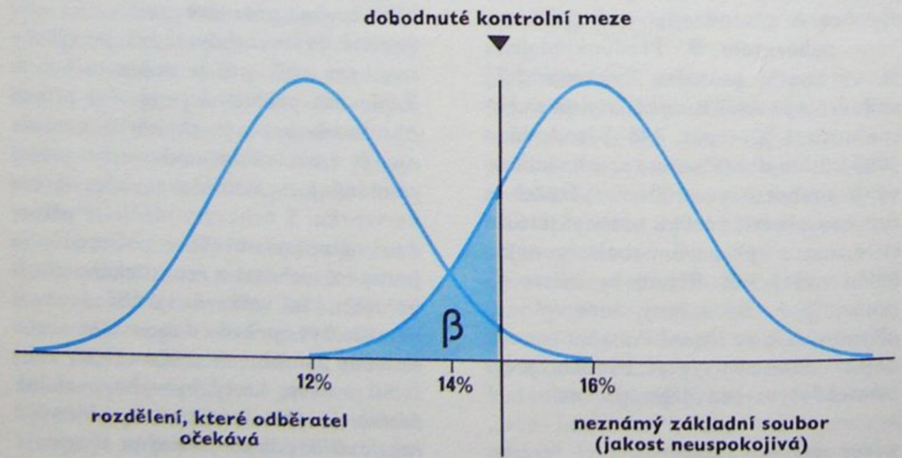
Ve více než jedné třetině případů bude zboží s 1 % vad vypadat jako odpovídající, ve čtvrtině případů bude odmítnuto, protože výběrový soubor bude obsahovat dva nebo více zmetků. Přitom je nutno zdůraznit, že odběratel je stejně neuvěřitelně tolerantní nebo naivní, když je ochoten souhlasit s 1 % závad ve vzorku, neboť se musí právem obávat, že celková četnost bude vyšší než ve výběrovém souboru.

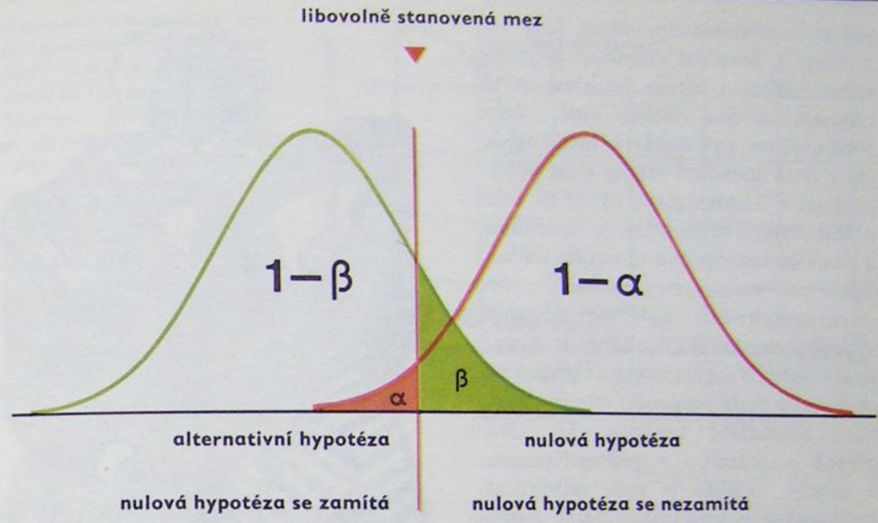
Na základě zásady „výrobní jakost přesně odpovídající požadované jakosti" se na delší dobu nevytvoří prospěšný obchodní styk. Skutečná četnost chyb ve výrobě musí být spíše *nižší* než ta, kterou akceptuje odběratel. Jen tehdy se vytvoří ono mezipásmo, ve kterém hra náhody (náhodný vzorek) může působit, aniž dojde k nadměrnému poškození zájmů některého z obou partnerů.

Snadné je to tehdy, když je výroba mnohem lepší, než je naprosto nutné. Zcela bez problémů je ideální případ absolutně bezvadných sérií, protože není-li v základním souboru ani jediný závadný kus, nemůže se objevit ani ve vzorku. S takovým ideálním případem nelze ovšem vůbec počítat. Spíše je nutno vycházet z realistického předpokladu, že veškeré vyráběné zboží nemůže být opravdu dokonalé, a proto se hledá způsob, jak nalézt takový zkušební postup, který by vyhovoval jak *odběrateli, tak dodavateli a především nebyl příliš nákladný*. Jinými slovy: je



Při šetření výběrových souborů se mohou stát dvě chyby. Výběrový soubor může zcela náhodou vypadat mnohem hůře, než jaký je skutečně celkový soubor — je to takzvaná „chyba alfa" nebo „chyba prvního druhu" (graf nahoře — dobré nebo správné se tím odmítá). Výběrový soubor však může být náhodou také mnohem lepší než základní soubor — „chyba beta" nebo „chyba druhého druhu" (graf dole — špatné, nesprávné, závadné se přijímá).





I testy hypotéz jsou stále ohrožovány chybami alfa a beta. Nulová hypotéza může být zamítnuta, ačkoliv je správná, nebo také nezamítnuta, ačkoliv je nesprávná.

třeba vytvořit přejímací kontrolu, která při pokud možno malém výběrovém souboru poskytne záruku, že odběratel — pokud dodrží svůj výrobní standard (obr. na str. 172) — dostane s největší pravděpodobností uspokojivou jakost a prodávající vysokou pravděpodobnost přejímky bez závad.

Oba, odběratel i dodavatel, musí podstoupit jediné riziko: odběrateli se může stát, že náhodný vzorek je podstatně lepší než skutečná jakost celé dodávky. Výrobci se může přihodit, že sice vyrobil zcela dobré zboží, ale že téměř všechny vadné exempláře proklouznou do výběrového souboru.

Předpokládejme např., že odběratel by byl ochoten akceptovat 2 % zmetků, zatímco výrobce ví, že jeho výroba jich má asi 1 %. V celkovém množství 1000 kusů je tedy jen asi 10 vadných. Vzorek o rozsahu $n = 100$ byl stanoven dohodou. Může klidně očekávat přejímací zkoušku? Ne zcela. Je dobře možné, že ve vzorku jsou tři vadné kusy:

$$p_3 = \frac{1}{e} \cdot \frac{1}{3!} = \frac{1}{6e} = 0{,}06$$

Šest ze stovky uskutečněných výběrů tedy bude obsahovat tři kusy zmetků, přinejmenším jeden z dalších dokonce čtyři nebo více. A k tomu ještě toto: žádný odběratel nebude souhlasit se stokusovým výběrovým souborem se dvěma vadnými kusy, když chce mít jistotu, že v celkovém množství není víc než dvě procenta zmetků.

Dříve než se budeme v krátkosti zabyvat početními postupy, které byly vymyšleny k dosažení vzájemně přijatelného rozložení rizika, musíme ještě objasnit povahu těchto rizik. Brzy totiž uvidíme, že v daném případě nejde jen o čistě obchodní vztahy mezi odběratelem a dodavatelem, nýbrž že celá oblast výběrové analýzy je ovládána a doslova zastiňována tématem rizika. Výběr poskytuje — to již konečně víme — jen upozornění, nikoliv však závazné výpovědi o celkovém souboru. *Nevykazuje* zpravidla přesně stejné proporce chyb nebo zcela správnou střední hodnotu základního souboru. To jsme přesně prokázali i v jednotlivostech u ořechů. Jestliže je tedy odběratel např. ochoten souhlasit s 15 % špatných ořechů a výběrový soubor vykáže 13 %, podstupuje stále ještě jisté vypočitatelné riziko, že bude mít v celkové dodávce přece jen více než 16 % vadného zboží.

Zcela opačný případ nastává, když vzorek vykazuje 16 % zmetků a odběratel právem dodávku odmítne, ačkoli by další zkouška ukázala, že skutečný podíl špatných ořechů činí dokonce jen 12 %.

Tento případ odmítnutí dobrého, správného nebo odpovídajícího na základě nesprávného vzorku se označuje jako výskyt *„chyb prvního druhu“*, někdy také nazývaných *„chyby alfa“*. Opačný případ, tj. přijetí špatného, nesprávného nebo neodpovídajícího na základě příznivého vzoru, je takzvaná *„chyba druhého druhu“* nebo *„chyba beta“*. Představují jakousi Scyllu a Charybdu analýzy výběrových souborů. Jak se jim lze vyhnout? (Viz obr. na str. 173.)

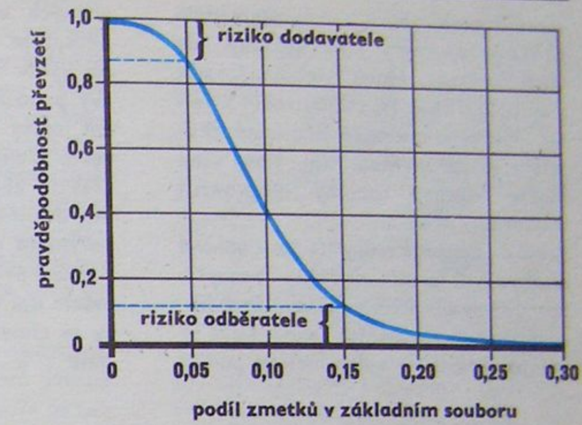
Jednu známou pravdu jsme již uvedli: není-li žádný exemplář vadný, nemůže se objevit ani ve výběrovém souboru. To je případ nejlepšího žáka: ať padne

Každá zkouška je výběrovou analýzou. Z „bedny“ mozku jsou taženy „kuličky“ vědomostí nebo nevědomosti a z toho se usuzuje na celkový stav vědomostí v mozku.

při zkoušce jakákoliv otázka, může na ni odpovědět. (Protože každá zkouška je přirozeně vzorkem — výběrem ze zásob znalostí zkoušeného.)

Druhý extrémní případ jsme již naznačili. Jestliže je kupující ochoten převzít 6 % špatného zboží (nebo uzná-li toleranci ± 1 cm), zatímco výrobce má jen 1 % zmetků (případně výrobní toleranci ± 0,3 cm), bude mít přejímací zkouška skoro vždy dobrý průběh. Třetí extrémní případ jsme již také načrtli. Jestliže zákazník žádá jako minimální kvalitu to, co je výrobce schopen dodávat jako průměrnou jakost, bude výběr skoro vždy uvádět v omyl a uspokojivý obchodní styk bude prakticky nemožný. Realita všedního dne a hospodářského života sotva zná tyto extrémy. Abso-

Plánovaná křivka zkoušky, operační charakteristika podle zobrazení Wallise — Robertse. Dodavatel věří, že nevyrábí více než 5 % zmetků, odběratel je ochoten převzít maximálně 15 % zmetků. Rozsah výběrového souboru je 40 a smějí v něm být nanejvýš tři vadné kusy. Při této přejímací kontrole jsou rizika odběratele a dodavatele přibližně stejně velká (v obou případech něco přes 10 %).



lutně bezvadná sériová výroba zůstává zbožným přáním. Kupující, který převezme zboží ze 6 % špatné (a podle toho zaplatí), nepůjde k výrobci, který vyrábí nepoměrně lépe (a podle toho si dá zaplatit). Kdo musí odmítnout každou druhou dodávku pro nedostačující jakost výběrového souboru, bude si muset hledat velmi brzy jiného dodavatele.

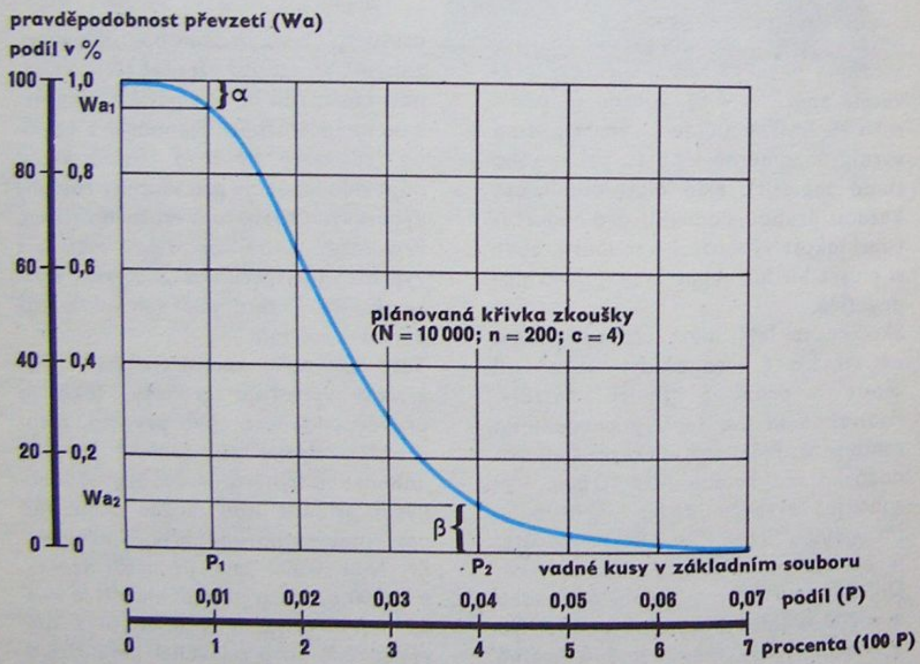
Skutečnost leží mezi těmito dvěma extrémy a v této oblasti se rozvíjí teorie a praxe přejímací kontroly. Podívejme se pro lepší pochopení na tento případ Hanse Kellerera: Celková dodávka má rozsah $N = 10\,000$. Výrobce je přesvědčen, že jen 1 % je vadné ($M = 100$, $M/N = P = 0{,}01$); odběratel je ochoten souhlasit se 4 % zmetků. Dojedná se výběrový soubor o rozsahu $n = 200$ a ještě přijatelný podíl chyb $m/n = 0{,}2$, tedy maximálně 4 špatné z 200.

V tomto případě se vypočítává „chyba prvního druhu“ (výrobcovo nebezpečí odmítnutí) ve výši 5 %, „chyba druhého druhu“ (riziko odběratele při nejisté kvalitě celkového souboru zboží) ve výši asi 10 %. Nechceme se v podrobnostech zabývat vzorci pro výpočet těchto hodnot. Na rozdíl od normální křivky neexistuje postačující jednoduchá tabulka, ze které by bylo možno vyčíst odpovídající údaje pro všechny rozsahy výběrových souborů a možného rizika. Pro každý jednotlivý případ musí být vypracována specifická *„plánovaná křivka zkoušky“*, která může mít velmi rozdílnou podobu.

Tato *„operační charakteristika“ (OC-křivka)* vyjadřuje graficky, jaká je pravděpodobnost chyb prvního, resp. druhého druhu při zvolené metodě náhodného výběru. V každém jednotlivém případě jsou možné četné (až nesčetné) plánované křivky zkoušky. Při $N = 10\,000$ jsme již určili vzorek $n = 200$ a v něm jsme pokládali $m = 4$ za nejvýš přijatelnou hodnotu s tím výsledkem, že prodávající nese riziko ve výši 5 % a kupující ve výši 10 %. Místo toho bychom mohli zvolit $n = 230$ (a ponechat $m = 4$); z toho by vyplynula poněkud změněná křivka — s větším rizikem pro výrobce a menším pro kupujícího. Mohli bychom zvolit menší rozsah výběru, což by znamenalo pro oba partnery nižší náklady, ale větší nejistotu. Mohli bychom dosadit $n = 1000$ a $m = 19$, což by vedlo k velmi „citlivé", ale také nákladné zkušební metodě. Mezi tím jsou však možné všechny rozsahy výběrových souborů.

Jde-li o „*nespojité znaky*" (také v obecně alternativní formě „dobrý — špatný"), nesmí se zvolit příliš malý rozsah výběrového souboru zvláště tehdy, když se má dosáhnout jen velmi malého podílu vadných exemplářů. Jestliže kupující akceptuje nejvýš 1 % zmetků, nebude asi výběr v rozsahu $n = 100$ uspokojivý, protože i když je v takovém výběru jen jediný vadný kus, má kupující již velké přejímací riziko. Prakticky může převzít zboží jen na základě úplně bezvadného vzorku, a to je zase velká nevýhoda pro výrobce.

V běžné praxi přejímací kontroly nejde ovšem ani tak o to najít postup, který by se choval přesně stejně ke kupujícímu i k prodávajícímu a který by přesně rozdělil riziko, ale přistupují k tomu faktory nematematické povahy. V závislosti na pozici, z níž se vyjednává (stav na trhu atd.), dosáhne někdy jeden, jindy druhý obchodní partner příznivějších přejímacích podmínek. Pro oba je ovšem důležité, aby přesně znali operační charakteristiku dohodnutého postupu zkoušení: co se na první pohled jeví jako fair, ukáže se při bližším zkoumání jako značně unfair, jak jsme již konečně viděli.

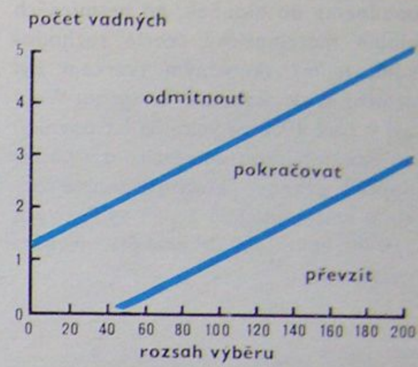


Operační charakteristika (plánovaná křivka zkoušky) pro výběrový soubor $n = 200$, ve kterém smějí být až čtyři ($c = 4$) vadné kusy. Levá stupnice udává pravděpodobnost převzetí výběru. Jestliže skutečný podíl vadných exemplářů je jen 1 %, vypočítá dodavatel své riziko (pomocí chyby alfa) na 5 %. Pro odběratele je riziko (chyba beta) asi 10 %, protože v souboru jsou 4 nebo více procent vadných exemplářů, ačkoliv vzorek udával jen 2 % (totiž 4 mezi 200). (Podle vyobrazení H. Kellerera.)

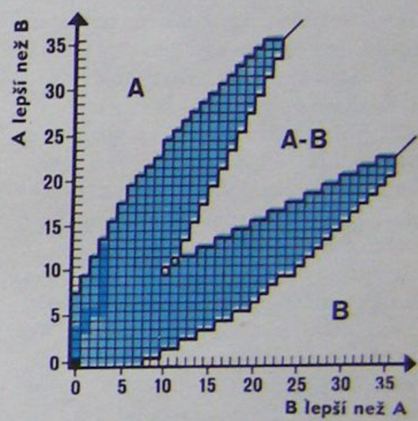
Někdy se také dohodnou „*sekvenční výběry*" (česky se též ne zcela přesně říká „postupný výběr"), protože se tak dosáhne zmenšení rozsahu vzorku, aniž se tím riziko příliš zvětší. Teoretické základy sekvenčních výběrových šetření, které vypracoval ve čtyřicátých letech v USA Abraham Wald, jsou z nejnovější doby a velmi obtížné.

Načrtneme jen krátce základní myšlenku, která je stejně jednoduchá jako zřejmá. Nejdříve se vybere poměrně malý vzorek a na jeho základech se ihned rozhodne v případech, kdy se dosáhlo extrémních hodnot: jestliže se např. ve vzorku o rozsahu $n = 50$ nenajde ani jeden vadný kus, dodávka se převezme. Jsou-li v něm dva nebo více zmetků, dodávka se odmítne. Jestliže je vadný jen 1 kus, rozšíří se vzorek o další kusy, přičemž musí být předem dohodnuta kritéria pro tento výběrový postup. Může se třeba předem dojednat rozšíření $n = 100$. Nevyskytne-li se přitom další vadný kus, dodávka se přijme; vyskytnou-li se dva vadné, odmítne se; vyskytne-li se jeden, provede se rozšíření např. na $n = 140$, a pak se již nesmí vyskytnout žádný vadný kus.

Ale tak, jak je zásada jednoduchá, je složitý přesný propočet rizika chyb prvního a druhého druhu. Nezbytně



Sekvenční výběrová analýza podle vyobrazení u Wallise-Robertse. Vyskytnou-li se dva vadné exempláře dříve, než rozsah výběrového souboru dosáhne $n = 40$, je zásadně odmítnut. Jestliže však naproti tomu při $n = 100$ je stále ještě jen jediný vadný exemplář, přijímá se.



Sekvenční výběrové analýzy podle vyobrazení Haseloffa a Hoffmanna. V řadě testů se srovnávají A a B. Vždy když je A lepší, postoupí se o jedno pole nahoru, když je lepší B, o jedno pole vpravo, jsou-li stejné, pokračuje se po diagonále dál. Na našem obrázku bylo při prvních třech pokusech A lepší, při čtvrtém lepší B, při pátém a šestém zase A, při sedmém a osmém B, při devátém a patnáctém opět A. Ukáže-li se A lepší i při dvou následujících pokusech, uzavírá se experiment s výsledkem „A je lepší než B".

upadneme do hloubek, ba nesmírných hlubin matematické teorie rozhodování, jejímž skutečným tvůrcem byl ostatně Wald. Rakouský emigrant Wald byl v roce 1938 ve věku 36 let povolán na Kolumbijskou univerzitu a v následujících dvanácti letech revolučně rozvinul matematickou teorii výběrových šetření. Roku 1950 zahynul při leteckém neštěstí.

6.12 Kontrolní karty

Z hlediska teoretického pojednání jsou mnohem pochopitelnější kontrolní karty, o kterých si nyní řekneme ještě několik slov. Jsou to různé formy záznamů výsledků vzorků z běžné výrobní kontroly; většinou se používá poněkud pozměněný systém souřadnic ve formě karet. Souřadnici tvoří dosavadní střední hodnota a každý výsledek měření se postupně zaznamenává nad nebo pod tuto „ideální čáru". Mimoto běží většinou s ní rovnoběžně po dvou čarách nad a pod souřadnici takzvaná „výstražná mez" nebo „akční mez". *Výstražná mez* probíhá převážně ve vzdálenosti 2σ od průměru, *kontrolní* nebo *akční mez* zpravidla $2{,}58\sigma$. To znamená, že pokud výrobní proces běží normálně, bude 95 % měřených hodnot uvnitř výstražných mezí a jen 1 % vně akčních mezí. Stejně dobře lze však říci, že rozsah spolehlivosti (95 %) probíhá mezi oběma výstražnými mezemi, nebo také že pravděpodobnost chyb (1 %) označuje prostor vně akční meze; při měření mimo tyto hranice se proto setkáváme s „významnými" odchylkami. (Viz obr. na str. 180.)

Kontrolní karta má však tu přednost, že ukáže i takové změny ve výrobním postupu, které ještě nevyvolávají nutnost „akce" — jako např. nepatrné posunutí průměru nahoru nebo dolů (ovšem uvnitř výstražných mezí).

Někdy se ukazuje užitečné zaznamenávat do kontrolních karet nikoli jednotlivé hodnoty měření, ale střední hodnoty z nejmenších vzorků. Výhoda je v tom, že jednotliví „zbloudilci" jsou do velké míry zneškodněni již uvnitř příslušného nejmenšího vzorku a nemusejí vyvolávat zbytečné starosti. Pak ovšem musí mít výstražné i akční meze také užší rozpětí, protože průměry vzorků mají menší rozptyl než jednotlivé hodnoty měření.

K ilustraci použijeme malého příkladu. Normovaná hodnota u kontrolní karty nechť je 100, rozptyl $\sigma = 3$, výstražné meze probíhají pak u $(100 - 6) = 94$, resp. $(100 + 6) = 106$, kontrolní mez u 92,2 a 107,8. Nyní máme těchto patnáct výsledků měření:

100,2	100,6	101,2
103,4	97,3	99,1
106,8	94,4	92,2
91,8	102,4	108,2
100,8	93,8	95,6

Máme dvě hodnoty nad (106,8 — 108,2), 3 hodnoty pod výstražnou mezí (92,2 — 91,8 — 93,8), 2 z toho dokonce mimo kontrolní mez (108,2 — 91,8), třetí přesně na kontrolní mezi (92,2). Nemůže tedy být pochyb: proces není statisticky pod kontrolou, všechno jde páté přes deváté.

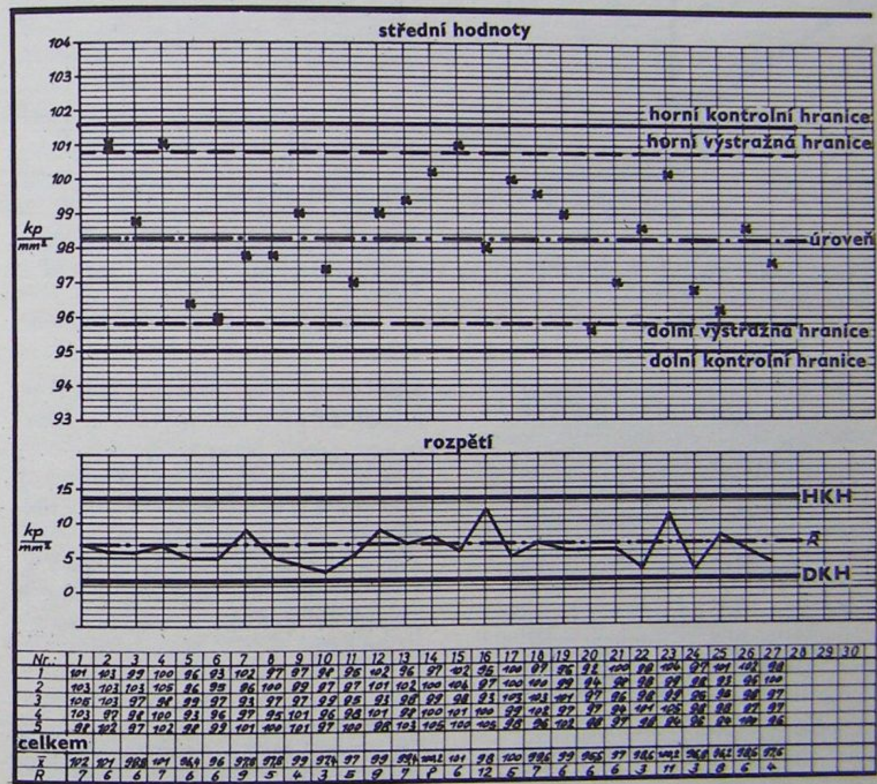
Vytvořme naproti tomu průměr vždy ze tří po sobě jdoucích měření, a už to vypadá mnohem lépe: 100,7 — 99,9 — 97,8 — 100,8 — 96,7.

Všechny hodnoty leží pohodlně uvnitř našich starých výstražných mezí. Ty však právě pro tento případ již nejsou vhodné a měly by probíhat mnohem blíž u sebe.

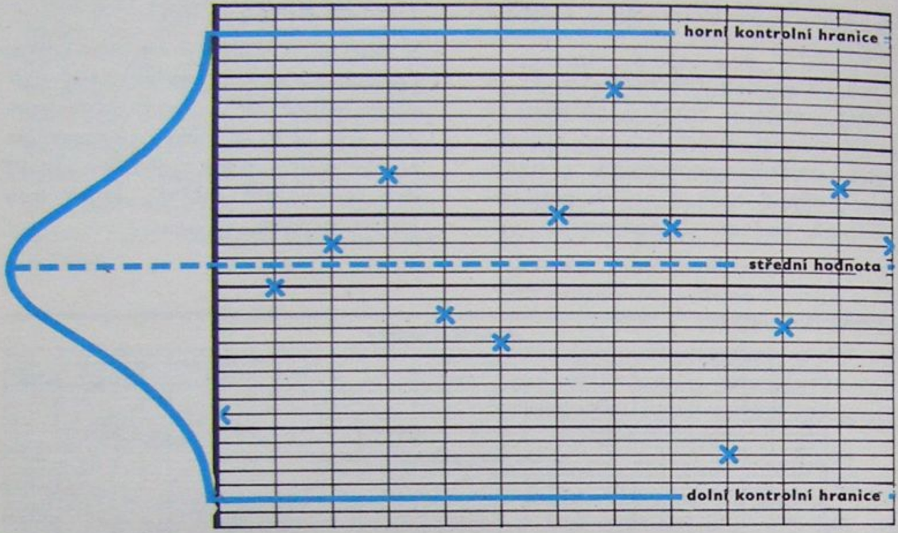
Konečně existují takzvané *R-karty* (R z angl. range — rozpětí), na nichž se uvádí rozpětí uvnitř jednoho výběrového souboru nejmenšího rozsahu. Zůstaneme-li opět u našich pěti tříčlenných hodnot, dostaneme:

1,0 — 6,1 — 14,6 — 16,4 — 7,0.

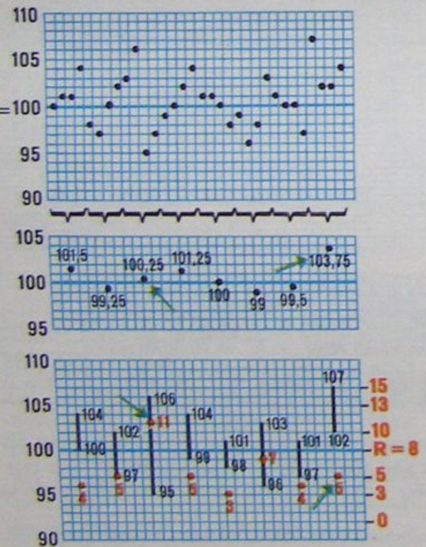
V daném případě se nic nezakrývá, můžeme se odvolat na normální rozdělení jednotlivých hodnot měření: rozpětí přes 12 by se mělo vyskytovat jen zcela zřídka — a my tu máme z pěti již dvě intervalová měření, která jsou podstatně vyšší.



Příklad použití kontrolní karty pro výrobní kontrolu průmyslového podniku. Měří se pevnost hřídele zadní nápravy. Záznamy se týkají průměru z pěti přezkoušených hřídelí zadní nápravy (horní kontrolní karta), případně jejího rozpětí v pěti jednotlivých měřeních (karta rozpětí dole). Střední hodnota („úroveň") výroby je 98,3 kp/mm².
Pozoruhodné je, že tato kvalita by musela být označena jako neodpovídající normě: kdyby pět zkušebních kusů výběrového souboru bylo úplně stejných (a snad dokonce na „úrovni"), kleslo by rozpětí na 0 a prolomilo tak spodní kontrolní hranici. Jestliže totiž náhoda nevyužije prostor, který jí byl poskytnut, je třeba vždy zvýšit pozornost.



Obrázek ukazuje souvislost mezi kontrolní kartou a normální křivkou. Měrné hodnoty blízké kontrolní hranici (HKH — horní, DKH — dolní kontrolní hranice) jsou extrémy a měly by se vyskytovat zřídka.



Proč akční mez probíhá právě na $2{,}58\,\sigma$? Za prvé tam neprobíhá, nýbrž se tam prostě často kladla (mohla by také být podle potřeby na 2,5 nebo $3\,\sigma$), a za druhé je to vítězství myšlení v procentech nad normálním rozdělením; uvnitř $\pm\, 2{,}58\,\sigma$ leží 99 %, mimo 1 %. Z těchto důvodů je také výstražná mez přesně na $+\,1{,}96\,\sigma$ a nikoliv na $2\,\sigma$ (= 99,5 %).

Tři kontrolní karty s měrnými údaji, které jsou rozptýleny kolem střední hodnoty 100. Nahoře 32 měření, pod tím stejné údaje ve skupinách po 4 měřeních — průměr výběrového souboru se rozptyluje nápadně méně než jednotlivá měření. Dole se měří na „R-kartě" rozpětí osmi čtyřčlenných skupin. Zde je mimo jiné zřetelně vidět, že třetí skupina (průměr výběrového souboru 100,25, tedy zcela nenápadný) je složena z extrémů — rozpětí činí 11!

6.2 Jednoduché testy hypotéz

Hypotéza znamená doslovně předpoklad či domněnku, že něco by mohlo být tak a tak nebo vysvětleno tak a tak. Hypotéza není ještě ani tezí, větou sloužící jako důkaz, jako tvrzení. Je to domněnka, která může vzniknout z okamžitého nápadu nebo může být vypreparována po dlouhých úvahách z určité pokusné řady: „Bylo by přece docela dobře možné, že..." nebo: „Předpokládejme, že je možná souvislost...".

Máme-li takovou hypotézu, je věcí temperamentu a svědomitosti, zda se plným hlasem prohlašuje za novou moudrost vyvracející všechny námitky, nebo zda se nejdřív potichu uvažuje, jak by bylo možno správnost dané hypotézy ověřit. Mezi těmito dvěma extrémy je ještě mnoho možností, z nichž jedna je zvlášť oblíbená: s úmyslem podpořit hypotézu se shromažďuje vše, co slouží k jejímu posílení, a kriticky se zkoumá vše, co ji vyvrací.

Kdo by však chtěl svoji (nebo také jinou) hypotézu zkoušet, bude se muset dříve nebo později uchýlit ke statistice — ledaže by šlo o hypotézu z oblasti, kde je sestavení statistických údajů zvlášť obtížné. S pomocí statistiky nebude např. možné tezi, že bohové byli kosmonauty, popřít, ale ani ověřit. Podobně je tomu s jinými historickými hypotézami, a to ze zcela zvláštního důvodu. Důkladné a přesné ověřování hypotéz vyžaduje testovací materiál, který je nutno opatřit jako doplněk k materiálu, jenž byl podnětem k vytvoření hypotézy. Jinak vzniká nebezpečí, že se pomocí matematických vzorců vybuduje pouze impozantní konstrukce kolem předem přijaté domněnky.

Matematické testování hypotéz je poměrně nový způsob statistické práce. Pochází z let krátce před vypuknutím druhé světové války a vytvořili jej především Jerzy Neyman a E. S. Pearson a dále pak rozvinul Abraham Wald. Od Neymana a E. S. Pearsona pochází také myšlenka testování hypotézy i představa o chybách prvního, resp. druhého druhu.

První test hypotéz však pochází již z těch raných dob statistiky, kdy se právě osvobozovala od biblické kletby a kdy se začal statisticky opěvovat božský pořádek. Britský lékař a satirik John Arbuthnot vykonstruoval *statistický důkaz boží existence*, když v roce 1710 prohlásil: V posledních 82 letech se v Londýně rodilo pravidelně více chlapců než děvčat. To nemůže být náhoda, protože kdyby narození chlapců a děvčat odpovídalo výskytu hlavy a orla při házení mincí, byla by tato stálá převaha chlapců nemyslitelná. Je tedy nutno v tom vidět boží ruku, která ve své moudrosti řídí rozdělení narozených tak, jak je to správné. Tato úvaha má *nedostatky teologické, pravděpodobnostně teoretické i logické*, ale v zásadě jde o testování hypotézy. Hypotéza „normální rozdělení narozených $p = q = 0{,}5$" byla vyvrácena faktickým materiálem.

Uveďme jiný příklad: v jedné věznici zjistím, že padělatelé směnek mají velmi vysoký kvocient inteligence. Vytvořím hypotézu, že tito podvodníci jsou nadprůměrně inteligentní. Jestliže použiji jen tento známý „materiál", mohu sice předložit svůj výsledek pozorování v nanejvýš působivé formě jako „statistiku", ale opravdové testování hypotézy lze provádět teprve tehdy, až prozkoumám nový materiál, např. vězně z jiné věznice nebo nový přírůstek podvodníků se směnkami v dalších třech letech.

Lze ovšem nalézt i jiné východisko, a to takové, že se řekne: Promiňte, prosím, ale vždyť já jsem od samého začátku měl hypotézu — ještě dřív než jsem se šetřením začal. Tato hypotéza spočívala v tom, že úroveň inteligence padělatelů směnek by se „rozdělila“ stejně jako v celku všech vězňů (nebo také celého obyvatelstva). Potom však výsledky mého šetření tuto „*nulovou hypotézu*“ vyvrátily a svůj výsledek mohu tedy označit jako výsledek ověřování hypotézy.

Ještě dříve než řekneme více o této důležité nulové hypotéze a o vypovídací schopnosti testů hypotéz, podíváme se na okamžik nazpět — na to, co jsme uvedli v předchozích oddílech. Ukáží se totiž pozoruhodné paralely toho, co nyní velkolepě nazýváme „ověřování hypotéz“.

Vždyť už jsme sami přece hypotézy ověřovali! Např. když jsme zkoumali pytel ořechů a přitom jsme si na počátku řekli, že řádově bude asi 20 % zkažených. Potom jsme udělali náhodný výběr o vhodném rozsahu, a tím jsme uvedenou hypotézu ověřovali.

Ještě jasnější to bylo v situaci, když jsme se zabývali kontrolními kartami a přejímacími zkouškami — jedno ověřování hypotéz za druhým. Na kontrolní kartě je ověřovaná hypotéza dokonce zakreslena: zde musí být střední hodnota, až sem a ne dál může být rozptyl, nemá-li být vyvrácena hypotéza „všechno probíhá normálně jako dosud“. A co se děje při přejímací zkoušce? Odběratel vlastně testuje tuto hypotézu: „Chtějí mne doběhnout, zboží je špatné“ — a rozhodne se pro převzetí zboží tehdy, je-li tato hypotéza vyvrácena. Dodavatel ověřuje hypotézu „mé zboží je vysoce hodnotné jako vždy“ a je zděšen, ukáže-li se něco jiného.

Nebylo vám nápadné, že jsme vytrvale mluvili o tom, že hypotéza se zamítá?

Neměli jsme být mnohem pozitivnější a říci: hypotéza se potvrdila? Ne, bohužel ne. Celé ověřování hypotéz spočívá koneckonců na tom, že se hypotéza zamítne — a to nikoliv bez důvodů. Máme totiž správnou míru pro nepravděpodobné: např. to, co je mimo trojnásobnou směrodatnou odchylku, má pravděpodobnost jen $p = 0{,}003$. Jestliže v této oblasti najdu více hodnot měření, než se tam „hodí“, mohu říci: tento druh normálního rozdělení nevystihuje situaci. Hypotézu „$\mu = 100$, $\sigma = 3$“ mohu zamítnout, jestliže mezi 100 měření bude 8 vyšších než 109. Ale co se tím pozitivně dokazuje? Hypotéza $\mu = 110$, nebo $\mu = 107$? Nebo může být hypotéza $\mu = 100$ správná, ale σ je mnohem vyšší, než se předpokládalo? Měrné hodnoty mohou upozornit, jaké by asi mohlo být rozdělení, ale pozitivně nic nedokazují.

Nejspíše je ještě možno vymezit oblast tím, že se např. konstatuje, že hypotézu $\mu \leqq 100$ (slovy: menší/stejná) i hypotézu $\mu \geqq 118$ (slovy: větší/stejná) lze pokládat za zamítnuté. Někde mezi tím musí být skutečný průměr. Kde přesně? Proto je možno uvést nanejvýš „*maximum likelihood*“ (někdy se překládá jako „*maximální věrohodnost*“), ale žádný důkaz. *Hypotézy je proto možno zásadně prohlásit za falešné (zamítnout je), nikoliv však dokazovat.*

V praxi se tyto těžkosti obvykle rády přehlížejí a předpokládá se toto: jestliže se nulová hypotéza zamítne, bude asi správná moje alternativní hypotéza. To je většinou oprávněno v těch případech, kdy alternativní hypotéza má již předem široký akční prostor nebo jestliže je v povaze věci, že otázka je omezena na prosté „buď — anebo“.

Tak je tomu především tehdy, jde-li např. o zkoušení nového léku, nového postupu apod. V daném případě nulová hypotéza zní takto: „Vše zůstane při starém, nový postup není ani lepší, ani horší než starý.“ (Zde je také etymologický základ „*nulové hypotézy*“, která říká, že změna se „*rovná nule*“.) Jestliže se tedy nulová hypotéza zamítne, může se alternativní hypotéza „je účinný“, resp. „je horší“ pokládat za potvrzenou a neurčený je pouze stupeň zvýšení účinnosti.

Především však nesmíme zapomenout na jedno: Protože se ověřování hypotéz zakládá na zkoumání výběrových souborů, podléhá také náhodným chybám prvního a druhého druhu. Proto se může vždy stát, že správná hypotéza bude zamítnuta a nesprávná přijata — především tehdy, když jde vůbec jen o chybné hypotézy. Jestliže se totiž ověřují jen nesprávné hypotézy, nemusí náhoda příležitostně jednu z nich označit za falešnou. Jestliže se naproti tomu ověřují jen hypotézy správné (alternativní), může sice jedna z nich být nedopatřením označena za falešnou, celkový výsledek je však uspokojující. Bylo tedy právem řečeno: „Statistické testy mají pouze jakýsi filtrační účinek, který způsobuje, že správné z přijatých hypotéz jsou poněkud obohaceny.“

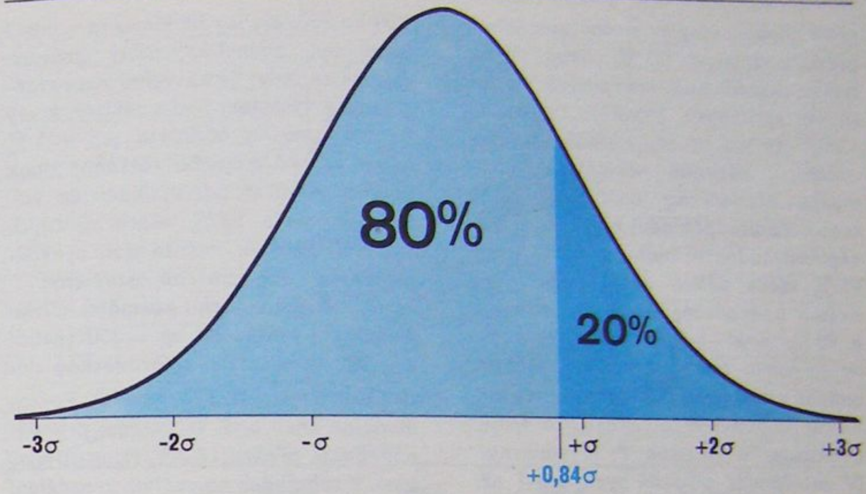
Postupy ověřování hypotéz jsou většinou počtářsky obtížné a nebudeme je proto uvádět v jednotlivostech. Některé z nich krátce nastíníme v poslední kapitole naší knihy, ale ten, kdo chce studovat matematické podrobnosti těchto testovacích postupů, může použít některé z učebnic statistiky, uvedených v příloze (viz str. 347—348).

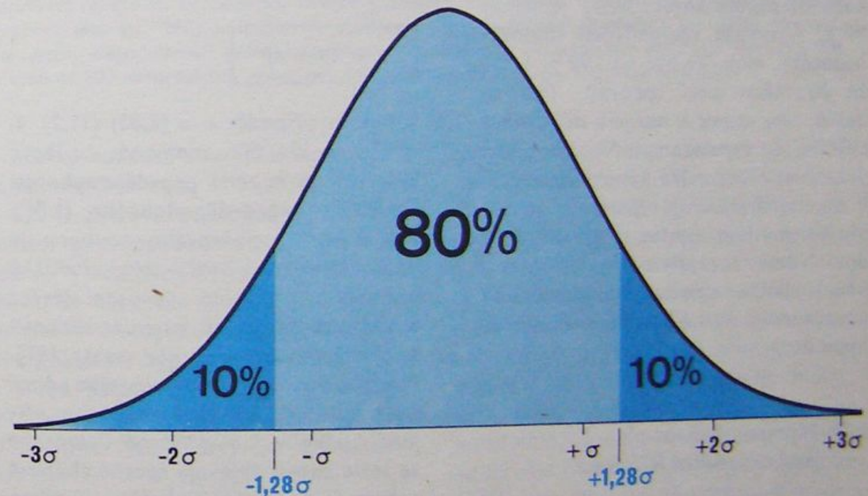
To však nevylučuje, abychom se alespoň v hrubých rysech neseznámili se základními myšlenkami ověřování hypotéz a statistického testování. Vzpomeňme si na výrobce A a kupujícího B, který nechce koupit zboží obsahující více než 2 % zmetků, přičemž si výrobce A myslí, že má svou výrobu pod kontrolou a že jeho zboží nemá více než 1 % zmetků.

Kupujícímu záleží na tom, aby „chyba druhého druhu“, „odběratelské riziko“, byla pokud možno malá. Jinak řečeno, záleží mu na tom, aby hypotéza „*více než 2 % zmetků v souboru*“ byla vyvrácena na tak dlouho, jak jen možno. Neboť pak — a jen tehdy — může předpokládat, že se zároveň potvrdí jeho alternativní hypotéza: „obsahuje 2 % nebo *méně* zmetků“. V tomto a v mnoha jiných případech jde o „*jednostranný test hypotéz*“. Otázka zní pouze, zda bude překročen určitý nejnižší nebo nejvyšší podíl. To však při normálním rozdělení znaků znamená, že zájem je soustředěn jen na jeden konec normální křivky. Chci například vědět, kolik šroubů je příliš krátkých (nikoliv příliš dlouhých), zda podíl chyb je více než 2 % (nikoli však, zda je pod 0,5 %) atd. To ale znamená, že se přípustná pravděpodobnost omylu soustřeďuje na tom konci křivky, který mne zajímá. Zatímco tedy při dvoustranném testu, který má dosáhnout 99% jistoty, se na každém z obou konců normální křivky považuje 0,5 % za „extrém“ a vede k zamítnutí hypotézy, soustřeďuje se při jednostranném testu — za předpokladu stejné 99% pravděpodobnosti — toto 1 % na *jedné straně*. Vzpomeňte si, jaké překvapující rozdíly jsme nejdříve měli (v části 3.33) při

Jestliže se „prosévají" jen správné hypotézy, odmítne se někdy i správná. Ověřují-li se jen nesprávné hypotézy, přijme se někdy i některá nesprávná. Ověřování hypotéz se obecně užívá ke zvýšení relativního počtu správných hypotéz tím, že většina nesprávných se zamítne a většina správných naproti tomu zamítnuta není.

vysvětlování tabulek normálního rozdělení a jak se v jedné tabulce říkalo, že „pravděpodobnosti se uvádějí pro horní konec" (viz obr. na str. 185).
Víme, že při uvažování obou konců hranice 99 % probíhá u $\pm 2{,}58\,\sigma$. Tabulka s *jednostranným* údajem pravděpodobnosti má však hranici 99 % u $2{,}33\,\sigma$, protože teď se soustřeďuje nepatrná „extrémní" plocha pod jedním koncem křivky, zatímco dříve byla pravidelně rozdělena vpravo a vlevo. Stejně je, jak známo, 95 % souboru mezi $+2\sigma$ a -2σ (přesně mezi $+1{,}98\,\sigma$ a $-1{,}98\,\sigma$), při uvažování *jedné* části křivky naproti tomu mezi $\pm 1{,}645\,\sigma$.
Zde je nutno co nejdůrazněji varovat





Rozdělení plochy pod normální křivkou se může provést dvojím způsobem. Pozorují-li se jen extrémní hodnoty v jednom směru (v našem příkladu jde asi o 20 %), shromažďují se na jednom konci; příslušná tabulka bude uvádět podíl plochy 0,80 pro $z = 0{,}84$. Mají-li se naproti tomu uvažovat extrémy v obou směrech, rozděluje se 20 % v našem příkladu zrcadlově na dvakrát 10 % a tabulka bude uvádět, že zbývající plocha 0,80 je omezena pro $z = 1{,}28$ (na obou stranách). Jsou-li v tabulkách uvedeny oba způsoby výpočtu, ukáže se první zobrazení většinou jako $\Phi(z)$, druhé jako $D(z)$. Zpravidla jde o omezení extrémů v obou směrech.

před nebezpečným nedorozuměním: pravděpodobnost 95 % nebo 99 % anebo každá jiná nevypovídá vůbec nic o správnosti hypotézy, naopak. Vždyť tím se označuje podíl volného prostoru, většinou normálního rozdělení. Myšlenkový postup je tedy tento: Za předpokladu, že μ nebo P má nějakou hodnotu, mohu si být jist, že 95 % všech dílčích údajů bude mezi dvěma směrodatnými odchylkami — a 99 % mezi $\pm$ 2,58 σ. Jestliže však najdu měrné hodnoty, které je možno umístit v tomto 99% prostoru, neznamená to ještě 99% jistotu, že moje domněnka o μ nebo P je správná. (V extrémním případě bych došel až k 100% „jistotě" — postačí, když přijmu každou hodnotu.)

Uvedených 99 % při testování hypotéz říká: Jestliže výsledek náhodného výběru umožňuje závěr, který je mimo 99 % intervalu spolehlivosti (nulové) hypotézy, mohu si být na „99 % jist", že hypotéza není správná. Tím se dostáváme znovu k našemu dřívějšímu zjištění, že *hypotéza nemůže být přímo dokázána, nýbrž může být jen zamítnuta jí odporující (nulová) hypotéza.*

Na to musíme myslet vždy, když se domníváme, že můžeme hypotézu dokázat. Ukážeme to na příkladu tím, že provedeme početně jednoduché ověření hypotézy.

6.21 Narození chlapců pod normální křivkou

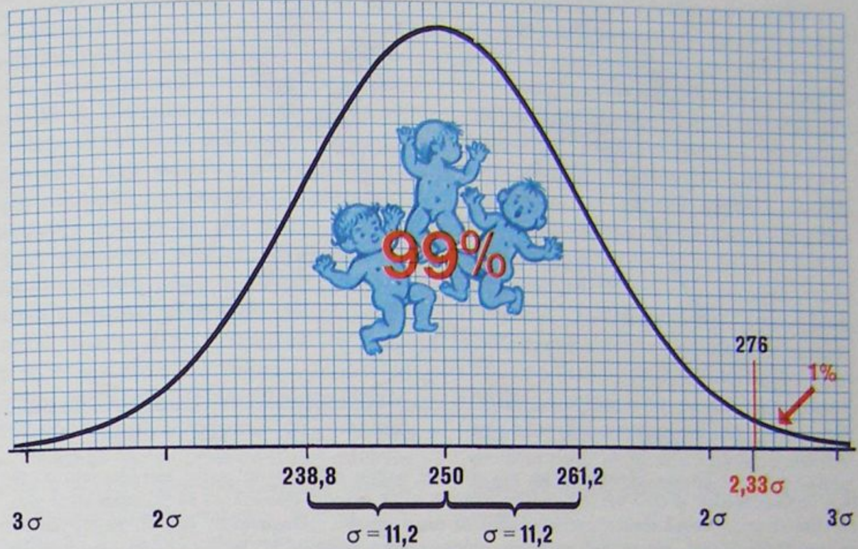
Ověřujeme hypotézu, že je stejný počet narozených chlapců a děvčat. Jako podklady náhodného výběru máme 500 po sobě jdoucích zápisů v matrice narozených, z nichž je 266 chlapců a 234 děvčat. Přesná rovnováha by byla 250 ku 250, ale my již víme, že náhodě musí být ponechán volný prostor. Ptejme se tedy: Je to velmi nepravděpodobný výsledek, třeba takový, který by bylo možno očekávat jen v 1 % všech vzorků stejného rozsahu? Jinak řečeno: Hodí se náš výsledek do velkého souboru 99 % všech možných výsledků vzorků, jestliže jsou pravděpodobnosti stejnoměrně rozděleny?

Načrtneme si v duchu normální křivku se střední hodnotou $np = 250$ (neboť $n = 500$, $p = 0{,}5$) a směrodatnou odchylkou $\sqrt{npq} = \sqrt{125} = 11{,}2$. Potom hledáme onen bod, ve kterém pravděpodobnost překročí 99 % (*jednostranný test*). V tabulkách normálního rozdělení najdeme pro jednostranný test u 99 % hodnotu $z = 2{,}33$. Nyní vypočítáme kritickou hodnotu, kterou označíme c, podle vzorce

$$\boxed{c = z\sigma + np}$$

V našem případě: $c = (2{,}33)\,(11{,}2) + 250 = 276$. To znamená: hodnota 276 leží na hranici pravděpodobnosti (99 %) a nepravděpodobného (1 %). 266 chlapců výběrového souboru je však daleko „uvnitř" této kritické hodnoty — hypotéza „narození děvčat a chlapců je stejně pravděpodobné" se tím jeví potvrzena (obr. na str. 187). Popravdě je ovšem tato hypotéza pouze *„nezamítnuta"*, což není totéž jako *„přijata"*, ačkoliv i v odborné literatuře se stále znovu objevuje špatné chápání tohoto „páru protikladů". Pojďme o krok dále a ptejme se, zda by se takový výsledek objevil i mezi 90 % „neextrémních" náhodných výsledků, kdyby p bylo opravdu 0,5.

K tomu zopakujeme dřívější výpočet, přičemž nyní vezmeme a dosadíme



Ve výběrovém souboru 500 novorozeňat bylo 266 chlapců a 234 děvčat. Je možno tento výsledek považovat za důkaz, že narození děvčat a chlapců není přesně rovnoměrně rozděleno? Krátký propočet (viz str. 186) ukazuje, že hypotézu „rovnoměrného rozdělení" tím nelze vyvrátit. Pro „nepravděpodobnou" nebo „zamítnutou" hypotézu by přinejmenším mělo být 277 nebo více chlapců na 500 novorozeňat, neboť pak činí pravděpodobnost jen 1 % nebo $p = 0{,}01$.

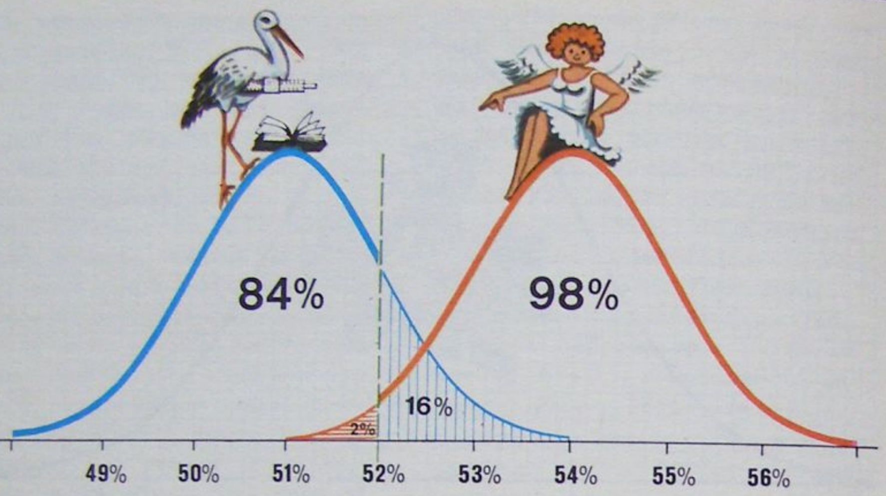
z tabulky hodnotu $z = 1{,}28$. Nová kritická hodnota $c_{90\,\%} = (1{,}28)\,(11{,}2) + 250 = 264$. Náš počet 266 je již mimo prostor pravděpodobnosti 90 % všech vzorků vymezených číslem 264. Na této úrovni se hypotéza zamítá.

Jestliže se v rozporu s dobrým svědomím pokusíme hypotézu „verifikovat", je třeba jistotu číst obráceně: 99 % neznamená více *bezpečnosti a jistoty*, nýbrž spíše více *velkorysosti a lhostejnosti*.

Považuji-li však domněnku stejné četnosti narození za nulovou hypotézu, lze říci: s 99% jistotou nemohu sice tuto nulovou hypotézu zamítnout, ale mohu to udělat snad s jistotou poněkud větší než 90 %. A dále: Skutečná četnost narození chlapců bude vyšší než 0,5. Kde přesně? Budu potřebovat větší vzorky a pak např. navzájem testovat hypotézy $p_1 = 0{,}52$ a $p_2 = 0{,}54$, abych poznal, která je pravděpodobnější, resp. která se dá dřív vyvrátit.

Z výběrových souborů o rozsahu $n = 500$ a $n = 2000$ bychom samozřejmě nikdy nezískali tak přesná čísla, k jakým se došlo v průběhu staletých pozorování na celém světě. Víme, že podíl chlapců na živě narozených činí ve střední Evropě (a podobně na celém světě) asi 51,5 %. Hypotéza $p = 0{,}5$ je tedy nesprávná a je možno ji — s příslušně velkým výběrovým souborem — kdykoli vyvrátit.

Dříve než budeme dále rozvádět myšlenku, jak se v tomto případě navzájem chovají základní soubor, vzorek a alter-



Typická nákupní situace přenesená na náš příklad podílu chlapců na počtu narozených. Anděl jako „objednávající" si chce být jist, že čáp nedodá příliš mnoho chlapců — 52 % nebo ještě více, obsažených ve „vzorové zásilce", se mu zdálo již příliš mnoho. „Dodavatel" čáp ví, že v průměru dodává 51 % chlapců, ale při tak přísných „přejímacích podmínkách" by bylo 16 % jeho pětisetkusové zásilky odmítnuto, ačkoliv „celková produkce" zcela odpovídá požadavkům. Otázka zní vždy znovu: jak se rozdělí riziko přejímky mezi dodavatelem a odběratelem? Zde na obrázku neriskuje anděl téměř nic.

nativní hypotézy, zabývejme se námitkou, kterou by mohli vznést velmi kritičtí, pečliví a pozorní čtenáři. Bylo by totiž možno oprávněně se zeptat, proč jsme právě před chvílí počítali vlastní „kritické hodnoty", když jsme přece již (od třetí kapitoly) mohli s mnohem menší námahou přesně vypočítat, jaký podíl pravděpodobnosti připadá na hodnoty pod 266, resp. nad 266 narození chlapců. Mohli jsme přece jen standardizovat normální rozdělení podle známého schématu:

$$\frac{\text{hodnota pozorování — hypotetický průměr}}{\text{směrodatná odchylka}}$$

V našem případě by vyšlo:

$$\frac{266 - 250}{11{,}2} = 1{,}425.$$

V tabulce standardizovaného normálního rozdělení bychom vyhledali 1,425 a dostali podíl plochy 92,3 %. Proč tedy ta rozvláčná srovnávání s 99, resp. 90 %?

Vysvětlení spočívá prostě v tom, že jsme chtěli „uvést" metodu, která sice v daném případě nebyla potřebná, ale již brzy bude jako jediná k dispozici — jakmile se totiž blíže seznámíme se *„Studentovým rozdělením t"*, které je založeno na *hodnocení výsledku srovnáváním s tabulkovou kritickou hodnotou* (viz část 6. 41).

Nyní se z těchto hledisek podívejme na scénku z „pohádky". Jedná se o dodávku dětí, kterou zařizuje příslušný anděl. V rámci populační exploze potřebuje nakoupit pro nejbližší dva týdny dva milióny dětí. Zmocněnec fy Čáp a spol. promptně dodávku přislíbí: „Je Vám známo, že náš podíl chlapců

činí spolehlivě téměř přesně 51 %." Anděl je však opatrný: „Měli jsme již několik reklamací. Chtěl bych mít úplnou jistotu, že podíl určitě nebude vyšší než 54 %. Dodejte nám nejdříve vzorek v rozsahu řekněme 500 kusů." Pak jsou povoláni dva podnikoví statistikové, aby dohodli podrobnosti přejímací zkoušky. Přitom by nás mohla napadnout zlomyslná otázka, jak by vůbec mohl statistik přijít do nebe (statistikové minulosti však byli, jak už víme, velmi zbožní lidé). Statistik fy Čáp a spol. nyní uvažuje takto: „Je-li normální rozdělení naší produkce — vztaženo na vzorek o rozsahu 500 — kolem četnosti 51 %, se směrodatnou odchylkou 1 %, pak jen dva z každé stovky předpokládaného vzorku budou mít podíl 53,3, neboli více chlapců." (To nemusí dlouho počítat, najde to v tabulce.) Nebeský statistik si řekne: „Rozhodně nechceme 54 % nebo více chlapců a musíme proto zajistit, aby nebyla převzata celá dodávka na základě pochybného vzorku. Je-li tedy produkce střední četnosti rozptýlena kolem 54 % se směrodatnou odchylkou 1 %, znamená pro nás podílová četnost 52 % nebo ještě méně ve vzorku to, že odběratelské riziko činí ještě celá 2 %, protože dva ze sta vzorků by mohly náhodou tak nízkého podílu dosáhnout."

Nebeský statistik je však člověk dobromyslný a svému partnerovi navrhne: „Dobře, jsme ochotni převzít celou dodávku, jestliže ve výběrovém souboru nebude podílová četnost chlapců větší než 52 %." Znalec fy Čáp a spol. nahlédne do tabulky a ztrápeně praví: „To od nás opravdu nemůžete žádat. Podívejte se přece sám: při této podmínce převzetí bychom museli počítat s tím, že pouhou hrou náhody s 16% pravděpodobností ztratíme objednávku, ačkoliv slíbenou jakost dodržíme." Nechme už nezodpovězenu otázku, zda se dohodli, zda oba měli stejnou naději na úspěch nebo zda pozice vyjednávání jednoho či druhého byla silnější, takže mohl vnutit partnerovi větší riziko. Jisté však je, že náš dřívější vzorek podle původního návrhu hodného nebeského statistika nevyhnutelně vedl k odmítnutí. K tomu by dokonce vedlo i „férové" rozdělení rizika. (Čáp a spol. měli právě s tímto vzorkem mimořádnou smůlu — podíl chlapců byl neobyčejně velký a mimoto firma špatně odhadovala svou produkci: skutečný průměr není 51 %, ale 51,5 %, a takový omyl již může vést ke značně chybným kalkulacím.)

6.22 Pravděpodobnost a hypotézy

V praxi ověřování hypotéz jde tedy většinou o takovou sestavu testu, aby zamýšlené úrovně odmítnutí bylo pokud možno přesně dosaženo. Je samozřejmě možné (pokud je k dispozici libovolně velký rozsah výběrového souboru) dosáhnout extrémní kritické hodnoty, a tím pravděpodobnosti 99 a 99,9 %, avšak k tomu je často třeba mít velký výběrový soubor, zvláště když se testované střední hodnoty nebo četnosti navzájem jen nepatrně odchylují. Proto se musíme nejdříve ptát, jaká je skutečně požadovaná přesnost testu, jinak vznikají zbytečné náklady a práce. Jestliže se testované hodnoty navzájem silně odchylují, postačí již malý výběrový soubor k tomu, aby bylo možno odmítnout nulovou hypotézu s postačující přesností. I bez provedeného propočtu je zřejmé, že nulová hypotéza („průměrná délka bude 5 cm") může být velmi rychle a beze zbytku odmítnuta výběrovým souborem $\bar{x} = 8$ cm, zatímco výběrový soubor $\bar{x} = 5{,}5$ cm bude dlouho svým rozdělením pronikat do rozdělení nulové hypotézy.

Testování hypotéz odpovídajícím způsobem je mimo jiné nutné při řešení těchto otázek: Odpovídá výrobek, který má být zkoušen, daným požadavkům na čistotu, tvrdost nebo zmetkovitost apod.? (Nulová hypotéza: Neodpovídá.) Je nový výrobek lepší než dosavadní? (Nulová hypotéza: Není lepší. Dodatečný požadavek: Musí se projevit jako „významně" lepší.) Dovoluje vzorek o průměru $\bar{x}$ závěr, který bude odpovídat výběru z určitého základního souboru o průměru μ? (Nulová hypotéza: *Nepochází* z tohoto základního souboru. Nebo, jde-li o jiný problém: Ověřuje se, zda toto $\bar{x}$ s malou pravděpodobností omylu nemohlo být převzato ze základního souboru.) Utvořil jsem hypotézu o určitých souvislostech mezi měrnými údaji — nebude experimentem vyvrácena? Jestliže nikoliv, potom: Mohly tyto výsledky vyjít, aniž by proto daná hypotéza ztrácela platnost?

V těchto případech může jít o ověřování středních hodnot, rozptylů, podílů, způsobů rozdělení, korelací atd. Téměř vždy bude důležitý rozsah výběrového souboru: S rozsahem vzrůstá přesnost všech testů hypotéz, zároveň však rostou náklady.

Při praktickém používání testů hypotéz vzniká proto vždy otázka, jaká má být spolehlivost výsledků, s jakou jistotou má být nulová hypotéza vyvrácena. Pfanzagl k tomu říká: „Základní pravidlo zní: hrozí-li velká ztráta, nepoznáme-li *skutečně existující* rozdíl, zvolí se malá pravděpodobnost. Hrozí-li velká ztráta, *předpokládá-li se existence rozdílu*, pak se zvolí velká pravděpodobnost." (Viz obr. na str. 191.)

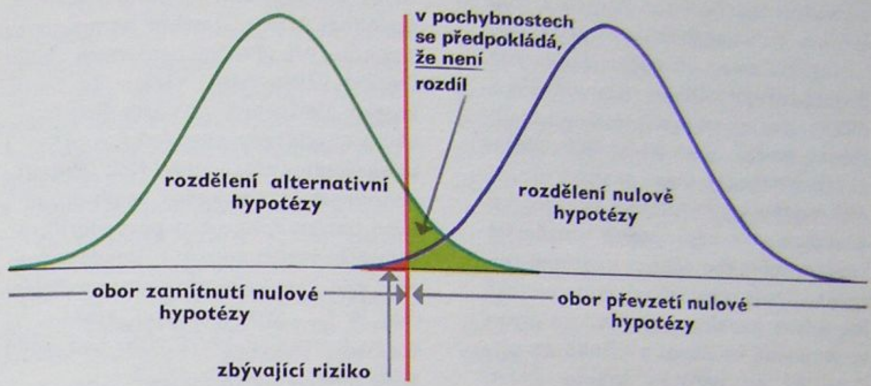
Toto tvrzení alespoň stručně vysvětlíme. V prvním případě předpokládejme, že vynaložím 100 000 DM, bude-li se nadále používat chemikálie A místo jen nepatrně lepší chemikálie B. V daném případě není nutno nulovou hypotézu („B není lepší než A") testovat příliš ostře, neboť lze říci: bude-li jen z 90 % — nikoliv však z 95 % — odmítnuta, je B snad přece jen lepší.

V druhém případě vynaložím 100 000 DM, jestliže použiji dražší B, protože se mylně domnívám, že je lepší než A. V daném případě budu velmi přísně ověřovat, zda hypotéza „B také není lepší než A" musí být zcela rozhodně odmítnuta. Teprve tehdy, mohu-li s 99% nebo větší jistotou říci, že nulová hypotéza může být pokládána za vyvrácenou, rozhodnu se pro nákladný přechod na B.

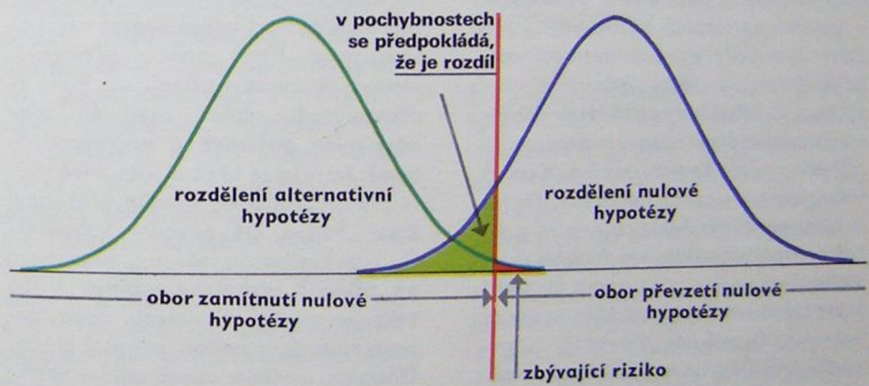
Později se ještě vrátíme k některým speciálním případům testování hypotéz. Nejdříve se však seznámíme se starým a neobyčejně svérázným postupem.

6.3 Bayesova věta a její použití

Ve většině učebnic statistiky se *Bayesova věta* neuvádí vůbec anebo jen zcela povrchně a jméno Thomase Bayese se neuvádí ani v rejstříku. Hned vysvětlíme, proč tomu tak je. Ať je názor na Bayesovu větu jakýkoliv, nelze popřít jeho význam pro teorii testování hypotéz. Tento anglický duchovní z druhé poloviny 18. století je jedním z předků a praotců induktivní statistiky, a tím i celé teorie výběrových soubo-



Podle potřeb situace se požaduje více nebo méně přísné vyvrácení nulové hypotézy. Kdyby to znamenalo vysoké náklady nebo nebezpečí, když se nulová hypotéza odmítne příliš snadno, budeme ji zkoušet velmi přísně a ponecháme jí jen nepatrný prostor ze zbývajícího rizika (obr. nahoře). Vzniknou-li však vysoké náklady nebo nebezpečí spíše tehdy, jestliže se správná alternativní hypotéza nedopatřením odmítne, omezí se toto riziko tím, že se nulová hypotéza bude testovat jen s malou pravděpodobností jistoty (obr. dole).



rů. V době, kdy počet pravděpodobnosti byl ještě v dětských střevíčcích, zabýval se již problémem, jak by bylo možno uvádět pozdější zkušenosti do souladu s původními předpoklady, resp. jak takové zkušenosti změní současné hodnocení situace proti původním předpokladům. Jde tedy do jisté míry o dynamické ověřování hypotéz se zabudovanou korekturou hypotéz, může-li se to tak říci.

Přestože přísně matematická statistika uvrhla Bayesovu větu do klatby, prožívá právě v posledních letech své znovuzrození. Profesor Pitman z Hopkinsovy univerzity řekl při příležitosti jubilejních oslav na počest Laplaceho, Bayese a Bernoulliho: „Jsou velké rozdíly v názorech na teorii statistické inference, umění nebo vědy činit platné závěry z náhodně pozorovaných prvotních hodnot... Pokoušíme se dělat výpovědi o více nebo méně neznámých rozděleních nebo takové výpovědi testovat. To je abstraktní matematická formulace problému. Z hlediska praxe to znamená usuzovat z účinků na příčiny. Až do počátku tohoto století spočívala nejdůležitější metoda řešení daného problému v tom, že se propočítalo podmíněné rozdělení příčin a posteriori, vycházející z daného nebo předpokládaného apriorního rozdělení. Nyní máme testy významnosti, obory spolehlivosti, důvěrné rozdělení, ale vzniká nespokojenost s těmito moderními nástroji a opět vážně uvažujeme o použití apriorních rozdělení."

Ještě dnes tedy existují důslední obhájci Bayesovy věty. Jedním z nejvýznamnějších je Leonard Savage. Vládnoucí učebnicový názor matematické statistiky však Bayese odmítá. Kromě toho existuje jakési zmatení pojmů o tom, co Bayes řekl, co myslel a co bylo pozdějším výkladem do jeho úvah vloženo. K tomu přistupuje podivně staromódní způsob, jímž byla uvedena tato věta (která není identická s *Bayesovým problémem* a je tudíž takto nazývána neprávem).

Říká přibližně toto: Představme si, že v pytlíku je 9 kostek — tři pravé a šest falešných (z nichž u tří padá číslo 6 vícekrát, u dalších jen zřídka). Vytáhnu-li jednu jinak stejnou kostku a víckrát házím, mohu na základě dosažených výsledků postupně snížit svou původní nejistotu (vytažení každé ze tří druhů kostek je stejně pravděpodobné) a říci: „Jestliže jsem s touto kostkou při dvaceti pokusech hodil šestkrát číslo šest, věřím, že to je kostka zfalšovaná na číslo šest."

Místo abych tedy původně („a priori") s pokrčením ramen musel říci: „Pravděpodobnost $^1/_3$ pro každý ze tří druhů", snad potom řeknu („a posteriori", po tomto krátkém testu): „Pravděpodobnost pro kostku se zfalšovaným číslem šest je $^2/_3$, pro správnou kostku $^1/_4$ a pro kostku, u které číslo šest padá zřídkakdy, je pouze $^1/_{12}$."

To všechno vypadá velmi vykonstruovaně a křečovitě (konečně jako všechny příklady s házením a tažením kostek a koulí v teorii pravděpodobnosti), avšak základní koncept je možno rozšířit daleko za obor vymyšlených a vykonstruovaných her. Myšlenkový pochod lze zcela obecně vyjádřit takto: Nejdříve („a priori") mám vždy o dalším problému, o každé otázce nějakou více nebo méně jasnou představu. Jsem-li nucen se rozhodnout, budu se chovat podle těchto, snad velmi neurčitých myšlenek a pravděpodobností. Jestliže se však dovím nové skutečnosti týkající se daného problému, které budou mé původní hodnocení situace korigovat, mám a posteriori, na základě těchto skutečností (nebo také jen pověstí a tvrzení), odpovídajícím způsobem změnit původní postoj. Dokonce nevím-li vůbec nic — například budu-li tázán, zda rychlost šíření neonového světla v pivu je větší nebo menší než slunečního světla ve vodě (Savageův příklad) — má se apriorní pravděpodobnost rovnat úplné rovnosti: $p = q = 0{,}5$.

Tato apriorní pravděpodobnost se změní dodatečnou informací. Např. z ilustrovaného časopisu se dozvím, že se světlo nešíří v žádné jiné tekutině tak rychle jako v mořské vodě. Na základě dosavadních zkušeností sice vědecké úrovni listu příliš nevěřím, nicméně přece jen bych byl nyní spíše nakloněn k domněnce, že neonové světlo se v pivu šíří pomaleji než sluneční světlo ve vodě, ačkoliv mimo jiné stále ještě nevím, zda se neonové světlo vůbec nešíří pomaleji než světlo sluneční. Předpokládám: a) světlo je světlo, b) tvrzení časopisu snad přece jen vychází z vědeckých pramenů. V důsledku toho se mé původní hodnocení $p = q = 0{,}5$ změní na $p = 0{,}6$, nebo dokonce na $p = 0{,}8$.

Početní schéma Bayesovy věty naznačíme na příkladu z lékařské statistiky. To není pouhá libovůle, vždyť právě v diagnostice pomocí samočinných počítačů se dnes Bayesova věta s oblibou používá a stále více se stává základem automatické diagnostiky. Představme si pacienta, který určitě trpí nemocí A nebo nemocí B. Podle našich dosavadních znalostí (anamnéza, klinický stav atd.) předpokládám, že pravděpodobnost nemoci A je 0,8, naproti tomu nemoci B jen 0,2. Pro získání dalších poznatků necháme, dejme tomu, provést zkoušku enzymů v séru, o níž víme, že u nemoci A je v 90 % případů pozitivní, u nemoci B naproti tomu jen ve 20 % případů. Test ukáže negativní výsledek. Jak je tím ovlivněna diagnóza?

V zájmu jednoduchosti ponechme stranou otázku, jak má být tento test vážen. Bylo by možno zaujmout stanovisko, že úplně mění situaci: protože v případě B vyjde s 80% pravděpodobností negativně a u A jen v 10 % případů, jde zřejmě jistě o B, ať už byla naše předchozí domněnka jakákoliv. Nicméně chceme tuto laboratorní informaci posuzovat „na stejné úrovni" s dosavadními domněnkami.

Poznamenejme ještě jednou: výchozí pravděpodobnost (apriorní pravděpodobnost) pro $A = 0{,}8$, pro $B = 0{,}2$. Do jaké míry je nyní pravděpodobné, že jde o A, jestliže laboratorní výsledek je negativní? Zřejmě: $(0{,}8) \,.\, (0{,}1) = 0{,}08$. Pro B (laboratorní výsledek negativní) dostáváme: $(0{,}2) \,.\, (0{,}8) = 0{,}16$. Tyto hodnoty z laboratorních zkoušek nyní dosadíme do vzorce takto:

$$\frac{0{,}08}{0{,}08 + 0{,}16} = \frac{8}{24} = \frac{1}{3}.$$

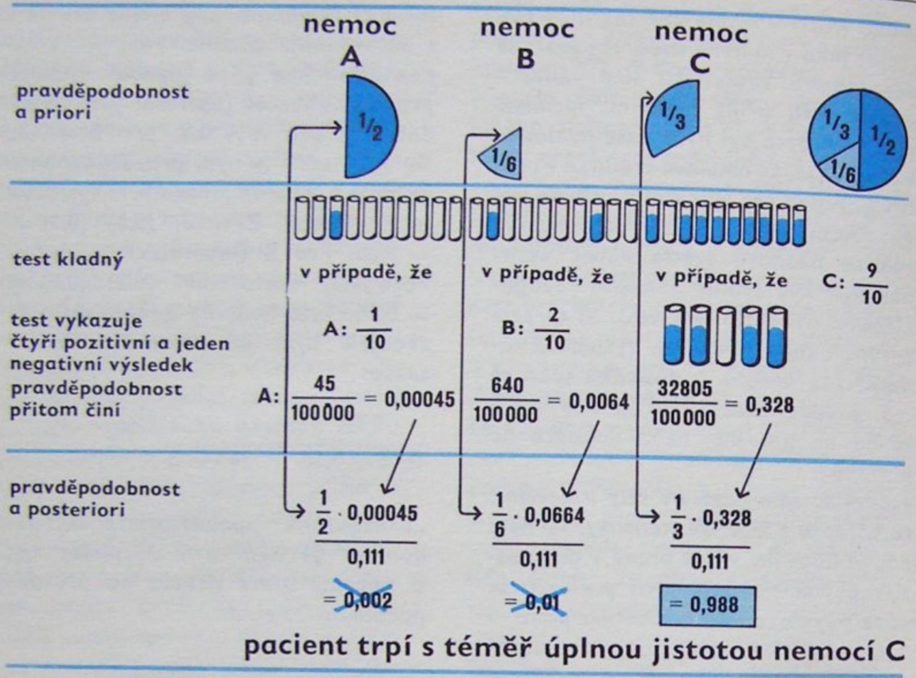
„Korigovaná" aposteriorní pravděpodobnost je nyní pro A pouze $^1/_3$. B naproti tomu získalo na pravděpodobnosti, neboť:

$$\frac{0{,}16}{0{,}08 + 0{,}16} = \frac{2}{3}.$$

Problém stejného druhu, ale již poněkud složitější, uvádějí ruští matematici Gněděnko a Chinčin. V jejich příkladu jde o tři nemoci, které označíme A, B, C a speciální lékařská vyšetření se provádějí pětkrát, přičemž čtyři výsledky jsou pozitivní a jeden negativní.

Výchozí údaje (apriorní pravděpodobnost nemocí) jsou tyto: $A = {}^1/_2$, $B = {}^1/_6$, $C = {}^1/_3$. Lékařský test je v případě nemoci A pozitivní pravděpodobností 0,1 ($p = 0{,}1$), v případě nemoci B je $p = 0{,}2$ a v případě nemoci C naproti tomu $p = 0{,}9$. Již víme, že z pěti testů vyšly čtyři výsledky pozitivní a jeden negativní. Do jaké míry je takový výsledek pravděpodobný?

V případě A (čtenář si jistě vzpomíná na naše pravděpodobnosti v kartách, uvedené ve třetí kapitole):



Schematické znázornění Bayesovy korektury hypotéz. Zpočátku se předpokládala pravděpodobnost, že pacient trpí nemocemi A, B nebo C, z $^1/_2$, případně z $^1/_6$ nebo z $^1/_3$. Na základě nových informací o známých pravděpodobnostních hodnotách vycházejí „potom" (a posteriori) podstatně změněné pravděpodobnosti diagnózy.

$$p_a = \binom{5}{4} \cdot (0{,}1)^4 \cdot 0{,}9 = 45/100\,000$$

V případě nemoci B dostaneme:

$$p_b = \binom{5}{4} \cdot (0{,}2)^4 \cdot 0{,}8 = 640/100\,000 =$$
$$= 32/5000$$

Naproti tomu v případě C je nanejvýš pravděpodobný tento výsledek:

$$p_c = \binom{5}{4} \cdot (0{,}9)^4 \cdot 0{,}1 = 32\,805/100\,000$$

Nyní musíme ještě uvést výsledky laboratorních šetření do vztahu k původním domněnkám o pravděpodobnosti nemocí. Vždy vychází zlomek, v jehož čitateli je apriorní pravděpodobnost nemoci násobená odpovídající pravděpodobností výsledku testu. Ve jmenovateli je ve všech třech případech součet tří čitatelů. V našem případě to znamená:

při výpočtu pravděpodobnosti a posteriori pro nemoc A je v čitateli:

$1/2 \,.\, 45/100\,000 \approx 0{,}0002$;

při B je v čitateli:
$1/6 \,.\, 32/5000 \approx 0{,}0011$;

při C je v čitateli:
$1/3 \,.\, 32\,805/100\,000 \approx 0{,}109$.

Součet těchto tří hodnot činí asi 0,111. A nyní dosazujme:

Aposteriorní pravděpodobnost pro A:

$$\frac{0{,}0002}{0{,}111} \approx 0{,}002.$$

Aposteriorní pravděpodobnost pro B:

$$\frac{0{,}0011}{0{,}111} \approx 0{,}01.$$

Aposteriorní pravděpodobnost pro C:

$$\frac{0{,}109}{0{,}111} \approx 0{,}988.$$

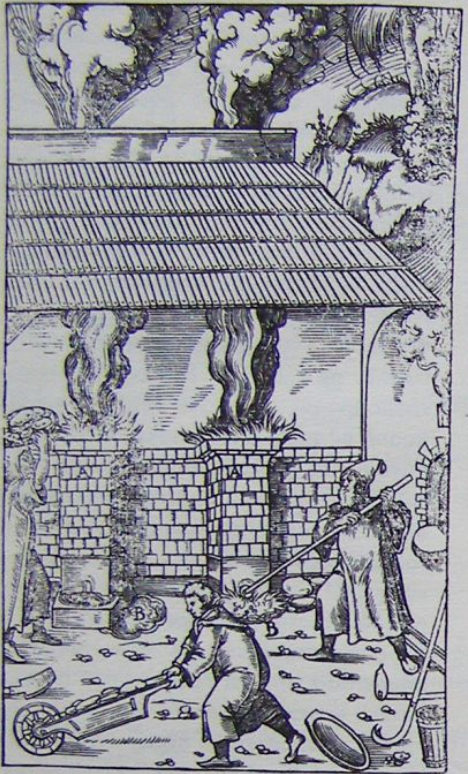
Je tedy téměř nezvratně jisté, že pacient má nemoc C. Vícenásobné opakování testu lze samozřejmě nahradit pěti (nebo také více) šetřeními, která mají vždy odlišnou pravděpodobnost pozitivního výsledku pro každou ze tří nemocí, ovšem výpočet se stává stále složitější. Není proto náhodou, že se Bayesova věta znovu uplatnila a že je uznávána v diagnostice pomocí samočinných počítačů: především proto, že jde většinou jen o odhady, a za druhé proto, že samočinné počítače si snadno poradí i s nejsložitějšími početními operacemi.

Je samozřejmé, že také u Bayese platí, že závěry nemohou být lepší než premisy, že výsledek nemůže být spolehlivější než odhadované pravděpodobnosti. V praxi je to však bohužel často tak, že pro apriorní pravděpodobnost jsou k dispozici jen zcela nespolehlivé odhady nebo protichůdné údaje. V diagnostice pomocí počítačů jsou např. žádoucí co nejpřesnější pravděpodobnosti tohoto druhu: „Jde-li o ledvinové kameny, jak často má pacient vysoký tlak?", „Kolik procent ledvinových kamenů je zjistitelných na rentgenovém snímku?" atd. Pro všechny tyto a četné jiné relativní četnosti jsou však dosud k dispozici jen odhadované hodnoty, takže jejich mnohonásobným použitím roste obava z nepřesnosti výsledků.

Převážná většina matematických statistiků, na rozdíl od pracovníků teorie rozhodování, Bayesovu větu zcela rozhodně odmítá, a to především z těchto důvodů: a) protože nemá pevnou teoretickou základnu, b) protože dosazené apriorní pravděpodobnosti, jak jsme také konečně viděli, jsou většinou určovány „od oka" a zakládají se na poměrně nečetných pozorováních.

Bayesovi stoupenci na tuto námitku odpovídají tím, že řada údajně objektivních metod konvenční vědy obsahuje mnoho subjektivních názorů a mínění. Jak se založí pokus, jaký výsledek se pokládá za „významný", a vůbec celá řada rozhodujících dílčích otázek je ponechána statistikovi na uváženou. K tomu však v praxi ještě přistupuje, jak stále znovu zdůrazňují stoupenci subjektivní pravděpodobnosti, že pro hodnocení stavu není velmi často dostatek statistických údajů, zato však je mnoho subjektivních ocenění, která mnohdy bývají použita přinejmenším jako výchozí základna.

Výchozím bodem Bayesových stoupenců je proto často sázka. I když o zítřejším počasí vůbec nic nevím, nemám revma v noze ani barometr, ani jsem neslyšel předpověď počasí, jsem však pravděpodobně ochoten — snad na základě včerejšího a dnešního počasí — uzavřít sázku 2 : 1, že zítra nebude pršet. Tím



„Vzorek“ je vlastně slovo převzaté ze staré řeči hutníků. Byl odebírán z tavicích pecí, aby se získal přehled o stavu tavby.

je dána apriorní pravděpodobnost, která může být nyní dále korigována. Dostali bychom se do hlubin filozofie statistiky a teorie pravděpodobnosti, kdybychom chtěli podrobně popisovat boj mezi Bayesovými stoupenci a zastánci teorie četností pravděpodobnosti. Omezíme se na dvě poznámky. Především je třeba na Bayesově větě ocenit dynamiku, která tvoří její součást — je to do jisté míry vzor regulérní technické pomůcky. Za druhé, subjektivní pravděpodobnosti a četnosti nejsou vždy ve vzájemném rozporu. Subjektivní pravděpodobnost představuje většinou spíše náhražku za chybějící informace o „skutečných“ četnostech. Pro teoretické rozhodovací modely je konečně subjektivní pravděpodobnost vůbec jedině rozhodující. Wilhelm *Krelle* např. říká: „Teorie rozhodování je proto od samého začátku... »bayesovská«, to znamená, že pracuje se subjektivními pravděpodobnostmi, totiž těmi stupni jistoty osoby, která rozhoduje, že jiné osoby učiní určité opatření nebo že nastanou jiné nejisté události.“

Tato subjektivní pravděpodobnost se ovšem v případě racionálně se rozhodujícího člověka co nejvíce přizpůsobí statistické pravděpodobnosti, pokud taková pravděpodobnost existuje.

A těmito objektivními metodami získávání informací na základě malých a nejmenších vzorků se budeme nyní zabývat.

6.4 Nejmenší vzorky

V nejlepší příručce z přelomu století, ve dvacetisvazkovém naučném slovníku, kterému se v mnohém ohledu od té doby sotva co vyrovná, budeme ještě marně hledat pojem „vzorek“. Jeho hutnický význam již zastaral („kdykoli tavič odpichuje tavbu, vždycky bere vzorek,“ říkalo se v jedné hutnické kronice ze 16. století o způsobu tavby ve vysokých pecích), a jeho nový, statistický význam ještě nepronikl mimo nejužší okruh vědců-odborníků.

Obsah pojmu, i když v poněkud jiném smyslu, existoval již ovšem déle. Už Thomas Bayes ve svém počtu pravděpodobnosti bral zřetel na jednotlivé

nové poznatky. Ještě dříve museli první političtí aritmetici pracovat s tak nedostačujícími údaji, že jejich práce, ačkoliv se zakládala na vyčerpávajícím zjišťování, měla jen vypovídací schopnost vzorku, mnohdy dokonce ani ne náhodného vzorku v dnešním slova smyslu; náhoda spočívala pouze v tom, zda zrovna získali nějaké údaje.

Později byly dílčí výsledky bezstarostně převedeny na výsledky celkové — Quételet vypočítal váhu všech obyvatel Bruselu prostě tak, že jednotlivé váhy příslušně násobil. Matematická teorie náhodných výběrů naproti tomu vznikla teprve ve dvacátých a třicátých letech našeho století a pak rychle nastoupila vítězný pochod světem. Vyčerpávající zjišťování existují ovšem i dnes, ale dynamika statistické teorie a praxe spočívá téměř výlučně na malých číslech, na vzorcích.

Tento vývoj začal dříve — a především se dříve prosadil — v anglo-americké oblasti, ve Francii a v Itálii než v německé jazykové oblasti, kde ještě hluboko do 20. století převládal názor Georga von Mayra, že každá řádná statistika musí být vyčerpávajícím zjišťováním. Dílčí a výběrová zjišťování pokládala tato škola pouze za nouzovou pomůcku. Ještě v polovině třicátých let si stěžoval Oskar Anderson, že pod vlivem tohoto směru „*je v Německu statistická teorie považována za vědu o státě, která se má převážně zabývat hromadnými jevy v lidské společnosti*“ a že „německý odborník vůbec nemůže rozumět např. anglickému teoretikovi a jeho odborným časopisům“. Tím se také vysvětluje, proč v našem výkladu moderní statistiky jsou neustále představováni Angličané nebo Američané. Rovněž *teorie malých čísel* vznikla téměř bez výjimky v anglo-americké oblasti.

Jak malá však mohou být tato čísla, rozsahy vzorků, aby mohla vypovídat něco užitečného? Prováděli jsme již propočty o potřebném rozsahu výběrových souborů a přitom jsme viděli, že optimální rozsah náhodného výběru je možno vypočítat *podle rozptylu* (resp. *četnosti znaků*) a v závislosti na *požadované přesnosti* (*úrovni významnosti*). Alespoň poněkud spolehlivé výsledky předpokládají výběrové soubory o rozsahu 500 až 3000.

Jak často však nelze — zejména ve vědeckém výzkumu — při nejlepší vůli zjišťovat tisíce hodnot? Máme např. testovat nový preparát proti třesavce — kolik pacientů, kteří přicházejí v úvahu, máme vůbec na klinice? Z nich však musí být část léčena jako kontrolní skupina dosavadní metodou! (Argument, že lék slibující úspěch se musí použít bez výjimky pro všechny pacienty, svědčí sice o vysoké etice, ale vede ke špatné statistice. Bez porovnání není nulové hypotézy ani ověřování hypotéz.) V jiných případech je potřebný homogenní materiál k dispozici jen v omezeném rozsahu. Teoreticky dosažitelné údaje a zjišťování ze zahraničí lze použít jen s velkou zdrženlivostí — zkrátka, většinou to už bývá tak, že z nějakého důvodu mohou být vytvořeny jen velmi malé vzorky. Jako statistický materiál je k dispozici 20, 15 a ještě méně případů. Je vůbec možné něco s tím počít?

6.41 Studentovo rozdělení t

William Sealy Gosset, který pracoval začátkem našeho století jako chemik v anglickém pivovaru Guinness, vymyslel postup, který měl poskytnout možnost dělat z malých a nejmenších vzorků použitelné závěry, přinejmenším však poznat, jak posuzovat vypovídací hodnotu takových vzorků. Jméno Gosset je už dnes téměř neznámé, avšak *každému statistikovi je zcela běžné „Studentovo t" a „rozdělení t" podle Studenta.*

Gosset se totiž pod svá průkopnická díla podepisoval pseudonymem *„Student"*, protože mu jeho firma nedovolila publikovat vědecké práce pod vlastním jménem. Snad by ani nesměl formou „vedlejšího zaměstnání" vůbec psát, a proto se úzkostlivě ukrýval pod pseudonymem. Výsledky jeho bádání, především rozdělení t, částečně doplněné a přepracované R. A. Fisherem, patří dnes k bezpodmínečné výzbroji moderní teorie výběrových souborů.

R. A. Fisher sám vtiskl statistice první poloviny tohoto století svoji pečeť. Alexander McFarlane Mood o něm v roce 1950 napsal: *„Fisher je obrem ve vývoji teoretické statistiky.* Jeho první příspěvek vyšel v roce 1912 a od té doby až dodnes neustále publikuje... Tomuto jedinému muži je nutno připsat nejméně polovinu celého důležitého zásadního vývoje, vycházíme-li z dnešního stavu teorie." A skutečně, téměř všechny moderní matematické metody statistické analýzy pocházejí od R. A. Fishera.

Již jsme uvedli, že zejména Bayesovi stoupenci uplatňují námitky proti čistě matematickým metodám, což platí zvláště o teorii nejmenších vzorků, *„small sample theory" — a rozdělení t může být omezeno na vzorek sestávající pouze ze dvou (!) hodnot!* Proti těmto metodám se ostře postavil zejména Richard von Mises, jeden z nejvýznamnějších teoretiků počtu pravděpodobnosti, který v roce 1950 napsal: „Tento vývoj vedl obdobím, kdy byla sledována pochybná myšlenka — z malého sledu pozorování dělat statistické závěry, tzv. »small sample theory«. Zdá se, že nyní jsou již tyto omyly v podstatě překonány..."

Nuže, „small sample theory" není zatím vůbec překonána, *Studentovo rozdělení t* má stále ještě vynikající úlohu, a proto tuto důležitou metodu alespoň krátce popíšeme. Náš popis nemůže ovšem zdaleka zachytit celou problematiku a už vůbec ne celou techniku použití tohoto postupu, neboť nejde sice o „vyšší", ale přece jen o — řekněme — „střední" statistickou metodu, k jejímuž ovládnutí je třeba více matematiky a teorie pravděpodobnosti, než si zde můžeme dovolit uvést. Nicméně vyjádření základní myšlenky by se nám mělo podařit.

Když jsme v páté kapitole mluvili o směrodatné chybě průměru a četnosti, upozorňovali jsme již na to, že vzorec vyžaduje σ, zatímco ve skutečnosti je k dispozici pouze s vzorku. Tehdy nebylo toto rozlišování ještě opravdu důležité, protože výběrový soubor byl stále dostatečně velký, aby se z něho dal přibližně přesně určit rozptyl. Tento předpoklad však neplatí, máme-li velmi malé vzorky. Máme-li jen tři měrné hodnoty (např. 18, 14, 13), mohli bychom opravdu seriózně vypočítat standardní odchylku? Podle již vyzkoušeného schématu? ($\bar{x} = 15$; $3^2 = 9$; $1^2 = 1$; $2^2 = 4$; $s = \sqrt{\frac{14}{3}} = 1{,}67$.)

Dokonce i když nedělíme *n*, nýbrž $n - 1$, jak je účelné i při větším rozsahu výběrového souboru, na což jsme již krátce upozornili v oddílu 2.5, je směrodatná odchylka $s = \sqrt{\frac{14}{2}} = 2{,}65$ pochybné ceny. K tomu však přistupuje ještě to, že nejen velikost rozptylu, ale i průměr vzorku je při velmi malém výběrovém souboru téměř bezcenný. Mohlo by se lehce stát, že rozdělení souboru vůbec nemá rozptyl $\sigma = 1{,}67$ (nebo také 2,65) kolem $\mu = 15$, nýbrž že by skutečný průměr mohl docela dobře být 12 nebo 7, skutečná směrodatná odchylka 1 nebo také 3.

Krátce řečeno, u vzorku o rozsahu 15 nebo ještě méně nemá smysl mluvit o „normálním rozdělení" a s tímto pojmem vážně pracovat. Ale co má přijít místo něho? *Na jeho místo přichází srovnání s tabulkou, která byla propočtena speciálně pro tyto případy, s tabulkou pocházející od Studenta.*

Předpokládejme hypotézu, že v určitém souboru skutečná střední hodnota činí přesně 100. Z tohoto základního souboru jsme vytvořili vzorek o rozsahu $n = 5$, abychom viděli, zda na tomto základě může být hypotéza zamítnuta. Pět potřebných měření dává tyto hodnoty: 82, 87, 104, 92, 90. Lze ještě uvést tento výsledek do souladu s danou hypotézou? Nejdříve vypočítáme $\bar{x}$, což je 91. Pak vypočteme známým postupem rozptyl: součet druhých mocnin odchylek od průměru výběrového souboru je $81 + 16 + 169 + 1 + 1 = 268$.

Rozptyl $s^2 = 53{,}6$, směrodatná odchylka výběrového souboru $s = \sqrt{53{,}6} = 7{,}32$. Tyto hodnoty dosadíme do vzorce:

$$\boxed{t = \frac{|\mu - \bar{x}|\sqrt{n}}{\sigma}}$$

Mohou překvapit svislé čárky, mezi nimiž je výraz $\mu - \bar{x}$; znamenají pouze *„absolutní hodnota"* a vyjadřují, že se nemusí dbát na znaménko. Výraz

„absolutní hodnota" jsme již ve stručnosti poznali (srv. oddíl 2.5).

Důležitější je něco jiného: *t*, jež nyní dostaneme, není totiž výsledek, který by třeba jako v případě normálního rozdělení bylo možno najít v tabulce a jehož přesnost výpovědi by se mohla přímo určit. Je to spíše testovací veličina, *„zkušební proměnná"*, jak jsme ji poprvé předvedli v příkladu Čápa a spol.

Nyní dosadíme naše hodnoty, přičemž nejdříve za σ prostě použijeme nahoře vypočítané s, které se ovšem ukáže jako příliš malé.

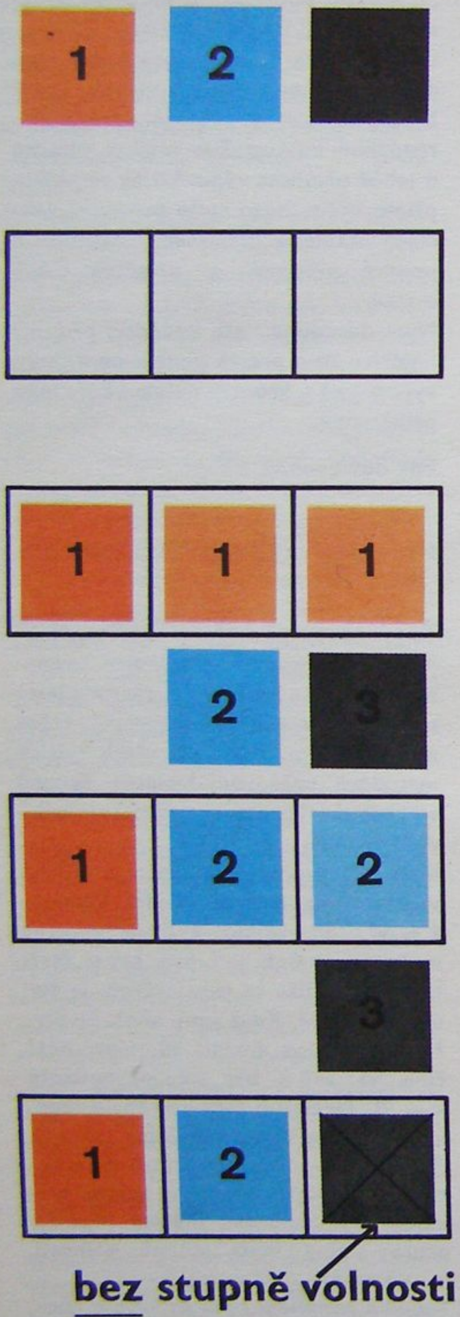
Pak dostaneme:

$$t = \frac{100 - 91}{7{,}32}\sqrt{5} = (1{,}23)\,(2{,}24) = 2{,}75.$$

To je naše první zkušební veličina. Nežli ji však porovnáme s hodnotou uvedenou v tabulce rozdělení *t*, musíme ještě stanovit tzv. *„stupně volnosti"*, které mají důležitou úlohu při všech vyšších metodách ověřování hypotéz. Stupně volnosti měří prostor (volnost) výsledků výběrů.

Myšlenkový postup je tento: z našeho vzorku jsme dostali střední hodnotu $\bar{x} = 91$. Jak vznikla? Existují nesčetné možnosti, avšak jen pro první čtyři měření: jestliže ta jsou určena, je tím určeno i páté. Když jsme např. ve čtyřech měřeních dostali 90, musí páté činit 95, má-li být střední hodnota $\bar{x} = 91$. Nemají-li tedy všechny *n* vzorky úplnou volnost, nýbrž jen $(n - 1)$, činí také číslo stupňů volnosti ve výběrovém souboru $n - 1$. V našem příkladu jsme měli pět měření ($n = 5$), proto máme čtyři stupně volnosti ($n - 1 = 4$).

Tabulka rozdělení *t* má sloupce a řád-



ky. Ve sloupcích je uvedena pravděpodobnost v procentech, v řádcích pak počet stupňů volnosti, na nichž se zakládá testovaný propočet. V průsečíku příslušného sloupku a řádku je srovnávací hodnota. Předpokládejme, že jsme chtěli 95% jistotu: vyhledáváme sloupec 95 % řádek 4 a tam najdeme: 2,13. Proti této tabulkové hodnotě je naše zkušební veličina ($t = 2{,}75$) poněkud větší, hypotéza ($\mu = 100$) je tedy na této úrovni zamítnuta. Naproti tomu u 97,5 % je tabulková hodnota 2,78 a na této úrovni se již hypotéza nezamítá. Náš výsledek je tedy mimo 95 % velikosti náhodného rozdělení kolem průměru $\mu = 100$, ale uvnitř 97,5 % velikosti rozptylu. Zda je to „významné", je ponecháno našemu úsudku.
Stupně volnosti pro rozdělení t se většinou udávají jen do 30 (nebo 50), protože nad tuto výši se již uplatňuje „bezmezná" volnost normálního rozdělení. (Mezní hodnota rozdělení t je totožná se směrodatnou odchylkou normálního rozdělení. Přitom se však musí vycházet z tabulek pro „jednostranné" pravděpodobnosti, protože i rozdělení t je jednostranné.)
Souvislosti mezi normálním rozdělením, směrodatnou chybou průměru a Studentovým t jsou pozoruhodné. Pro standardizaci normálního rozdělení jsme použili vzorce (srovnej oddíl 3.32)

Tři měrné hodnoty (nebo čísla, hrací kameny atd.) se rozdělí libovolně na tři volná místa, přičemž místo pro první kámen lze ještě vybírat. Také pro druhý je ještě nějaká „volnost" volby. Místo pro třetí je však již pevně určeno provedenými volnými rozhodnutími. Číslo „stupně volnosti" je tedy o jednu menší než počet míst, která jsou k dispozici.

..	1) 15	15	15	15
..	..	2) 22	22	22
..	..	..	3) 18	18
..	..	..	..	4) 21
..	..	..	..	24
100	100	100	100	100

Mám-li z pěti libovolných hodnot dostat celkový součet 100, jsou k dispozici jen 4 stupně volnosti: první až čtvrté číslo mohu volit libovolně, páté však už je nutně určeno. Jestliže jsem zvolil 15, 22, 18 a 21, nezbývá mi již žádný stupeň volnosti: páté číslo nemůže být žádné jiné než 24.

$$\text{normovaná hodnota} = \frac{\text{daná hodnota minus aritmet. průměr}}{\text{směrodatná odchylka}}$$

nebo, vyjádřeno v symbolech,

$$z = \frac{x_i - \mu}{\sigma}.$$

U Studentova t známe již vzorec

$$t = \frac{|\mu - \bar{x}|}{\sigma}\sqrt{n},$$ který lze napsat také takto: $$\frac{|\mu - \bar{x}|}{\frac{\sigma}{\sqrt{n}}}.$$

Hodnotu $\sigma/\sqrt{n}$ však známe (z oddílu 5.4) jako „*směrodatnou chybu průměru*". Můžeme tedy definovat:

$$\text{zkušební veličina } t = \frac{\text{hypotetický průměr souboru — průměr výběrového souboru}}{\text{směrodatná chyba průměru}}$$

Potíž spočívá v tom, že konstantu σ hypotetického základního souboru můžeme odhadovat jen z náhodné proměnné s vzorku. Avšak právě u malých a nejmenších vzorků může být rozptyl výběrového souboru velmi nespoleh-

	Rozdělení t			
n \ α	95 %	97,5 %	99 %	99,5 %
1	6,31	12,71	31,82	63,66
2	2,92	4,30	6,96	9,92
3	2,35	3,18	4,54	5,84
4	2,13	2,78	3,75	4,60
5	2,01	2,57	3,36	4,03
6	1,94	2,45	3,14	3,71
7	1,89	2,36	3,00	3,50
8	1,86	2,31	2,90	3,36
9	1,83	2,26	2,82	3,25
10	1,81	2,23	2,76	3,17
11	1,80	2,20	2,72	3,11
12	1,78	2,18	2,68	3,05
13	1,77	2,16	2,65	3,01
14	1,76	2,14	2,62	2,98
15	1,75	2,13	2,60	2,95
16	1,75	2,12	2,58	2,92
17	1,74	2,11	2,57	2,90
18	1,73	2,10	2,55	2,88
19	1,73	2,09	2,54	2,86
20	1,72	2,09	2,53	2,85
21	1,72	2,08	2,52	2,83
22	1,72	2,07	2,51	2,82
23	1,71	2,07	2,50	2,81
24	1,71	2,06	2,49	2,80
25	1,71	2,06	2,48	2,79
26	1,71	2,06	2,48	2,78
27	1,70	2,05	2,47	2,77
28	1,70	2,05	2,47	2,76
29	1,70	2,05	2,46	2,76
30	1,70	2,04	2,46	2,75
34	1,69	2,03	2,44	2,73
40	1,68	2,02	2,42	2,70
50	1,68	2,01	2,40	2,68
70	1,67	1,99	2,38	2,65
100	1,66	1,98	2,36	2,63

Tabulka srovnatelných hodnot rozdělení t. V levém sloupci je číslo stupňů volnosti n.

livý. Bude totiž menší než rozptyl v základním souboru a tento rozdíl bude zvlášť rušit u nejmenších vzorků.
Jsou dvě možnosti, jak korigovat příliš malé hodnoty vzorků, přičemž obě vedou ke stejnému výsledku. Ukážeme je na našem příkladu, protože se tím také vyjasní někdy rozdílné definice s. Již v druhém oddílu jsme upozornili, že směrodatnou odchylku výběrového souboru lze matematicky správně získat tím, že součet umocněných odchylek — $\Sigma (x_i - \bar{x})^2$ — se nedělí počtem měření (n), nýbrž jen ($n - 1$).
Měli jsme tedy součet odchylek 268, dělený ($n - 1 = 4$), což dává $s^2 = 67$ a $s = \sqrt{67} = 8{,}18$.
Dosaďme tuto hodnotu do našeho vzorce místo neznámého σ, čímž dostaneme správnější zkušební hodnotu (která již přihlíží k nepatrnému rozptylu výběrového souboru):

$$t = \frac{(100 - 91)\sqrt{5}}{8{,}18} = \frac{20{,}16}{8{,}18} = 2{,}46.$$

Můžeme však použít také jiného způsobu a říci: odhadneme neznámou σ tak, že nejdříve vypočítáme rozptyl výběrového souboru, jako kdyby šlo o rozptyl základního souboru (dělíme tedy $n = 5$), pak však použijeme opravného koeficientu. Tento koeficient zní $\sqrt{\frac{n}{n-1}}$, v našem příkladu tedy $\sqrt{5/4} = \sqrt{1{,}25} = 1{,}12$. Náš původní výpočet dal příliš malý rozptyl $s = 7{,}32$. Násobíme jej 1,12 a dostaneme jako nahoře 8,18. A tím opět $t = 2{,}46$.
Moderní učebnice statistiky téměř bez výjimky používají definice s^2 a s, která již zvýrazňuje charakter výběrového souboru, to znamená s dělením ($n - 1$). Ve starší literatuře se však můžeme setkat s tím, že se rozptyl výběrového souboru počítá stejně jako rozptyl základního souboru — a pak je nutno u malých výběrů použít „opravného koeficientu", jehož oprávněnost ještě dále prokážeme.
V této knize jsme dříve zacházeli s rozptylem vzorku poněkud „naivně", ale nyní, když se zabýváme malými a nejmenšími vzorky a kromě toho se stále více věnujeme matematickým testům hypotéz, musíme s^2 a s počítat přesně.
Poznamenejme si tedy: naše zkušební veličina je nyní $t = 2{,}46$. Zkušební veličina se nachází „pohodlně" uvnitř 97,5 % náhodného rozptylu — žádný div, vždyť jsme poněkud zvětšili korekturu rozptylu.
Skoro to vypadá tak, jako by se čitatel zkušební veličiny při rostoucím n stále zvyšoval, avšak to nesouhlasí, protože při rostoucím rozsahu vzorku směřuje i rozdíl ($\mu - \bar{x}$) k nule, neboť vzorek nevychází z hypotetického základního souboru, což má být mimochodem také testem dokázáno.
V mezním případě výběrového souboru jen se dvěma měřeními (tedy minimem pro tvorbu střední hodnoty) je stále možno použít rozdělení t, které je pak ovšem krajně „velkorysé", protože v takovém případě se jeví jako možné téměř všechno. Předpokládejme, že bychom z výběrového souboru o rozsahu $n = 2$ dostali 2 výsledky: 4 a 6, tedy $\bar{x} = 5$. Nyní bychom rádi věděli, zda se to snáší s hypotézou $\mu = 12$, což se na první pohled zdá nesmyslné. Avšak dostáváme

$$t = \frac{(12 - 5)\sqrt{2}}{\sqrt{2}} = 7.$$

V tabulce pro stupeň volnosti nacházíme však při 95 % již 6,31, při 97,5 %

dokonce 12,71 — naše hypotéza je tedy těsně zamítnuta jen na úrovni 95 %, ale už vůbec ne na úrovni 97,5 %.
Kdo již listoval ve statistické literatuře, narazil možná na *vzorce*, které se od zde uváděných *poněkud liší*.
Hlavní důvod spočívá v tom, že — jak jsme již několikrát poznamenali — *při používání symbolů a také v terminologii není vždy úplný soulad* mezi autory různých zemí a dob, nemluvě už o rozdílné matematické úrovni čtenářů.
Mimoto neznamená n v tabulkách rozdělení t zpravidla rozsah vzorku n, nýbrž číslo stupňů volnosti. Dále je pravděpodobnost někdy „zrcadlovým obrazem": 1 % znamená 99 %; co je tedy zamítnuto na úrovni 95 %, má pravděpodobnost více než 0,05. Ještě horší však je, že stejně jako u normálního rozdělení je třeba pečlivě zkoumat, zda tabulka platí pro jednostranné nebo dvoustranné testy — zpravidla jen pro jednostranné s odpovídající „koncentrací" pravděpodobnosti na jednom konci křivky rozdělení t (tato křivka totiž přece jen existuje, ale má pro každý stupeň volnosti poněkud jinou podobu, až nakonec přechází do normální křivky).
Na tomto místě bylo důležité jen to, aby se podařilo načrtnout zásadní význam rozdělení t a testů t. Kdo se tím chce vážně zabývat, udělá dobře, bude-li ještě studovat odbornou literaturu.
Studentův test je užitečný především tehdy, když se má zkoumat, zda dva vzorky pocházejí ze stejného základního souboru. U jednoho výběrového souboru vychází např. $\bar{x}_1 = 94$, u druhého $\bar{x}_2 = 103$. Nyní je třeba ověřit, jaká je pravděpodobnost, že oba pocházejí ze základního souboru se skutečným průměrem $\mu = 99$. Máme zde jakýsi druh hry na ověřování hypotéz (nulová hypotéza: „Mezi oběma vzorky není žádný významný rozdíl"), a to při použití velmi malých vzorků, jako je tomu vždy u rozdělení t.
Takové testy jsou velmi poučné zvláště tehdy, když podmínky pokusu u obou vzorků byly rozdílné. Erwin Kreyszig („*Statistické metody a jejich používání*") k tomu uvádí příklad z anglického odborného časopisu. Při ručním třídění fazolí („Popelčina metoda") byly vyzkoušeny dva způsoby. Při jednom se vadné fazole vkládají do dlaně a odloží se, až je dlaň plná. Při druhém se špatná fazole vezme a hned odhazuje. Devět dělníků zkoušelo obě metody s těmito výsledky:

Dělník č.	První metoda	Druhá metoda
1	214	281
2	253	279
3	276	260
4	215	230
5	238	267
6	221	253
7	210	205
8	229	265
9	269	299

Podmínkou pro platnost testu při tom ovšem je, že při obou pokusných řadách jsou všechny fazole stejného druhu.
Nulová hypotéza zní: „Oba způsoby jsou stejně dobré." Může být vyvrácena? První výběr má průměr $\bar{x}_1 = 260$ a rozptyl $s_1^2 = 796{,}4$, druhý $\bar{x}_2 = 236$ a $s_2^2 = 607{,}1$.
Zkušební veličina t se nyní vypočítává podle vzorce:

$$\boxed{t = \frac{|\bar{x}_1 - \bar{x}_2|\sqrt{n}}{\sqrt{s_1^2 - s_2^2}}}$$

Dosadíme a dostaneme:

$$t = \frac{(260 - 236)\sqrt{9}}{\sqrt{796 + 607}} = \frac{3,24}{\sqrt{1403}} =$$

$$= \frac{72}{37,4} = 1,91.$$

V tabulce najdeme (při úrovni 95 % v řádku dvakrát osm = 16 stupňů volnosti) 1,75.

Naše zkušební hodnota je vyšší, nulová hypotéza se tedy zamítá, způsoby zřejmě nejsou stejně dobré. (Pravděpodobnost, že jsou přece jen stejně dobré, je menší než 5 %.) Předpokládejme proto jako alternativní hypotézu, že druhý způsob výběru je lepší.

7 Proč statistiky lžou?

7.1 „Statistiky", které nejsou statistikami

Jestliže je statistika (jako metodika nebo jako vědní obor) často posuzována s pochybnostmi a odmítavě, může se za to poděkovat především statistikám, které ve skutečnosti statistikami nejsou. Je to stejné jako kdyby nemocného člověka léčil mastičkář, zřízenec nebo kuchyňský personál kliniky a nemocný pak mrzutě konstatoval: „Medicína není vůbec žádná věda; všichni lékaři jsou šarlatáni."

Zákon však dbá o to, aby jako lékař působil jen ten, kdo dokončil příslušné studium, vykonal zkoušky a získal praktické zkušenosti. O statistice se naproti tomu bezmála každý domnívá, že zde může působit bez zvláštních předchozích znalostí kdokoliv — od ctižádostivého vedoucího prodeje až po šéfa s geniálními myšlenkami, intrikán a demagog stejně jako naivní pošetilec a dobromyslný ignorant.

Již několikrát jsme měli příležitost ukázat, že stejnou věc je možné pozorovat z různých hledisek a podle toho statisticky různě vyjádřit. Vždycky jsme však přitom mlčky předpokládali, že — za prvé — číselný materiál, který je připraven k použití, je alespoň do určité míry správný a že — za druhé — se nelže vědomě.

Jestliže se tedy mluví o lži ve statistice, je nutno vždy zjistit, o jaký druh lži jde. Existuje především *zdánlivá lež*, která není v podstatě nic jiného než nesprávně pojatá přesná statistika (je zcela možné, že je lstivě zaměřena na oklamání naivních lidí, ale sama o sobě svými údaji a tvrzeními je nenapadnutelná). Dále existuje *odvozená, sekundární lež*, charakterizovaná tím, že se statistik zaměňuje s komentátorem. Je to zvlášť rozšířený druh a vzniká většinou tak, že nějaký tajemný manipulátor, demagog nebo podvodník se seznámí — nebo je seznámen — s čísly, která vhodně vyložena se zdají dokazovat jeho myšlenkový koncept. Zmocní se tedy v podstatě správných čísel a buduje kolem nich konstrukce lži, která je nespornými čísly znamenitě udržována a posilována. To se stává především tehdy, když se k číslům vyhledávají „příčiny" nebo „následky". Až budeme později mluvit o korelacích a zdánlivých korelacích, bude ještě možné k tomu ledacos vážného i veselého dodat.

A konečně existuje *úplně nemotorná forma lži*, při níž však lze postupovat statisticky korektně jak při zpracování, tak při výkladu. Bývá tomu tak tehdy, jestliže se hned od začátku zfalšuje prvotní materiál. Můžeme si zde uvést starou anekdotu o dvou cikánech, v které se jeden velice diví tomu, že ten druhý může prodávat koše pletené z vrbového proutí ještě laciněji než on sám. „Jak je to možné?" ptá se první. „Tak lacino nemohu prodávat, ačkoliv vrbové proutí kradu." Druhý se smíchem pokrčí rameny: „To já kradu celé koše."

Jestliže se tedy najde vhodný číselný materiál, opravdu není zapotřebí namáhat se s vymýšlením rafinovaných statistických operací nebo s obtížnými vychytralými výklady. Ernst Wagemann, v letech 1923 až 1933 prezident Říšského statistického úřadu a autor četných populárně vědeckých knih, uvádí ve svém díle „*Číslo jako detektiv*" několik příkladů ze světových dějin, které — někdy po tisíciletí — vedly k nesprávným výkladům a závěrům: Xerxovo vojsko mělo podle Hérodota 4,2 miliónu mužů — ve skutečnosti však snad jen 20 tisíc, s vozatajstvem 60 až 70 tisíc. Švýcaři odhadovali ztráty Karla Smělého u Gransonu na 7000 mužů — ve skutečnosti padlo jen 7 rytířů a několik vojáků. V německých osvobozovacích válkách prý bylo v bitvě u Hagelsbergu pobito 4000 Francouzů — ve skutečnosti jich však padlo pouze 30. Použitím nesprávných výchozích údajů — nebo dokonce vědomým falšováním „prvotního záznamu" — je možno dokázat všechno. I nekritickému čtenáři nebo posluchači bude těžké namluvit, že 4 a 5 je 6, ale jestliže nejdříve zfalšuji 5 na 2, mohu potom plným právem tvrdit, že 2 a 4 je 6.

Proto se vždy doporučuje nejdříve zjistit, zda se setkáváme s vrátným, ošetřovatelkou nebo šéflékařem, zda s pseudostatistikou nebo se statistikou. Rozlišení není na první pohled vždy snadné. I vrátný může v bílém plášti a v brýlích prohodit několik latinských slov a laikovi předstírat lékaře. Ještě snadněji se podaří mladému muži v osobním oddělení vypočítat na počítači procenta s přesností až na desetinná místa, a tak vyrobit zdánlivě pravou „statistiku".

Jak se ubránit vylhaným sloupcům čísel a fantastickým křivkám, jak je poznat? Kdo dočetl tuto knihu až sem, je už značně imunní. Kdo ji dočte až do konce, může se pak s klidným svědomím setkat se správnými i nesprávnými statistikami. Existuje však několik docela prostých zásadních pravidel, která ovšem — jako ostatně každé pravidlo — zaručují jen relativně malou pravděpodobnost jistoty. Vždy je ještě možno např. statistiky a výsledky anketních šetření, shrnuté v pěti (nebo i v deseti) řádcích noticky v denním tisku, označit za „ne zcela důvěryhodné". V nejlepším případě jde o původní číselný materiál, který byl někým interpretován a v této podobě předán tisku. Potom nějaký novinář nikoli ve zlém úmyslu, nýbrž aby nerozzlobil čtenáře bojícího se čísel, část údajů vynechá a z komentovaného textu zdůrazní to, co působí aspoň trochu senzačně.

W. J. Reichmann komentuje např. zprávu vytištěnou tučným písmem: „Každá druhá žena si stěžuje na bolesti v zádech", a uvádí pak, že již původní statistika obsahovala několik slabin. Předně nešlo o základní soubor „ženy" (mimo jiné by bylo třeba vyjasnit, zda například šestnáctileté dívky mají být zahrnuty nebo nikoliv), nýbrž o pacientky. Ženy, které navštíví lékaře, jsou ovšem bezpochyby v průměru méně zdravé a trpí více bolestmi než ty, které nejsou v čekárnách. Tedy jen: „Každá druhá pacientka si stěžuje na bolesti v zádech."

Dále se ukázalo, že tento výsledek nebyl získán z reprezentativního anketního šetření mezi praktickými lékaři, nýbrž byl výsledkem soukromé statistiky jediného lékaře. Tedy jen: „Každá druhá pacientka doktora X si stěžuje na bolesti v zádech." Reichmann k tomu již zlomyslně poznamenává, že dotyčný lékař provozuje svou praxi buď ve velmi vlhké krajině, nebo má v čekárně dost nepohodlné židle.

Ale to zdaleka není všechno. Každá

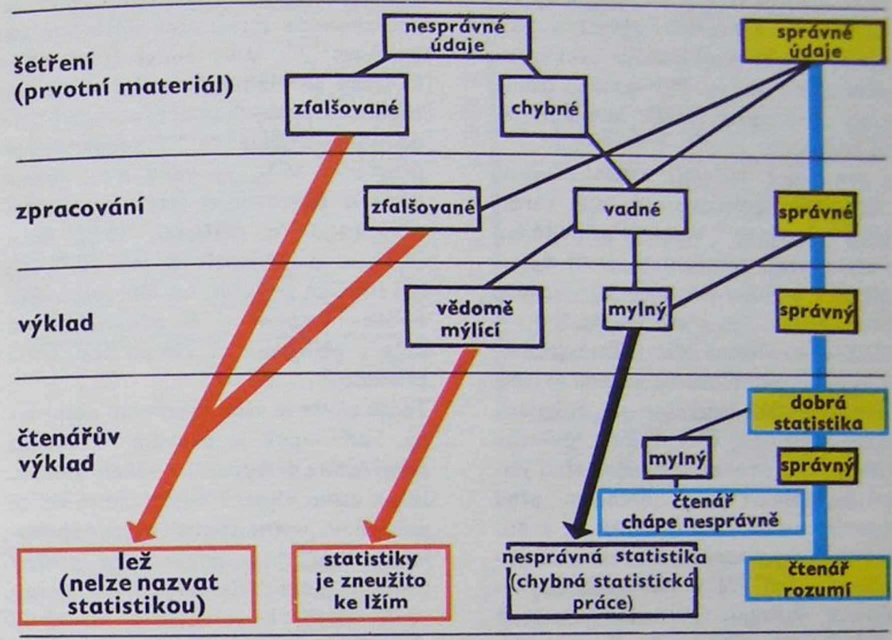


Schéma zdrojů chyb a možností klamání v průběhu sestavování a zpracovávání statistik.

druhá pacientka doktora X si na bolesti v zádech sama od sebe nestěžovala, nýbrž odpověděla „ano" na otázku, zda trpí bolestmi v zádech. Nu a koho někdy nebolí v zádech? Nebo hlava? Nebo kdo čas od času není nervózní? To všechno jsou tak rozšířené symptomy, že diagnosticky mají určitou cenu teprve tehdy, jsou-li intenzívní a pacienti je uvádějí jako značné obtíže. Ten, kdo teprve na otázku „Máte někdy bolesti v zádech" odpoví „ano", si na to vlastně nestěžuje a není tím „významně" omezen ve svém zdravotním stavu. *Tak se scvrkává „statisticky podložené" tvrzení*, podle něhož si každá druhá žena stěžuje na bolesti v zádech, na mnohem méně působivou skutečnost, že někde v Anglii je nějaký lékař, polovina jehož pacientek na otázku, zda také mají bolesti v zádech, odpovídá „ano".

Tuto zprávu doplňovalo sdělení, že též na jedné ortopedické klinice si třetina pacientů stěžovala na bolesti v zádech, což Reichmann komentoval protiotázkou, zda je to na takové klinice překvapující, a skončil kousavým vtipem, že jednoho dne se s hlubokým dojetím zjistí, že v tuberkulózních sanatoriích trpí 100 % pacientů tuberkulózou.

V tomto případě bylo alespoň možné vystopovat na základě původní zprávy všechny zdroje chyb, ale co má dělat čtenář, kterému se předkládá pod uvedeným titulkem hustý text, než domnívat se, že opravdu každá druhá žena v Anglii si stěžuje na bolesti v zádech?

Ještě horší je to s novinářským vyhodnocováním anketních šetření a „statistik“, které mají pomoci překlenout okurkovou sezónu. Pak vznikají titulky typu „Průměrný člověk líbá v životě 56 807krát“.

Jeden ještě tučnější titulek hlásal: „Statistika dokazuje: šílenství chrání před rakovinou“. V tomto případě bohužel chybějí věcné údaje téměř úplně, ale již z prvního odstavce, upraveného jako „trhák“, je přece jen možno vytušit, o co vlastně jde: „Demografové a lékaři kroutí hlavou a neznají vysvětlení. Nejnižší úmrtnost na rakovinu mají chovanci ústavu pro duševně choré.“ Už z toho je přinejmenším jasné, že šílenství vůbec „nechrání“ před úmrtím na rakovinu, nýbrž že z duševně chorých umírá na rakovinu poměrně méně než z veškerého obyvatelstva. Zjištění, které si nepochybně zaslouží přesnější ověření hypotézy (pokud v této formě opravdu vznikla), však vůbec nepodává „statistický důkaz“ o tom, že duševně choří neumírají na rakovinu.

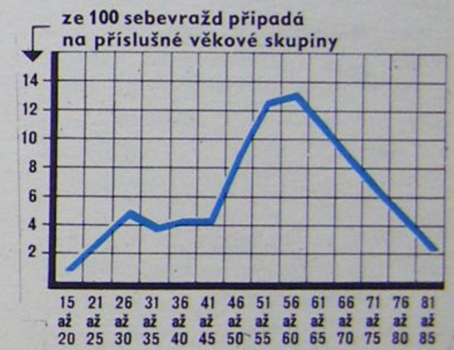
Z popisu „průměrného Američana“ je možno se mimo jiné dozvědět toto: „Průměrný Američan je 1,75 m vysoký, váží 71,5 kg, navštěvuje všechny zápasy baseballu a košíkové a v hazardních hrách víc prohrává než vyhrává.“ Váha a výška jsou zřejmě hodnoty mediánu (viz str. 38). Poznámka o hazardních hrách je otřepanou pravdou, že všechny hazardní hry se zakládají na tom, aby přinášely provozovateli zisk, a proto může hráč „v průměru“ jen prohrávat, stejně jako průměrný Středoevropan prohrává v loterii a sázkách více, než vyhrává. Co znamená „*všechny* baseballové zápasy“? Nelze se pak divit, že inteligentní čtenář na „statistiku“ nadává...

Horší je naproti tomu text, který se zdůrazněnou serióznosti následuje za titulkem „Ve stáří budeš šťastnější“. Doslova se říká: „Policejní statistika západoevropských zemí překvapivě jednomyslně zjišťuje: ze 100 sebevrahů je průměrně 45 % ve věku mezi sedmnácti a pětadvaceti lety, 30 % mezi pětadvaceti a čtyřiceti, 15 % mezi čtyřiceti a padesáti a jen 10 % (!) lidí starších padesáti let má dost svého života. Rozhodnutí k *sebevraždě* má tedy s přibývajícím věkem sestupnou tendenci.“

Tento závěr je však absolutně nesprávný. Spíše opak je pravdou. Procento sebevrahů s přibývajícím věkem stoupá. Jak k tomu vlastně došlo? To je velice poučné: v „optimistické“ verzi o spokojeném stáří byly pozorovány příčiny smrti v jednotlivých věkových skupinách. Mladí lidé většinou neumírají na rakovinu, srdeční infarkt, mozkovou mrtvici nebo sešlostí věkem, umírají vůbec zřídka. Když umírají, tedy především v důsledku úrazů, ale také sebevražd. Ve skupině 15 až 40letých se sebevraždy podílejí na úmrtích 14 %. A na co umírají lidé po šedesátce? Na rakovinu, na infarkt, na mozkovou mrtvici, sešlostí věkem a — umírají všichni! Relativní podíl sebevrahů, představující asi 1 %, je velmi malý. Když se na to díváme takto, pak mladí lidé se odhodlávají k sebevraždě „relativně častěji“.

Správnější je provést srovnání s celkem věkové skupiny, a pak se ukazuje stoupající četnost (na 100 000 žijících stejného věku): ze tří u mladistvých pod 15 let na 44 u mužů nad 60 let. (U žen jsou čísla podstatně nižší, mají však stejnou tendenci.)

Podobných příkladů by bylo možno uvést ještě libovolný počet. Spokojíme



Tak se „lže“ statistikou.

První „statisticky dokázané zjištění“: sebevražednost se zvyšuje s rostoucím věkem. Průběh křivky grafu zřetelně ukazuje, že počet sebevrahů v roce a dané věkové skupině se plynule zvyšuje a ve věkové skupině „70 a starší“ činí více než desetinásobek skupiny mladistvých.

Druhé „statisticky dokázané zjištění“: ve středním věku dosahuje sebevražednost vrcholu, vzácné jsou naproti tomu mezi mladými a nejstaršími. Průběh křivky grafu zřetelně ukazuje, že z každých 100 sebevražd připadá asi $^1/_4$ na věkovou skupinu 51 až 60, avšak na sedmdesátileté (71—80) jen asi 10 % a na skupinu 21 až 30 let pouze 8 %.

Třetí „statisticky dokázané zjištění“: sebevražednost je nemocí mládí a její význam klesá s rostoucím věkem. Průběh křivky grafu zřetelně ukazuje, že podíl sebevrahů na celkovém počtu úmrtí jedné věkové skupiny klesá z 25 % mezi mladými lidmi (15—25 let) na méně než 1 %, jakmile se věk blíží sedmdesátce.

V závislosti na volbě základu — a) pro žijící stejné věkové skupiny, b) pro celkový počet sebevrahů, c) pro úmrtí ve stejné věkové skupině — vzniká zcela rozdílný obraz. Naše tři grafy se sice týkají tří různých zemí a různých období, a nejsou tedy v jednotlivostech srovnatelné, jsou však srovnatelné — a to je v daném případě důležité — charakteristickým průběhem křivky. (První graf znázorňuje sebevražednost mužů v Anglii a Walesu v roce 1956, druhý sebevražednost mužů v Hamburku v letech 1945—1958 a třetí sebevražednost ve Vídni v roce 1963.)

se však s těmi, které jsme uvedli, protože nám šlo jen o to zdůvodnit naše varování před příliš „lidově" zjednodušenými „statistikami".

Na které statistiky se tedy můžeme spolehnout? Zpravidla na *úřední statistiky*, na statistiky *velkých institucí a organizací*. Především však na ty, které uvádějí absolutní údaje, udávají rozsah výběrového souboru a pokud možno i některé údaje o způsobu zjišťování a pravděpodobnou teoretickou spolehlivost vzorku. I v těchto případech je však nutno připomenout: Nedoporučuje se číst mezi řádky! Kdo statistiku interpretuje, dělá to na vlastní nebezpečí. Kdo si nechá výklad nekriticky nakukat, nepotřebuje se na čísla ani dívat. Dobrá statistika poskytuje přehledně zpracované údaje, případně matematické souvislosti mezi těmito čísly, uvádí průměrné hodnoty a směrodatné odchylky, meze chyb, případně vysvětlující poznámky, nedokazuje však žádné hypotézy — ani vědecky „čisté", ani demagogicky „špinavé".

Protože však demagogové a skrytí manipulátoři statistiku tak rádi používají, je užitečné zeptat se v případě, že vznikají pochybnosti o seriózním vyjádření, ještě jednou nedůvěřivě: *cui bono* (komu to prospívá)? Tato otázka pomáhá již po staletí odhalovat zločiny a osvědčuje se často jako velmi užitečná i při odhalování statistických podvodů.

7.2 Počtářské umění demagogů

Proč jsou tak oblíbené a časté — a bohužel také často tak úspěšné — lži pomocí statistik? Je tomu tak proto, že průměrný občan vyrostl v *uctivé plachosti před čísly*, která jsou obklopena posvátnou, ale nenapadnutelnou přesností matematiky, a ve světě kolísavých představ jsou jediná absolutně spolehlivá. *Lichtenberg* sice řekl: „Jednou je nutno o všem pochybovat, i o tom, že dvakrát dvě jsou čtyři." Ale koho by vážně napadlo pochybovat o malé násobilce.

Vzhledem k tomu, že statistika pracuje převážně s čísly, přenáší důvěřivý občan svůj vztah k počtům také na čísla statistiky — ačkoli vedle toho může docela dobře obstát ze zkušenosti získané přesvědčení, že „statistiky lžou". Ve skutečnosti je obojí správné. Statistika používá matematických metod a matematické přesnosti a — statistika lže.

První, pozitivní představa mimochodem převládá, jinak by také nebylo tolik pokusů lhát pomocí statistik. Představa, že „čísla dokazují", není přes veškeré špatné zkušenosti překonána. Proto demagogové, manipulátoři a lháři používají statistiky všeho druhu obzvláště rádi, stejně jako fotografií, protože fotografie prý nelžou. V době amatérského filmování je ovšem možné seznámit se s tím, jak lze obrazy a scény sestavovat a manipulovat s nimi — avšak o statistických metodách ví laik strašně málo.

Již jsme vyčlenili ty druhy „statistických" lží, které jsou jen pouhé lži a vůbec ne statistiky, jako např. *falšování prvotních materiálů*. Jestliže chci proti svému lepšímu přesvědčení tvrdit, že v minulém měsíci se nevyskytl žádný případ cholery, nemohu kvůli tomu sestavit statistiku nemocí, která by zasluhovala tento název. Jestliže tvrdím, že můj nový přípravek je na základě laboratorních zkoušek o 65 % lepší než všechny ostatní preparáty (ačkoli ve skutečnosti je sotva stejně dobrý), je to pouhá lež, maskovaná jako statistická výpověď.

To však ponechme stranou; budeme se zabývat takovými manipulacemi, které přece jen mají základ ve skutečnostech, i když pochybný. Fakta velmi často nejsou přesně to, co se někomu právě hodí do konceptu, a proto je nutno vhodně je upravit a přizpůsobit (první příklad jsme ostatně poznali již v oddíle 4. 3, kde jsme nechali fiktivní zaměstnavatele a zaměstnance provádět kejkle s indexy).

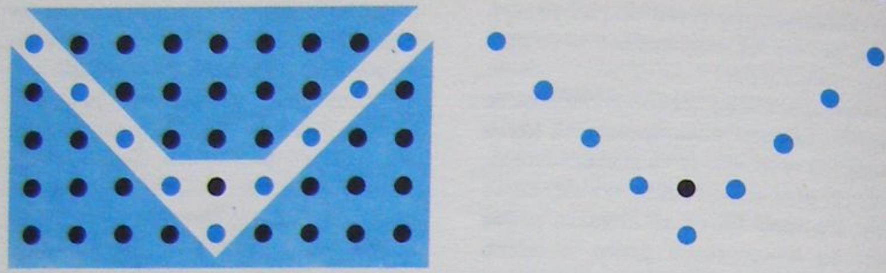
Pro tyto manipulace je charakteristické, že čísla a početní metody jsou sice úplně správné, že však se tím u čtenáře nebo posluchače vyvolává docela jiný, skutečnosti neodpovídající dojem. Dělá se to proto, aby se skrylo něco nepříjemného, předstíralo něco pozitivního, zkrátka ze všech možných důvodů, jichž reklama a propaganda užívají, aby se něco ukázalo zájemci v co nejlepším světle.

Při přepočtu je ovšem možno produkovat nesmysly různým způsobem. To se stává většinou tehdy, když se něco přepočítá a pak „seřadí za sebou". Lze na to uvést asi tento příklad: „Kdyby se všechny vlasy, které se denně ostříhají v mnichovských holičstvích, seřadily za sebou, sahaly by ze Země až ke Slunci (nebo k Marsu, event. k Jupiteru)." To je hra, která sice může řadu lidí značně obveselit, ale většinou vzniká na základě takových „srovnání" spíše větší než menší zmatek.

Na tyto přepočítávací hříčky existuje velké množství příkladů. Kdybychom chtěli papír, na kterém jsou vytištěny, slepit za sebou, dostali bychom se asi až na Měsíc. Věnujme se však raději vážnějším věcem a dejme si pod drobnohled velmi zvláštní druh procent, a to ta, která se vyskytují pod titulky typu „Devět z deseti generálních ředitelů používá tužkových ořezávátek Lignix" nebo „Tři z deseti hospodyň čistí stříbro Swobodovým argentorizidem". Na základě prodejních statistik však velmi dobře vím, že můj argentorizid používá stěží 1 % hospodyň. Lhal jsem tedy? Ano i ne. Mohu se např. prostřednictvím maloobchodu dostat do styku se dvěma nebo třemi desítkami žen, které tento čisticí prostředek kupují. To by bylo hned třicet spotřebitelek a k tomu připočtu ještě 70 výsledků anketního zjištění z náhodného výběru a tady snad také ještě přibudou jedna nebo dvě. Zkrátka: ze sta dotázaných hospodyněk jich nejméně třicet můj prostředek používá. Tři z deseti!

Jestliže naivní čtenář z toho usoudí, že jsou to vždy tři z každých deseti nebo vždy třicet z každého sta (což je v této formulaci podle smyslu stejné), zkrátka že asi třicet procent všech hospodyněk používá můj prostředek na čistění stříbra, mohu ještě vždy s politováním pokrčit rameny a hlasem vyjadřujícím hluboké přesvědčení ujistit: „To jsem netvrdil! O »z každých« nebylo ani řeči. To je nedorozumění." Samozřejmě žádné politováníhodné nedorozumění, spíše předpokládané, ba dokonce vyvolávané. Ale s touto formulací mám čisté ruce, zůstal jsem takříkajíc čestným člověkem.

Možná že by se dalo jít ještě o krok dále. Daruji např. padesáti generálním ředitelům nový druh ořezávače tužek a dotazy u jejich sekretářek pak zjistím, že devět generálních ředitelů tu věc opravdu zkusilo. A teď mohu říci: „devět z padesáti", mohu však také říci „devět z deseti". Že je to devět z deseti dodatečně vybraných, opatrně zamlčím. (Viz obr. na str. 212.)



9 ze 45

9 z 10

Mezi 45 kuličkami jsem našel devět modrých — mohu proto právem tvrdit, že bych mezi 10 kuličkami našel 9 modrých? Vždyť přece netvrdím, že mezi každou desítkou kuliček najdu 9 modrých! Takové slovní hříčky na pokraji serióznosti vrhají na „statistiku", která s tím nemá nic společného, špatné světlo.

To je metoda, která se ukáže správnou teprve až se dotyčné osoby identifikují. V inzerátu britského výrobce plynových spotřebičů bylo možné číst: „Devět z deseti proslulých šéfkuchařů používá k vaření plyn. Proč?" K tomu bylo připojeno devět fotografií a jmen kuchařů z renomovaných anglických hotelů. Avšak i to souvisí s otázkou solidnosti nebo nekalé soutěže více než se statistikou. Je přece veřejným tajemstvím, že známé osobnosti, především herci, dávají svá jména k dispozici pro reklamní účely. Samozřejmě za odpovídající honoráře.

Obtížnější je, když se povolají vědecké ústavy jako korunní svědci. Tu vzniká především otázka: Je daný ústav vůbec zařízením, které lze brát vážně? Za druhé: Nejsou z posudku citovány jen jednotlivé části? Za třetí a hlavně: Co je vlastně v posudku přesně obsaženo? Již dříve (v oddílu 4.4) jsme poukázali na možnost libovolného výběru srovnávaných hodnot. Upozornili jsme tam na mnohostrannou možnost mluvit, podle sledovaného účinku, raději v procentech nebo absolutních číslech a ukázali jsme, že by bylo celkem lehké zařadit např. Stuttgart jako páté velké město mezi deseti západoevropskými městy. Darmstadt může být největší mezi deseti největšími západoněmeckými městy — nebo také nejmenší — zcela podle libosti.

Volbou systému srovnání je rovněž možno dosáhnout různých účinků: Co jsou například „největší" noviny? Jsou to snad ty, které mají největší náklad, nebo ty, které mají největší odbyt nebo největší počet stran anebo nejvíce inzerátů (event. také obojí: jedny s největším počtem inzerátů, druhé s největší inzertní plochou nebo největším počtem stran); kromě toho to mohou být i noviny „nejčtenější", které nemusí mít největší náklad.

Další dva důležité manipulační obory ještě poznáme, až se budeme blíže zabývat grafickým znázorněním statistických údajů, a pak — v deváté kapitole — až budeme mluvit o korelacích všeobecně a o zdánlivých korelacích zvláště.

Někdy se statistiky zneužívá *k té největší demagogii*, jindy se jí však také dojímavým způsobem šidí. Tak např. potulný cirkus, jak jednou uváděl Rolf Wagenführ, velkými písmeny oznamoval „32 účinkujících". Skutečně pak vystoupilo dvanáct artistů, čtyři koně, šest opic a deset psů.

„Lež" této statistiky spočívá v tom, že si čtenář vykládá pojem „účinkující" zpravidla jinak než autor textu, který s tím počítá. Co se v daném případě jeví jako malá podivnost, postihuje i nejserióznější vědecky přesné statistiky. Již jednou jsme poukázali na to, že kategorie „nezaměstnaní", „smrtelné nehody", „studenti" a mnohé další jsou v jednotlivých zemích určovány často rozdílně, což prakticky znemožňuje jakékoliv srovnání. Kdo by se vážně pokoušel přesně srovnat počet „akademiků" v NSR a v USA?

„Odbory požadují zvýšení mezd o 15 %!" — „Zaměstnavatelé nabízejí jen 5 %!" Jsou tyto pozice zcela neslučitelné? Ne vždy. Uvedených 15 % se může týkat např. minimálních mezd podle kolektivních smluv, zatímco 5 % vyplácených mezd, které jsou možná podstatně vyšší než minimální mzdy.

Cirkus jednou na plakátech oznamoval „32 účinkujících": na tom nebylo vůbec nic vylhaného, pokud se za „účinkující" nepokládali jen artisté — lidé. Jakmile se setkají čísla s reklamou, je třeba nejvyšší opatrnosti. Příliš snadno vzniká představa „statistického souboru", o němž se vlastně nedá vůbec mluvit.

Mezi hrubou hodinovou mzdou za normální pracovní hodiny a průměrnou čistou hodinovou mzdou se zřetelem na nezdaněné příplatky, přesčasové hodiny, náhrady za dovolenou, přídavky na děti atd. jsou takové rozdíly, že je možno podle zvolené definice sestavit naprosto rozdílné statistiky a učinit zcela rozdílné závěry.

K tomu přistupuje ještě také skutečnost, že během určitého časového období *dochází k posunu pojmů:* průměrná „kriminalita" se samozřejmě zvýšila v důsledku růstu dopravních přestupků. Stane-li se homosexualita (nebo nějaký jiný delikt) beztrestná, počet deliktů automaticky klesne.

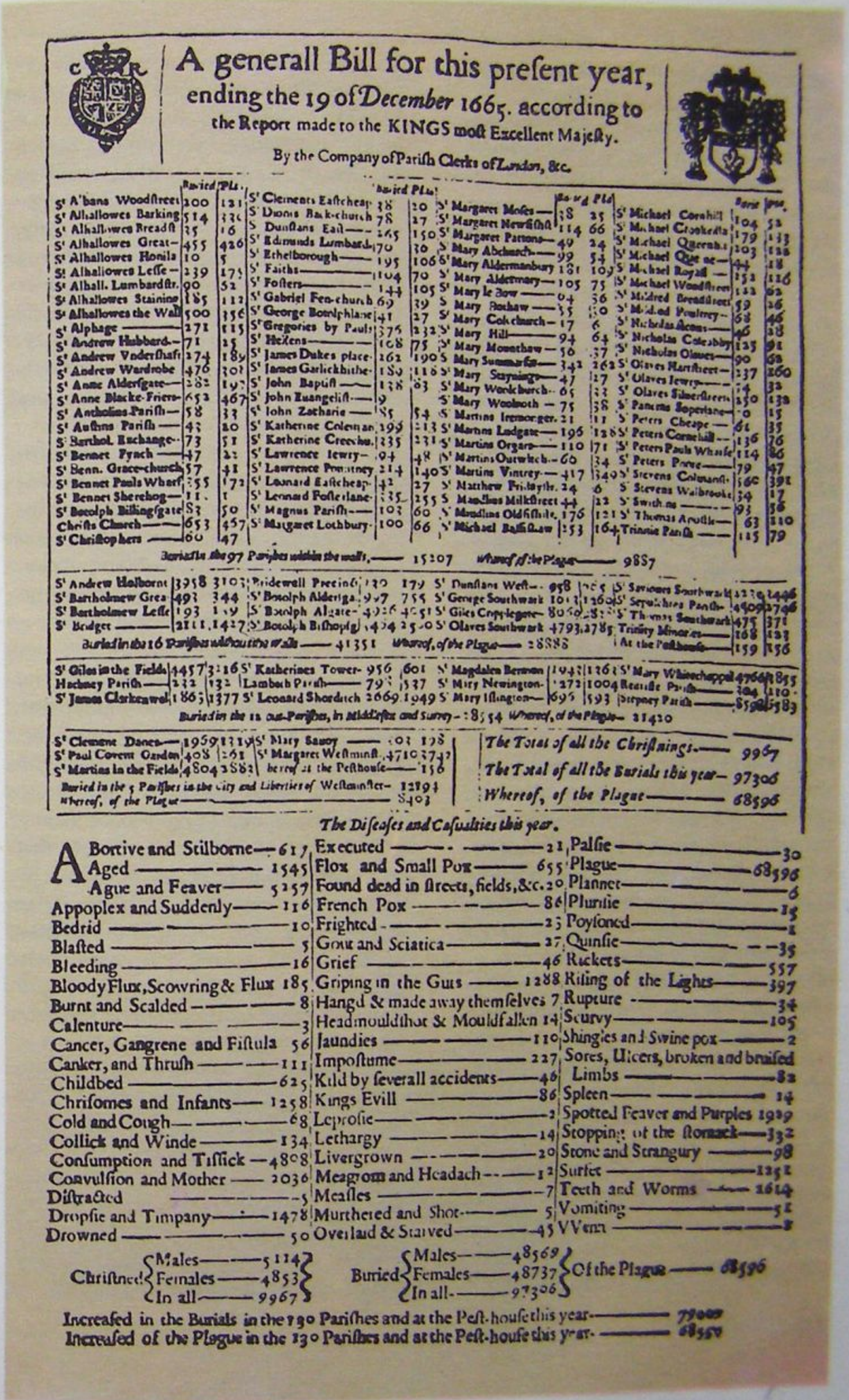
Také vzrůst počtu úmrtí na rakovinu nemusí „statisticky dokazovat" bezmocnost medicíny, stejně jako zvýšení průměrného věku nelze vysvětlovat jen jejím pokrokem. Může prostě jít o přesnější diagnózy. Seznamy příčin úmrtí před 100 lety a dnes nelze vůbec srovnat. V jednom středoevropském velkoměstě zemřelo v roce 1863 asi 17 000 osob — z toho 864 sešlostí věkem, 600 na fyzickou slabost, 4840 na souchotiny, 1135 na průjem, 180 v šestinedělí, 590 na tyfus, 1433 na zápal plic, 300 na černé neštovice a 302 na spálu. Jestliže pomineme dětskou a kojeneckou úmrtnost, zůstávají souchotiny, zápal plic, sešlost věkem a tyfus nejdůležitějšími příčinami smrti, zatímco o rakovině ani o chorobách krevního oběhu není ani zmínky. (Obr. na str. 215.) — Jako kuriozitu k tomu dodáváme, že Abraham de Moivre, kterého jsme poznali jako jednoho z velkých předků naší vědy, zemřel podle úmrtního listu na „somnolence", na ospalost.

Jiné nepřípustné srovnání: „Průzkumy jedné švédské automobilové firmy ukázaly, že ze 42 813 osob, které měly autonehody, jich 52 zemřelo; z toho 49 nemělo bezpečnostní pásy." Proč to vůbec nic neříká? Protože chybí údaj o procentech uživatelů bezpečnostních pásů. Jestliže například pásy používá jen 1 % automobilistů, je spíše pravděpodobné, že 3 z 52 obětí pásy nepoužívaly. Kapitola sama pro sebe jsou válečné statistiky. *K nespolehlivosti odhadu* se druží více nebo méně vědomé *falšování údajů vojenskými* a úředními místy. V únoru 1968 oznamoval časopis „Newsweek", že mrtví Vietkongu prý byli „často počítáni vícekrát", mimoto prý byli obvykle „na každou ukořistěnou zbraň počítáni 3 až 4 mrtví nepřátelé". Ministr obrany McNamara mluvil jednou o 165 000 mrtvých nepřátel za minulý rok, přičemž ovšem dodal, že tyto údaje se zakládají zčásti na „faktorech posuzování" (judgment factors) a musí se brát s rezervou.

V bouřlivém vnitropolitickém sporu v Rakousku o údajně „manipulované statistice socialistů" se vedle oprávněných námitek argumentovalo mimo jiné i takto: „Zatímco na začátku roku 1966 byl počet příjemců nejnižších příjmů (pod 1000 šilinků měsíčně) asi 230 000,

---

Statistická porovnání za dlouhá období jsou prakticky nemožná. Během několika staletí se úplně změnily dokonce i příčiny úmrtí — částečně v důsledku rozdílů v diagnostice. Matrika zemřelých z Londýna vykazuje v morovém roce 1665 přes 68 tisíc zemřelých na mor z celkového počtu 97 tisíc úmrtí. Většina dalších příčin smrti je dnes téměř neznámá nebo se nazývá jinak: střídavá zimnice, krvotok, psotník, božec, souchotiny, padoucnice, ujímání, bolesti v břiše, skvrnitý tyf a „surfet" snad bylo obžerství. Naproti tomu „cancer" není zřejmě rakovina v dnešním smyslu slova, nýbrž těch několik případů „cancer, gangrene and fistula" se týká hnisavých nádorů a snětí.



A generall Bill for this present year, ending the 19 of *December* 1665. according to the Report made to the KINGS most Excellent Majesty.

By the Company of Parish Clerks of *London*, &c.

The Total of all the Christnings — 9967

The Total of all the Burials this year — 97306

Whereof, of the Plague — 68596

The Diseases and Casualties this year.

Abortive and Stilborne	617	Executed	21	Palsie	30
Aged	1545	Flox and Small Pox	655	Plague	68596
Ague and Feaver	5257	Found dead in streets, fields, &c.	20	Plannet	6
Appoplex and Suddenly	116	French Pox	86	Plurisie	15
Bedrid	10	Frighted	23	Poysoned	1
Blasted	5	Gout and Sciatica	27	Quinsie	35
Bleeding	16	Grief	46	Rickets	557
Bloody Flux, Scowring & Flux	185	Griping in the Guts	1288	Rising of the Lights	397
Burnt and Scalded	8	Hangd & made away themselves	7	Rupture	34
Calenture	3	Headmouldshot & Mouldfallen	14	Scurvy	105
Cancer, Gangrene and Fistula	56	Jaundies	110	Shingles and Swine pox	2
Canker, and Thrush	111	Impostume	227	Sores, Ulcers, broken and bruised Limbs	82
Childbed	625	Kild by severall accidents	46	Spleen	14
Chrisomes and Infants	1258	Kings Evill	86	Spotted Feaver and Purples	1929
Cold and Cough	68	Leprosie	2	Stopping of the stomack	332
Collick and Winde	134	Lethargy	14	Stone and Strangury	98
Consumption and Tissick	4808	Livergrown	20	Surfet	1251
Convulsion and Mother	2036	Meagrom and Headach	12	Teeth and Worms	2614
Distracted	5	Measles	7	Vomiting	51
Dropsie and Timpany	1478	Murthered and Shot	5	Wenn	8
Drowned	50	Overlaid & Starved	45		

Christned: Males — 5114, Females — 4853, In all — 9967

Buried: Males — 48569, Females — 48737, In all — 97306; Of the Plague — 68596

Increased in the Burials in the 130 Parishes and at the Pest-house this year — 79009

Increased of the Plague in the 130 Parishes and at the Pest-house this year — 68590 činil počátkem roku 1969 již jen 120 000; zmenšil se tedy asi o polovinu."

Co se stalo? Začátkem roku 1966 byly nejnižší důchody těsně pod hranicí 1000 šilinků. V dalších letech se postupně zvyšovaly v souladu s vývojem indexu životních nákladů a v roce 1969 byly těsně nad 1000 šilinků. V tomto vývoji se zatím neodráží nic jiného než plíživé znehodnocení měny, ale již vůbec ne ústup chudoby — k tomu by bylo třeba prokázat, že nejnižší příjmy stouply relativně více než za prvé ceny a za druhé průměrné příjmy.

Dnes již zřídka rodí ženy doma. Ještě před několika desítiletími však bylo možno „statisticky dokázat", že porody doma jsou bezpečnější než na klinice. Z jednoduchého důvodu: jestliže hrozilo nebezpečí komplikací, doporučil ošetřující lékař porod na klinice. Proto porody na klinikách netvořily „reprezentativní průřez", nýbrž představovaly nadprůměrně vysoké procento komplikovaných těhotenství. Také na zápal plic se doma umírá méně často, protože těžké případy se posílají do nemocnic. V klimatických lázních umírá nadprůměrně mnoho lidí na tuberkulózu, v lázních, kde se léčí poruchy krevního oběhu, nadprůměrně mnoho na srdeční infarkt. Ti, kdo potřebují léčení, jsou přirozeně více ohroženi.

Nesmí se však také zapomenout, že právě provedené srovnání úmrtnosti u porodů doma a na klinice v ještě starších dobách vyjadřovalo skutečně něco správného. Předtím než Ignác Semmelweis objevil horečku omladnic a dokud se v porodnicích systematicky nepoužívala sterilizace, umíralo skutečně více matek na klinikách. Nebezpečí nákazy tam bylo opravdu větší než doma.

Na *hru s indexy* jsme již upozornili. K tomu ještě doplněk z americké vnitřní politiky. Wallis a Roberts uvádějí: „Před volbami v roce 1936 demokraté tvrdili, že prý zaměstnanost a výroba podstatně stouply, aniž se zvýšily životní náklady. Při svém srovnání vycházeli pro zaměstnanost a výrobu z roku 1933, ale pro životní náklady z let 1925 až 1929. Republikáni naproti tomu tvrdili, že prý životní náklady stouply, ale zaměstnanost a výroba se nezvýšily. Při výpočtu životních nákladů vycházeli ze základu roku 1933, u zaměstnanosti z průměru let 1925—1929."

Krásný případ působivé pseudostatistiky poskytuje také úvaha, kterou jednou pronesl americký rozhlasový komentátor v souvislosti *s projektem reformy kalendáře*. „Dokázal", že náš nepohodlný kalendářní systém stojí jen samo město New York přes 5 miliónů dolarů ročně. První důvod spočíval v tom, že je prý zatím třeba asi 1 minuty k tomu, aby se k libovolnému datu našel v kalendáři odpovídající den v týdnu. Jako druhý důvod uváděl, že to dělá každý obyvatel New Yorku každý druhý den. Tím vzniká časová ztráta přes 58 300 hodin denně. Průměrná dělnická mzda činila tenkrát jen 25 centů za hodinu, ale i s tímto číslem došel tento statistik-amatér k denní ztrátě 14 583 dolarů a k roční ztrátě přes 5,3 miliónu. Podle dnešních průměrných mezd by tento naivní propočet ukázal ztrátu přes 40 miliónů dolarů.

Ponecháme na čtenáři, aby si sám vzpomněl na další hrubé logické a statistické chyby. Uvedený příklad je spíše zábavný než demagogický, ale jeho schéma přesně odpovídá schématu, které se někdy používá, ať již z nerozumu nebo z demagogie, také při vážnějších problémech. Podobně jako výhodu zjednodušeného kalendáře je

možno „dokázat" téměř všechno. *Hranici mezi zábavou a vážným pokusem o oklamání pomocí pseudostatistik* lze často stanovit jen obtížně.

Mohli bychom donekonečna uvádět takovéto „výkvěty", ležící na fascinujícím pomezí demagogie a naivity. Závěrem snad ještě jeden příklad výpočtu ve dvou verzích. Nejdříve však čísla: Za prvé, spolková vláda a zemské vlády NSR vydaly v roce 1968 asi 6 miliard DM na úsek „veřejná bezpečnost, pořádek a právní ochrana". Za druhé, počet obyvatel činil zhruba 60 miliónů. Za třetí, dani ze mzdy a z příjmů podléhalo asi 28 miliónů osob.

První statisticky doložená skutečnost: Každý poplatník musí v průměru ročně zaplatit hodně přes 200 DM na policii a soudy. Druhá skutečnost podložená stejnými údaji: Na jednoho obyvatele se na soudnictví a udržení veřejného pořádku denně nevynakládá ani 30 feniků (toto druhé číslo je možno podle potřeby chápat buď tak, že se od občana vybírá jen tato zanedbatelná částka, nebo i tak, že udržení pořádku a právní jistoty státu nestojí za víc než za těch pár grošů).

Chci-li však ukázat, že podíly daní na tuto kapitolu rozpočtu jsou mimořádně nízké, uvědomím si, že jen asi 70 % státních výdajů NSR se kryje z daňových výnosů — a pak již připadá na jednoho obyvatele jen asi 20 feniků daní na den.

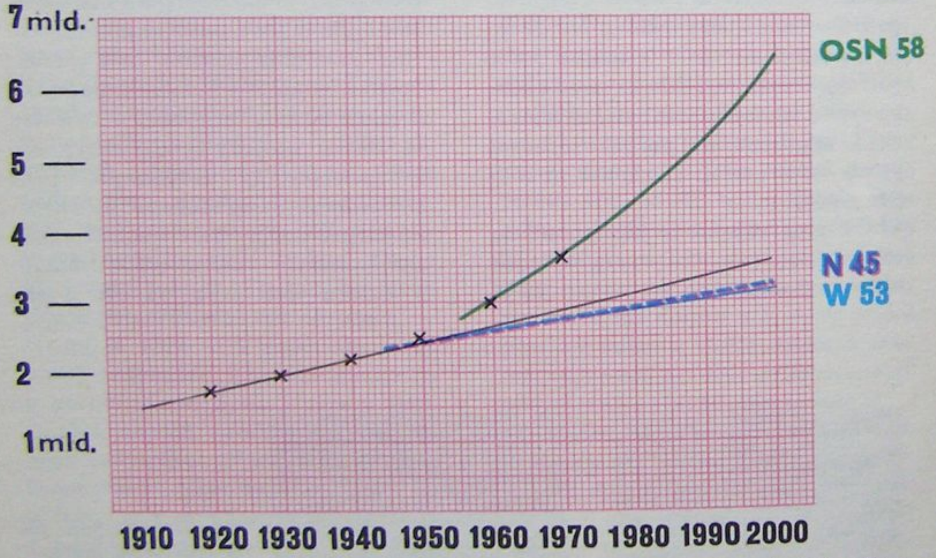
7.3 Od extrapolace k futurologii

„Budoucnost není možno plánovat ve světle minulosti," řekl prý kdysi Edmund Burke, jinak tradicionalistický a konzervativní státovědný teoretik. Odmítne-li se „světlo minulosti", ocitáme se v úplné „temnotě" nebo jsme odkázáni jen na magickou zář fantazie. Chceme-li však přistoupit k dešifrování budoucnosti na základě statistického materiálu, musíme si být vědomi toho, že *naše čísla a údaje pocházejí z minulosti*. I když naším úmyslem není *„průzkum budoucnosti"* v užším slova smyslu, přece jen se používá statistických údajů téměř bez výjimky jako *pomůcky pro rozhodování*. Studují se relativní čísla přítomnosti a minulosti k získání poznatků o jejich budoucí podobě. Po celá léta se sledují „trendy", aby mohly být *extrapolovány* do budoucnosti. Také sledování „stálého souboru" není určeno jen pro archívy nebo k zajištění plné zaměstnanosti statistických úřadů. V daleko větší míře má být orien-

Každý daňový poplatník	—	**Na jednoho obyvatele**
musí platit	—	**se vynakládá**
ročně	—	**denně**
daleko více než	—	**ani ne**
200 marek	—	**30 feniků**
na policii a soudy	—	**na právní ochranu a udržení pořádku**

Stejnou skutečnost lze vyjádřit velmi rozdílnými slovy, takže podle přání zní „pozitivně" nebo „negativně". Kvůli tomu však ještě „statistika" nelže.

tační pomůckou pro rozhodování všeho druhu.
Avšak minulost znamená vždy jistotu, budoucnost vždy jen možnost a pravděpodobnost. Co bylo po dvacet let, nemusí již být zítra, ale přece jen to bude pravděpodobné. Jestliže nemoci srdce a krevního oběhu tvořily v posledních letech v průměru 30 % příčin úmrtí, je možno předpokládat, že letos bude číslo podobné. Dokonce to lze předpokládat i pro rok 1980 nebo dokonce 2000.
Podívejme se na zcela jednoduchý (tak se alespoň zpočátku jeví), dobře známý a velmi diskutovaný příklad — *populační explozi*. Obyvatelstvo světa čítalo v roce 1800 asi 775 miliónů, v roce 1850 1,1 miliardy, v roce 1900 1,6 miliardy a v roce 1950 2,4 miliardy. Tento růst populace můžeme vyjádřit graficky a dostaneme křivku, kterou lze proložit do budoucnosti. Přesněji řečeno neobdržíme jedinou jednoznačnou křivku, nýbrž je možné — buď podle matematických znalostí, nebo temperamentu, intuice a informovanosti — promítnout do budoucnosti křivky velmi rozdílné. V závislosti na tom obdržíme i pro rok 2000 velmi rozdílné hodnoty. Tým futurologů Kahna a Wienera počítá s asi 6,5 miliardy, ponechává však dostatek prostoru i pro jiná čísla a čtenáři poskytuje pro určení počtu obyvatel Číny v roce 2000 výběr mezi hodnotami 922 a 1600 miliónů. V daném případě



Extrapolace obyvatelstva světa. V roce 1945 a pak v roce 1953 vypočítali Notestein (N) a Woytinski (W) na základě tehdy dostupných údajů, že v roce 2000 bude na světě více než 3 miliardy lidí. V roce 1958 se již zřetelně projevila populační exploze v rozvojových zemích, a tak extrapolace, kterou v tomto roce provedla OSN, odhaduje již počet obyvatel světa v roce 2000 na více než 6 miliard.

nelze moc zkazit, skutečná hodnota bude někde mezi tím. Německý fyzik a futurolog Wilhelm Fucks dochází k 6,6 miliardy. Statistická ročenka NSR uvádí pro rok 1970 na podkladě údajů Statistického úřadu Spojených národů počet obyvatel světa v roce 2000 ve výši 6 129 700 000, což je poněkud podivné vzhledem k tomu, že by muselo dojít k šťastné náhodě, kdyby se tento počet ukázal správným alespoň na stamiliónové místo.
Jiné odhady z novější doby se pohybují mezi 6 miliardami a více než 7 miliardami a může se zdát, že bylo dosaženo dosti značného souladu a že by proto bylo dobře možné extrapolovat dosavadní vývoj uvnitř poměrně úzkých hranic. Při bližším zkoumání se však tato čísla objeví v jiném světle. Mnozí autoři zabývající se rokem 2000 mohou totiž stěží sami provádět statistická šetření o růstu počtu obyvatel světa, a důvěřují proto údajům o sčítání lidu a odhadům Statistického úřadu OSN a jeho extrapolačním metodám.
Použití poněkud starších odhadů není bez nebezpečí, jak hned uvidíme. Dennis Gabor píše ve své knize „*Inventing the Future*“ (Vynalézání zítřka) docela správně: „Kdyby se populační vývoj v roce 1940 extrapoloval z údajů, které tehdy byly k dispozici, došlo by se pro rok 2000 při extrapolaci trendu třicátých let k číslu 4,4 miliardy, s přihlédnutím k celkovému průběhu od roku 1920 do roku 1940 pouze k číslu 3,5 miliardy — a toto číslo bylo tehdy skutečně pokládáno za nejspolehlivější odhad.“
Dobrý a svědomitý odhad Franka W. Notesteina v roce 1945 uváděl pro rok 2000 počet obyvatel světa pouze ve výši 3,35 miliardy, a v roce 1953 pečlivě zpracovaný odhad Woytinských dokonce ještě nižší: 3,25 miliardy! Německý hospodářský statistik Ernst Wagemann psal v tomto smyslu ještě v roce 1952: „Když počet obyvatelstva poroste i nadále stejným tempem, obývalo by za sedmdesát pět let Zemi již 5 miliard lidských bytostí.“ Za 75 let — to znamená, že se to vztahovalo k roku 2025!
Ať se zdá tato prognóza jakkoli falešná, přece jen je správná, čteme-li přesně, co Wagemann tehdy napsal: „Když počet obyvatelstva poroste i nadále stejným tempem (tj. jako v posledních desetiletích před rokem 1950)...“ Kdyby se byl tento předpoklad splnil, byl by uvedený odhad správný. Wagemann sám hned ještě poznamenal: „Jak lehce lze při populačně politických prognózách upadnout do omylu, ukazují nejlépe dřívější pokusy v této oblasti.“
Tak Gregory King odhadoval v roce 1696, že do roku 3500 bude v Anglii žít asi 22 miliónů lidí. (Dnes má Anglie a Wales asi 48 miliónů.) King počítal takto: v roce 1260 bylo napočteno 2,75 miliónu lidí, v roce 1695 jich bylo dvakrát tolik — 5,5 miliónu. K zdvojnásobení bylo tedy zapotřebí 435 let; z toho vyplývá i dalších 435 let na zdvojnásobení na 11 miliónů (v roce 2130) a „již“ v roce 2565 další zdvojnásobení na 22 miliónů obyvatel. Gregory King se domníval, že podobný prudký růst je na dlouhou dobu nemyslitelný, že pro budoucí zdvojnásobení se musí předpokládat asi dvakrát delší časové období než dosud — tedy 11 miliónů lidí kolem roku 2565 a 22 miliónů kolem roku 3440.
Ještě jeden citát z Wagemanna: „Každá prognóza se vaří jen z vody, tj. zkušenosti z minulosti, tak jak vznikaly ve své době, transponuje do budoucnosti. Nemáme proto žádný důvod k povýšenému úsměvu, jestliže velký kritik spo-

lečnosti, Voltaire, v 18. století prohlásil, že považuje za fantazii předpoklady o možném zvyšování počtu obyvatelstva byť i jen o 5 % za sto let. Totéž platí o Montesquieuově pesimismu, když byl přesvědčen, že se počet lidí den ze dne zmenšuje a že za tisíc let lidstvo zcela přirozeným způsobem zmizí s povrchu zemského."

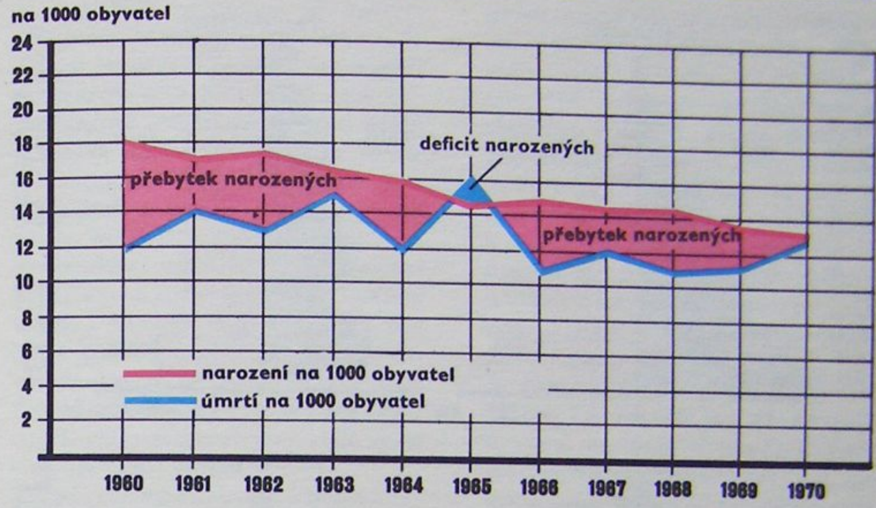
Dnešní futurolog se pravděpodobně přece jen bude povýšeně usmívat a prohlašovat, že při vší úctě k Voltairovi a Montesquieovi nelze oba pány přece jen považovat za vědecké futurology. Především prý také tehdejší statistický podkladový materiál byl nedostačující, a technika i teorie extrapolace prý na tom byly obzvláště špatně. Futurologové jsou ochotni připustit, že ani kolem roku 1900 nebylo možno udělat relativně přesné prognózy vývoje na příštích sedmdesát let. Právě jsme však viděli, že ještě v roce 1950 byly demografické prognózy zcela nesprávné. Přesto však nás odborníci ujišťují, že nyní se pracuje přísně vědecky, a proto také spolehlivě. Ještě se k tomu vrátíme, nejdříve se však podíváme na jinou pozoruhodnou extrapolaci.

W. S. Woytinski uvádí ve své knize „*Limits of Mathematics in Statistics*" (Meze matematiky ve statistice) příklad, který je plný ironie, protože jsme právě poznali Woytinského v souvislosti s velmi nešťastnou prognózou z nedávné minulosti. Woytinski píše o matematikovi N. S. Pritchettovi, který koncem minulého století extrapoloval demografický vývoj Spojených států v letech 1790 až 1890 tak, že těchto jedenáct ročních hodnot vynesl do systému souřadnic, tyto body pak proložil parabolou třetího stupně a — matematicky bezvadně — propočítal počet obyvatel Spojených států až do roku 2900. Stanovil, že počet obyvatel USA v uvedené době dosáhne 41 miliard. Nuže, Woytinski se Pritchettovi smál, ale my dnes čteme — zčásti zděšeně, zčásti lhostejně — že sice nikoliv obyvatelstvo Spojených států samých, ale obyvatelstvo světa dosáhne počtu 40 miliard — nikoliv ovšem až v roce 2900, nýbrž již ve druhé polovině příštího století, tedy kolem roku 2075.

Dříve než přikročíme k závěru a poněkud blíže osvětlíme statisticko-matematické aspekty tohoto problému, zodpovíme otázku, kterou by snad čtenář mohl položit: proč se vůbec tak podrobně zabýváme právě vývojem obyvatelstva? Máme pro to více důvodů. Za prvé, projevujeme tím úctu *nejstarší větvi statistiky*, statistice obyvatelstva. Za druhé, *populační exploze* je dnes obecně uznávána jako *životně důležitý problém* a doporučuje se vědět něco o možnostech a mezích těchto prognóz. Za třetí a především — tento příklad jsme zvolili proto, že je sice poměrně jednoduchý, přesto však naznačuje nekonečně *obtížné problémy prognóz*.

Jak to, že statistika obyvatelstva je poměrně jednoduchá? Především je možno s pomocí úmrtnostních tabulek zjistit, kolik z dnes žijících lidí bude ještě naživu za deset, dvacet nebo třicet let. Za druhé je možno s ještě větší pravděpodobností určit, kolik osob dosáhne v příštích dvaceti letech sňatkového věku. Za třetí lze — ovšem již s velmi omezenou pravděpodobností — předpovídat počty narozených v nejbližších letech a snad i desetiletích.

Jak to tedy, že je statistika obyvatelstva přesto tak obtížná a plná nejtěžších chyb? Abychom to mohli objasnit, musíme si především uvědomit, jak vůbec dochází k „přirozenému pohybu obyvatelstva".

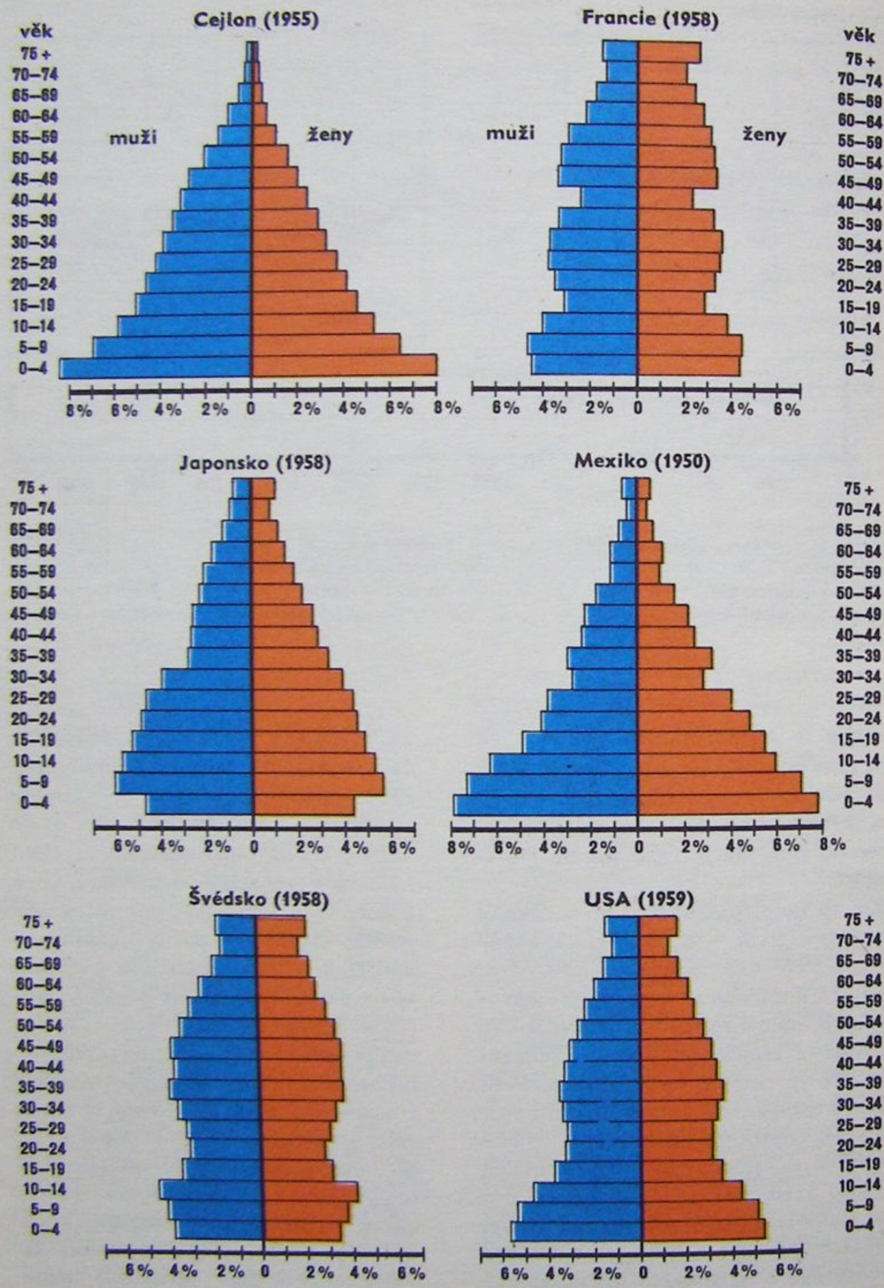


Přebytek a deficit narozených jsou početním výsledkem dvou „proměnných" — počtu narozených a počtu úmrtí. Jestliže každá rodina má 7 dětí, vznikají veliké počty narozených; jestliže z nich však 5 předčasně zemře, jsou také vysoké počty úmrtí a nevzniká „populační exploze". (Náš graf je pouze ilustrativní, udává však pro střední Evropu charakteristické počty narození a úmrtí.)

Často se nesprávně předpokládá, že populační exploze je výsledkem zvýšené porodnosti. To je úplně, nebo řečeno opatrněji, téměř úplně nesprávné. Skutečnou příčinou je spíše snížení úmrtnosti.

Zní to nesmyslně vzhledem k zákonu lidského života, který zní, že každý musí dříve nebo později zemřít. Ten, kdo se dnes objeví ve statistice narozených, objeví se jednoho dne také ve statistice zemřelých. Čím je však základna pyramidy obyvatelstva širší, tím menší je relativní úmrtnost (úmrtí na 100 obyvatel): ve vrcholku několik zemře, v základě se mnoho narodí. Tento vývoj, tak charakteristický pro dnešní rozvojové země, vzniká tehdy, když se výrazně snižuje úmrtnost kojenců, dětí a mladistvých, avšak zůstává zachován počet porodů připadajících na ženu v plodném věku. (Obr. na str. 222.)

Proč se teprve nyní volá po „*kontrole porodnosti*" a „*pilulce*"? Protože až do doby před několika desetiletími „*kontrola smrti*", prováděná přírodou, vznik tohoto problému vůbec ani nedovolila: jestliže ze sedmi živě narozených dětí v rodině jen dvě dosáhnou plodného věku, nemůže vzniknout populační exploze. Jestliže se však všech sedm dětí dožije věku sňatku, jsou oba rodiče nahrazeni 3,5 manželského páru a každý z nich — za předpokladu stejné plodivosti — je opět nahrazen 3,5 páru atd. až — až dokdy? Dokud nezačne působit některý z kontrolních mechanismů — kontrola porodnosti nebo kontrola smrti. Stromy nerostou do nebe; obyvatelstvo Spojených států



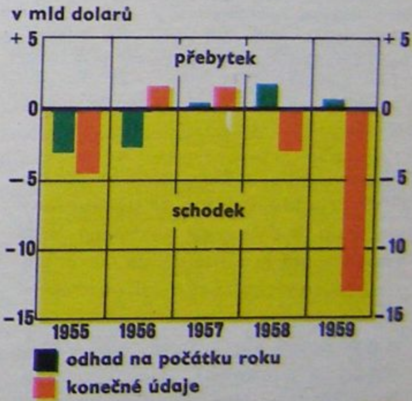
v roce 2900 nedosáhne 49 miliard. Pro statistika to především znamená, že nestačí pozorovat jen přírůstek obyvatelstva, ale především vývoj dvou křivek, jejichž „rozevření" teprve udává přírůstek obyvatelstva: *počet porodů* (absolutní a relativní) a *počet úmrtí* (absolutně a relativně). Teprve pak přijde mnohem obtížnější část: jak extrapolovat tyto údaje do budoucnosti? V jaké míře respektovat kolísání vývoje z poslední doby nebo více přihlížet k dlouhodobějšímu vývoji? V třicátých letech se počet obyvatelstva Evropy téměř nezměnil, v padesátých letech silně vzrostl — jaký bude vývoj v osmdesátých letech?

Zde končí matematická statistika a začíná dobrodružství domněnky — „*ars coniectandi*" (umění dohadu), jak kdysi nazval Jacob Bernoulli svou knihu o počtu pravděpodobnosti. V daném případě nejde o početní teorii pravděpodobnosti, ale o odhady a předpovědi sociologického, hospodářského a technického vývoje, které spíše patří do oblasti tvůrčí fantazie než do statistické analýzy.

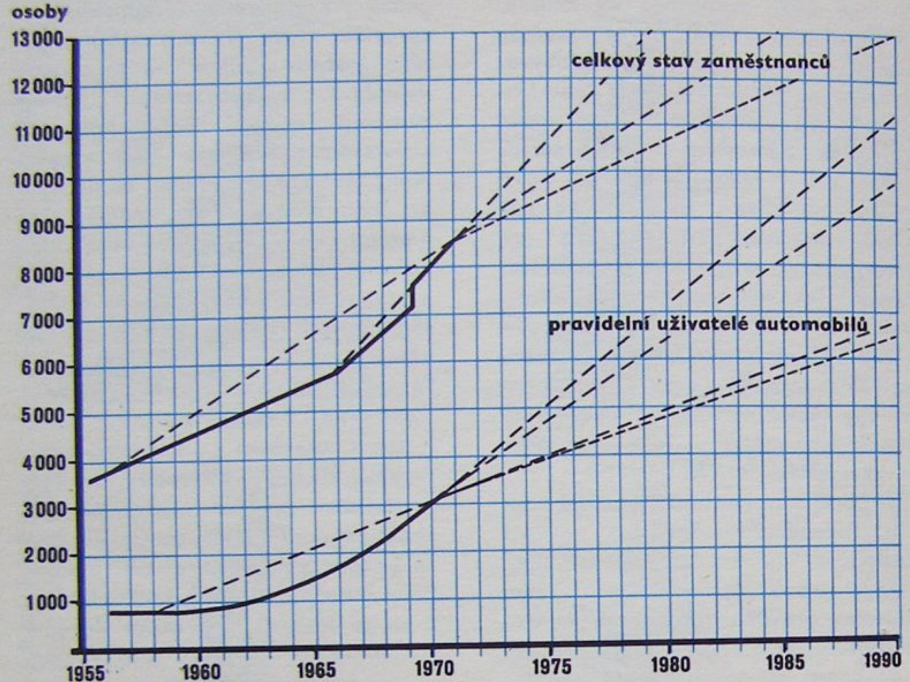
Statistika nemůže v této situaci dělat nic jiného, než propočítávat matematické důsledky předpokladů podle schématu: „za jinak stejných okolností pak...", nebo: „když růst obyvatelstva světa bude plynule klesat až na roční přírůstek 1,3 %, pak...". Každá prognóza vychází z určitého předpokladu — zpravidla přinejmenším z mlčky přijatého „ceteris paribus" — že experimentuji-li s určitou veličinou, ostatní zůstávají neměnné. Každá prognóza obyvatelstva pro rok 2000 vychází mimo jiné z předpokladu, že nedojde ke třetí světové válce, zejména ne k atomové válce. Nelze totiž vůbec odhadnout, jak by taková válka skončila a kdy, zda by nebyly vyhubeny $^4/_5$ lidstva a zbytek z převážné většiny smrtelně nemocný.

Je proto každá prognóza nesmyslná? Jistěže ne. Protože, za prvé, slouží k tomu, abychom si namáhali hlavu tvorbou vhodných alternativ, a v tom koneckonců spočívá rozhodující význam průzkumu budoucnosti. Za druhé, prognózy jsou o to reálnější a prakticky cennější, čím jsou krátkodobější, a to ze zcela jednoduchého důvodu. V „dohledné" době je prudký zlom

◄ Věkové pyramidy různých struktur. Pro rozvojové země je charakteristický Cejlon se stále ještě vysokou úmrtností dětí a mládeže. V Japonsku se ve slabé věkové skupině pod 4 roky projevuje účinek kontroly porodnosti. Nepatrná úmrtnost v mladém a středním věku je zřejmá ve Švédsku a v USA.



Především to vždy dopadne jinak — pětiletí rozpočtu Spojených států v odhadech a skutečnosti. Ani jinde nejsou prognózy podstatně spolehlivější.



Švýcarský chemický koncern koncem roku 1970 „extrapoloval" celkový stav svých zaměstnanců a podíl pravidelných uživatelů automobilů až do roku 1990. Z tohoto odhadu vyplynul pro rok 1990 počet 13 000 zaměstnanců, z toho 6500 pravidelných uživatelů automobilů. Graf ukazuje, že grafickými extrapolacemi je možno dojít ke zcela jiným údajům. Pokud jde o celkový stav zaměstnanců, může být jeho výše samozřejmě určována dlouhodobým plánováním podniku. Podíl uživatelů automobilů pro rok 1990 se však dá stěží určit.

vývoje příliš málo pravděpodobný. (Velmi rozšířený názor, že střednědobý průzkum budoucnosti poskytuje lepší výsledky než krátkodobý, je nutno považovat za významový omyl: od krátkodobé prognózy požadujeme podrobnosti a přesnost, které vůbec nelze očekávat od střednědobé.)

Nakonec lze ještě namítnout, že na vývoj obyvatelstva snad působí iracionální síly, zatímco při extrapolaci technického a hospodářského vývoje není nutno mít obavy před nečekanými „zlomy". To však bohužel není správné. Stát s centrálně plánovaným hospodářstvím může sice dalekosáhle řídit spotřebu a výrobu, nicméně zkušenost ukázala, že plánovaných hodnot se nikdy přesně nedosáhlo: v nejlepším případě jsou daleko překročeny, v nejhorším jich nebylo ani zdaleka dosaženo.

K tomu ještě přistupuje, jak již víme, skutečnost, že vysoké procentuální přírůstky při nízké absolutní úrovni jsou zcela dosažitelné, při vyšší základně jsou

však nemožné. Některá země může výrobu automobilů v průběhu jednoho roku zvýšit ze 100 000 na 200 000 jednotek, je však vyloučeno, aby se výroba z 10 miliónů zvýšila za stejnou dobu na 20 miliónů.

Posuzujeme-li hospodářské prognózy poněkud skeptičtěji než prognózy obyvatelstva, má to svůj zvláštní důvod. Obvykle vývoj hospodářství úzce souvisí s vývojem obyvatelstva. Tak např. lze již dnes docela dobře odhadnout potřeby školních prostor na nejbližších 5 nebo 10 let. Nyní však víme, že ani vývoj obyvatelstva nelze přesně vypočítat předem.

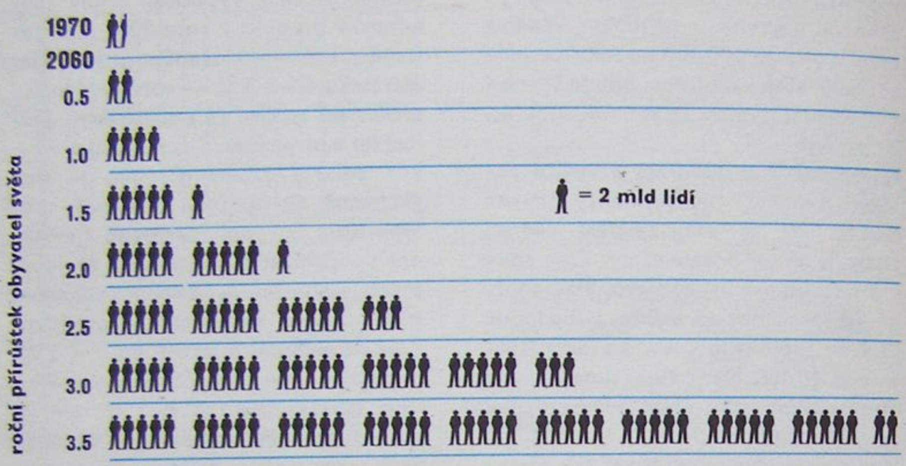
Mimoto zde vzniká problém, jak racionálně nebo také iracionálně (např. móda je skoro vždy iracionální) se bude obyvatelstvo chovat za 2 roky, za 5 let, za 10 let, o co bude usilovat a co bude odmítat. Při nedostatku zboží lze ještě takové propočty dělat s určitou pravděpodobností, sotva je to však možné ve společnosti, která má přebytky zboží. Co spotřebitele omrzí, které reklamě podlehne, které netušené psychologické faktory se objeví?

Před 10 lety, i když v podstatě již existovaly, zůstávaly zcela bez povšimnutí problémy znečištění ovzduší, řek, devastace životního prostředí a další škody napáchané civilizací. Ještě před 5 lety sotva pronikaly do vědomí průměrného občana. Neovlivní to však za 5 let jeho spotřebitelské chování, a jak to bude vypadat za 10 let? Neztratí do té doby platnost i neúprosný zákon hospodářského pokroku, který nás vždy vedl k přesvědčení, že se musí vyrábět stále více a laciněji (i když ne vždy lépe a zdravěji)? Jak se vyvine světový obchod, když se dospívající generace zřekne hektického růstu spotřeby?

Není žádným uměním, nýbrž pouze početním úkonem vypočítat hrubý společenský produkt v roce 2000 alternativně při předpokládaném průměrném přírůstku 4,5 a 5 % — *opravdové umění předpovědí spočívá ve vizionářském, prorockém a utopickém.* Pro statistika vznikají takové problémy, jako je sled vzájemně navazujících intervalů spolehlivosti. To např. znamená (jestliže pro současnost zvolíme opět základní index 100), že nejdříve se odhaduje růst hrubého společenského produktu v obou nejbližších letech na 3 a 5 %; pro další 3 roky se počítá sice se stejným růstem (mezi 3 a 5 %), ale již s menší pravděpodobností (asi 90 místo 95 %), pro dalších 5 let s růstem mezi 2,5 a 4,5 % (s pravděpodobností 85%). I když se nepřihlíží k tomu, že můj odhad pravděpodobnosti nemá matematický podklad, nýbrž je velmi subjektivní, co to říká o situaci, která nastane za 10 let? Jestliže vývoj zůstane převážně v dolní části mých odhadů, bude za 10 let index asi 134. Jsou-li správnější optimistické odhady, může vystoupit na 158.

To je již velmi široký interval spolehlivosti. Avšak i ten předpokládá, že zvolené pravděpodobnosti jsou téměř zcela správné. Nesmí se vyskytnout systematická chyba, nesmí se přihodit něco neočekávaného, ani se nesmí nějakou náhodou (které např. při 85% pravděpodobnosti zůstává přece jen vyhrazeno 15 %) ukázat odhadovaný rozsah jako příliš úzký. Zkušenost ostatně vždy znovu ukazuje, že hospodářské prognózy se nesplnily ani pro nejbližší jeden až dva roky. (Dodatečně se také vědělo proč, ale my jsme to chtěli vědět předem.)

I když jsou některé statistické materiály a extrapolace obdivuhodně přesné, mohou takovými být jen za předpokla-



Alternativy pro svět zítřka. Růst obyvatelstva světa do roku 2060, vyjádřený obrázkovou statistikou (podle „Population Bulletinu“) za předpokladu sedmi různých procentních přírůstků.

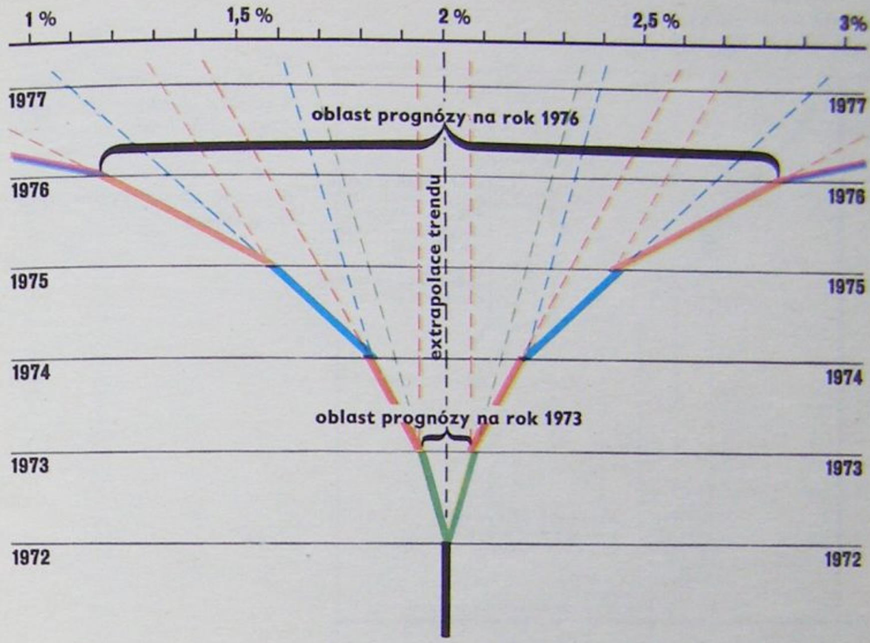
du, že vývoj bude přibližně stejný jako dosud. Je-li tento předpoklad nesprávný, čísla nemohou souhlasit. Pak statistika zase jednou lže...

Jiné nebezpečí ukvapené extrapolace stručně ukážeme na příkladu z nedávné minulosti. V roce 1969 vyvolala velké vzrušení zpráva amerického fyzika z oboru záření Sternglasse, který na základě velkého počtu údajů a grafických zobrazení ukazoval, že relativní počet *mrtvě narozených dětí* v americkém státě *Missouri* od počátku pokusů s atomovou bombou v Nevadské poušti poklesl podstatně méně výrazně než v uplynulých 15 letech, ba dokonce se téměř nezměnil.

Jedním z nejpůsobivějších obrázků byla křivka extrapolovaného průběhu za léta 1935 až 1950, porovnaná se skutečnými údaji o mrtvě narozených dětech.

Na základě předložených materiálů nelze zcela bezpečně vyloučit možnost, že zvýšené záření stroncia 90 „významně“ ublížilo ještě nenarozeným dětem. Avšak hypotéza „*atomové pokusy jsou vinny*“ byla již znehodnocena absurdní extrapolací křivky vývoje. Nikdo nám totiž nemůže zabránit, abychom tuto křivku neprodloužili ještě dále, a pak by — i bez atomových pokusů — již v roce 1980 nebyl ani jediný případ mrtvě narozeného dítěte. (Po tomto datu by dokonce muselo docházet ke křísení mrtvých nebo k živému narození nezplozených dětí.)

Přehlédla se a zanedbala skutečnost, že určité úkazy mohou sice dosahovat výrazných vzestupů nebo poklesů, ale že *nemohou být nekonečně velké, ani že se nemohou rovnat nule*. Na počátku třicátých let operovali v Anglii odpůrci očkování podobnou extrapolací jako profesor Sternglass. „Dokázali“, že od zavedení očkování začátkem třicá-

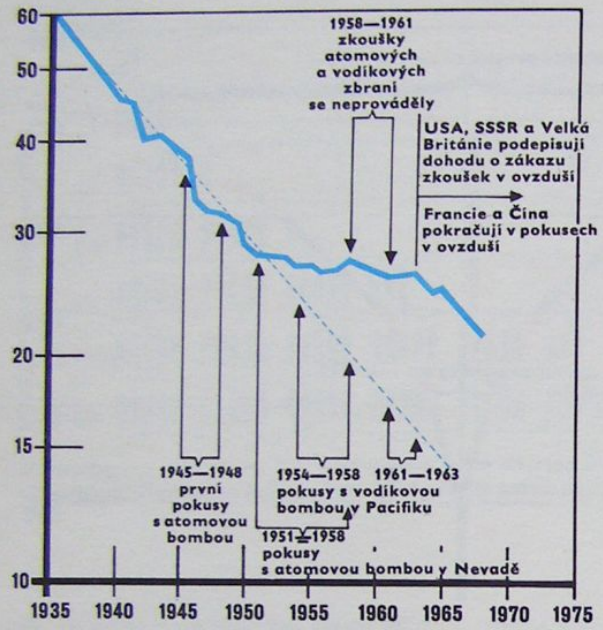


Každým rokem se stále více rozevírá vějíř prognózy. Je-li pro příští rok ještě možno vymezit s vysokou spolehlivostí relativně úzký rámec prognózy, pak pro pozdější období je stále širší (nebo stále méně spolehlivý). Výchozí předpoklad: přírůstek (např. obyvatelstva) v roce 1972: 2 %.

tých let prý poklesla úmrtnost celého obyvatelstva podstatně méně (ba vůbec ne) než v desítiletích od roku 1890 do roku 1930. Dokonce extrapolovali, jak hluboko by úmrtnost podle jejich mínění musela klesnout do roku 1940. Extrapolace do roku 1950 by i v tomto případě vedla ke „křísení mrtvých“. Tento příklad je ještě podivnější než Sternglassův. Je zcela dobře možné si představit, že kdysi snad vůbec nebyly případy mrtvě narozených dětí, ale je jen stěží představitelné, že by někdy nebyly v celém obyvatelstvu žádné případy úmrtí a že by populační exploze byla navíc podporována věčným životem na Zemi.

Na závěr je nutno ještě krátce upozornit na filozofický problém předpovědí; zabývat se jím podrobněji není však na tomto místě vhodné. Je ovšem zcela možné zastávat názor, že budoucí dění je předurčeno — ať již boží vůlí, „vestavěnou“ automatikou vesmíru nebo historickou zákonitostí. Aby bylo možno předpovídat budoucnost, vzniká v tomto případě snaha objevit *početní vzorce této zákonitosti*, a to podobně jako se první statistikové domnívali, že mohou dokazovat boží vůli z údajů o statistice

Úmrtnost amerických dětí do 1 roku na 1000 živě narozených



Tak by to mohlo jít dál, kdyby... a to chce ukázat uvedený graf. Čárkovaná linka extrapolace vývoje má jednu malou chybu — ještě v tomto století by bylo dosaženo prakticky nedosažitelné nulové hranice.

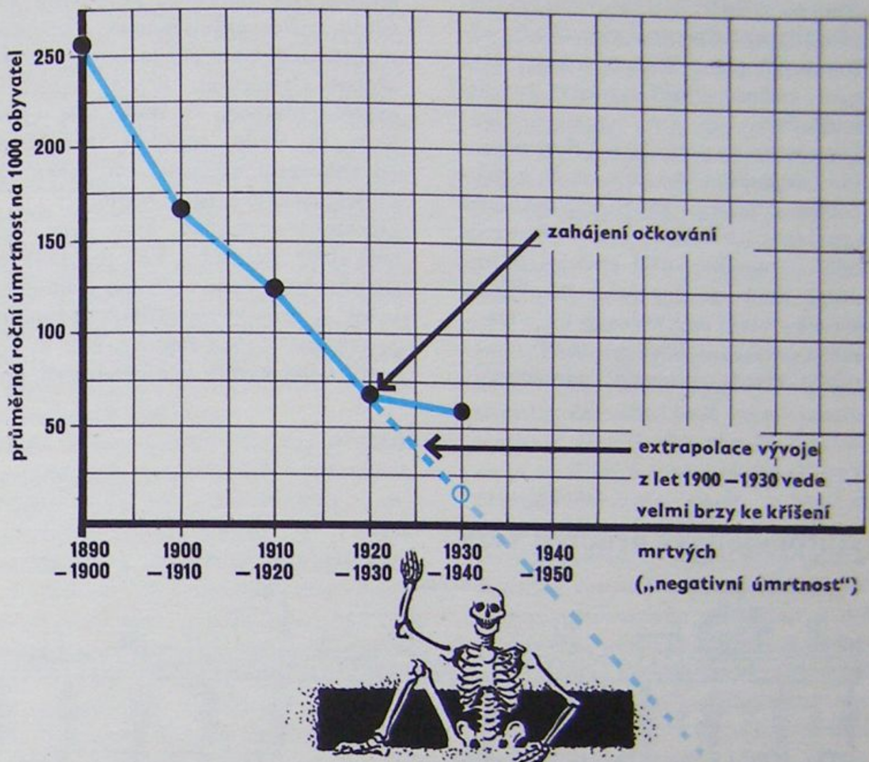
obyvatelstva. Kdyby se podařilo nalézt matematický zákon, musela by být prognóza absolutně přesná.

Lze však také ve větší nebo menší míře uznávat svobodu lidské vůle a jednání, a pak bude prostor pro svobodné individuální nebo kolektivní rozhodování i *prostor pro náhodu vždy větší*, ovšem přesná prognóza bude zásadně nemožná, případně se musí pokládat za případ náhodného výsledku.

Reálné možnosti statistických metod propočtu budoucnosti leží mezi oběma uvedenými extrémy. Používají pozorovaných zákonitostí minulosti, avšak ponechávají náhodě a nepředvídanému určitý prostor — někdy příliš velký, jindy příliš malý, zřídkakdy ve zcela správné míře.

Praktický význam statistické extrapolace nespočívá tedy především v prodlužování časových řad a v extrapolaci vývoje, nýbrž mnohem více v důmyslném *vyhodnocování výběrových souborů* pro budoucí rozhodování. To proto, že např. výsledek ověřování hypotézy „lék A je významně účinnější než lék B" je založen na vzorku, který byl vytvořen v minulosti a nyní vyhodnocen, zahrnuje však i fiktivní základní soubor „minulých i budoucích pacientů s touto nemocí". Praktické vyhodnocení testu hypotézy neříká nic jiného než toto: „Protože v nedávné minulosti byl lék A účinnější než lék B, bude účinnější i v dohledné době."

Kouzlo extrapolovaných časových řad však působí na laika většinou silněji, je-li graficky působivě vyjádřeno.



Protivníci očkování jednou „dokázali", že od zahájení očkování úmrtnost klesala pomaleji než v minulých desetiletích. Extrapolace, tak jak si ji představovali, by byla již ve čtyřicátých letech vedla „k likvidaci smrti" a pak snad až k zmrtvýchvstání zemřelých. — Nerozumná extrapolace vede i v méně průhledných případech k pošetilým závěrům.

7.4 Grafická zobrazení: obraz lže víc než tisíc čísel

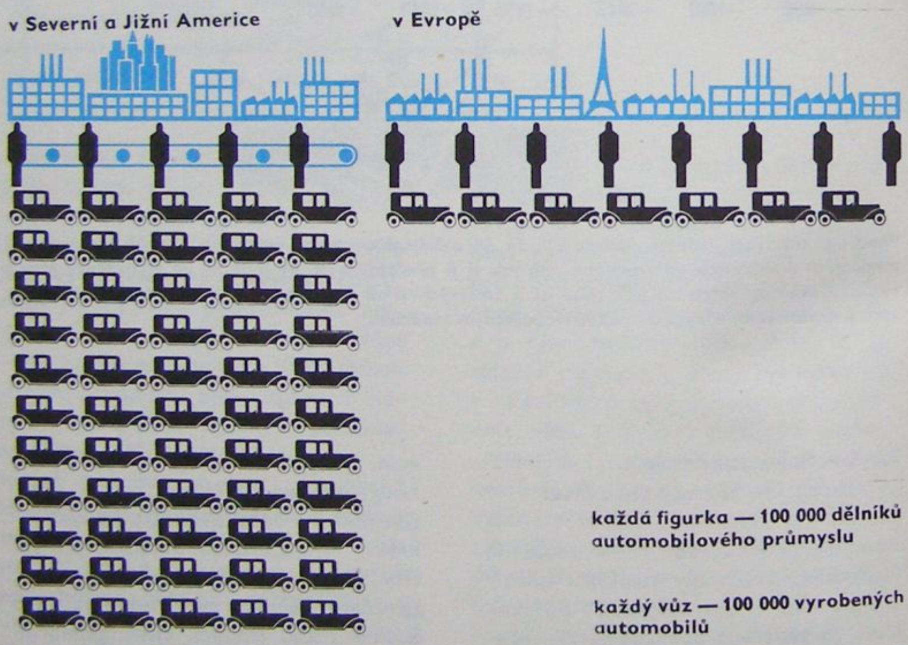
S rozšířením fotografie a zdokonalením reprodukční techniky vzniklo i úsloví: „*Obraz říká víc než tisíc slov.*" Považujeme za zbytečné zabývat se zde rozmanitými psychologickými, teoreticko-informačními a kulturně historickými hledisky poměru mezi slovem a obrazem. Musíme se omezit pouze na ty body, které platí pro poměr mezi číselnou a obrázkovou statistikou.

Prvotním materiálem statistických šetření jsou kvantifikované informace; zpracování tohoto materiálu pak dává číselné řady, sloupce čísel, poměrná čísla, korelace, obory spolehlivosti, pravděpodobnosti, indexy, časové řady... Číslo lze — někdy velmi rozdílně — vyjádřit kresbou. Můžeme rozlišovat mezi *obrazovou statistikou* v užším smyslu a *grafickou statistikou*, i když není možno přesně vymezit hranici mezi nimi.

Pro *obrazovou statistiku* v užším smyslu je rozhodující přání, aby se čísla převedla na optický vjem, jsou to takzvané „cartoons" statistiky a za svůj vznik také především vděčí snahám o poučení. Ke konci dvacátých let našeho století vznikla pod vedením Otto Neuratha *„vídeňská škola obrazové statistiky"*, která sledovala popularizaci statistických informací především použitím stylizovaných figurek. Obrazová statistika je tedy především a hlavně učební pomůckou, která nezůstala ovšem omezena na školy, ale velmi brzy se stala součástí populárního zobrazení statistických materiálů i v hospodářské části novin.

Obrazová statistika říká „víc než tisíc slov" jen tomu, kdo má hluboce zakořeněný odpor k číslům a údajům, avšak je ochoten zamyšleně pozorovat pestré stylizované figurky. Kdo si nemůže nebo nechce nic představit pod

Automobilový průmysl 1929



„Klasická" ukázka obrazové statistiky z doby Otto Neuratha a „vídeňské školy obrazové statistiky". V Evropě tenkrát vyrábělo 800 000 dělníků jen 700 000 automobilů.

čísly nebo křivkami, tomu obrazová statistika představuje mužíčky a žínky, stylizovaná auta, krabičky od cigaret nebo měšce s penězi. *Základní osnovu obrazové statistiky představuje legenda:* „každá figurka — 20 000 lidí" nebo „každý vůz — 100 000 vyrobených automobilů". Přitom nelze zabránit tomu, aby se vyskytly a byly zobrazeny také jiné situace než průměry, jako např. 2,34 dítěte na jednu rodinu, které laik pozoruje s nedůvěrou: jestliže jedna figurka představuje 20 000 lidí a obrázkem se má vyjádřit 128 000 lidí, je to šest celých a čtyři desetiny figurky v obrazové statistice.

Hned se dostaneme i k dalším otázkám populárních statistických vyobrazení. Nejdříve však ještě několik slov k *„legitimnímu statistickému"* grafickému zobrazení. V daném případě jde o to, aby zpracovaný statistický materiál poskytl již na první pohled *dojem*, který je jinak dosažitelný až na základě *srovnání více čísel a údajů*. Tak např. křivka vývoje ukazuje v jediném okamžiku všechno, co by se jinak muselo vyhledat ze dvou desítek čísel několikanásobným srovnáním. Taková vyobrazení lze proto často nalézt i v odborných vědeckých publikacích, nicméně v nich jsou figurky nahrazeny úsečkami a sloupky. Věková pyramida sestavená statistikem se skládá z *úseček* a pyramida sestavená obrazovým statistikem z malých figurek.

Obojí statistické znázornění — obrazová statistika i vědecké statistické grafy — mají své výhody a nevýhody. Brzy se poznalo, že rčení „obraz říká víc než tisíc slov" by často stejně dobře mohlo a mnohdy dokonce muselo znít: *„obraz lže víc než tisíc slov"*, protože obraz má schopnost vyřadit svým emocionálním působením rozumovou výchovu, která je většinou spojena s čtením textu. Obraz může utkvět v paměti a ovlivnit názor na určitý problém víc než deset úvodníků (především tehdy, když se úvodníky nečtou).

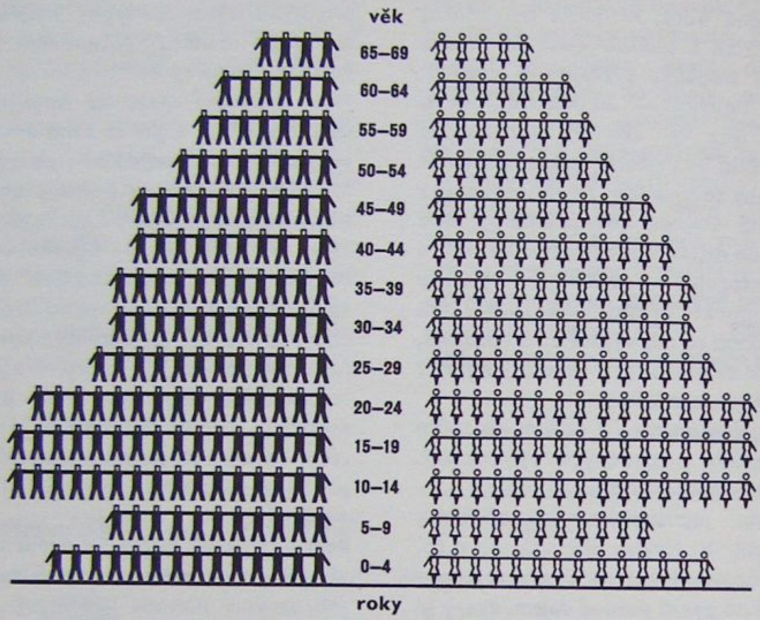
Také statistický obraz lže často víc než tisíc čísel. Proč a jak se toho dosahuje, prozkoumáme poněkud podrobněji.

Krajní případy nejsou zvlášť problematické: tabulka s deseti figurkami a pod ní vysvětlivka: „každá figurka — 5000 dělníků" je pouze ilustrací čísla. Zobrazení regresní křivky nebo operační charakteristiky či sezónně vyrovnaného klouzavého průměru v nějakém statistickém díle také není problematické: čtenář věci rozumí a ví též, co je to dvojitá logaritmická síť nebo pravděpodobnostní papír a proč se používá.

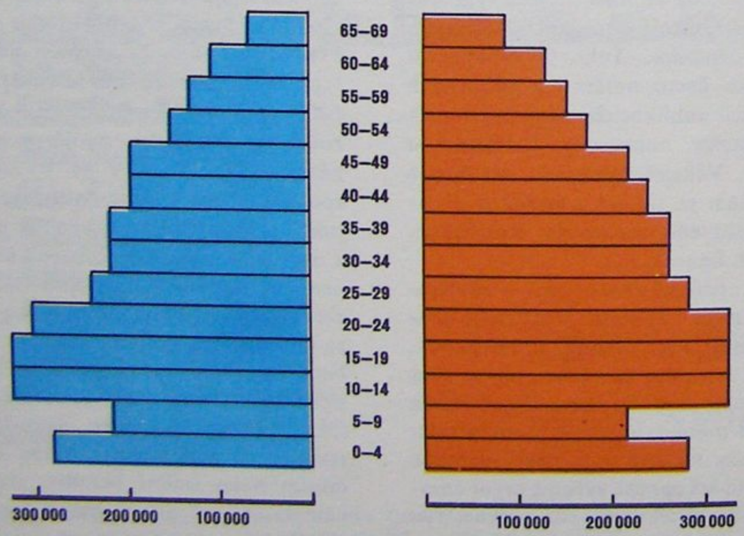
Rozsáhlý okruh zmatků a nedorozumění začíná tam, kde v podstatě legitimní zobrazovací metody vyvolávají u nepozorného nebo nedostatečně informovaného pozorovatele *jiný dojem*, než by vlastně odpovídalo číselnému materiálu. Nejprostším případem je silně stoupající, resp. plochá křivka vývoje. Předpokládejme, že bychom měli pro čtyři různá období tyto hodnoty: 7500, 8000, 8700 a konečně 9400 — např. roční výroba šicích strojů v určitém závodě nebo úspory (v tisících) ve spořitelně apod. Tato čísla zakreslíme jako body do systému souřadnic a jedna z nich (abscisa) tvoří časovou osu. Na osu x zakreslíme čtyři stejně dlouhé intervaly. Na pořadnici (osa y) se nanese měřítko pro měřené hodnoty (vyrobené šicí stroje, případně úspory v tisících apod.).

Okamžitě vzniká otázka, mám-li začít stupnici na dolním konci nulou — zda musím nebo smím? Nikoliv, stupnice nemusí začínat nulou, avšak ten, kdo

Věková pyramida Rakouska v roce 1923



každá figurka — 20 000 lidí



To, co vyjadřuje obrazová statistika v symbolech, mění se v méně pitoreskním zobrazení většinou na sloupky. (Obrázek nahoře: vyjádřeno „obrazovou statistikou" Otto Neuratha.)

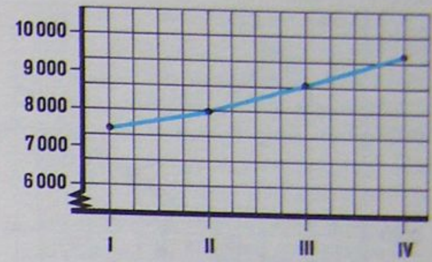
◄

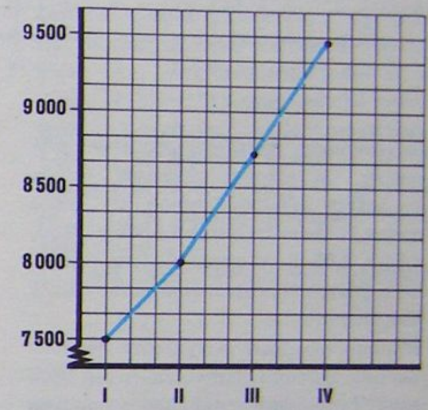
je zcela přesný, naznačí pomocí vroubkované čáry ležící mezi nulou a první hodnotou na stupnici, že se od nuly plynule nenanášelo ve stejném měřítku. Začneme na stupnici u 6000 — jednak pro úsporu místa, jednak také proto, aby křivka nemusela začínat až někde dole. (Pozor: to je už libovolné rozhodnutí, pro které neexistují žádné závazné předpisy!)

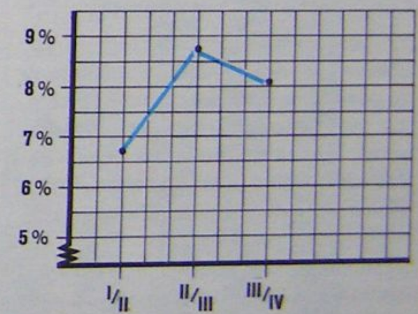
Na základě čeho provedeme volbu *velikosti intervalů* na obou osách? Z estetických nebo typografických důvodů si zvolíme čtvercový diagram — v tomto případě bude vzdálenost pro jednotlivé roky na ose x odpovídat 1000 jednotek na ose y. Také to je libovolné rozhodnutí. Nic a nikdo mi nemůže zakázat, abych neudělal vyobrazení široké a nízké — roční interval třeba 2500 jednotek. V tomto případě bude růst menší, optický dojem působí spíše jako tlumený růst. Se stejným oprávněním mohu meziroční interval upravit na šíři 500 nebo ještě méně jednotek z osy y — a křivka již stoupá strmě do výše.

Tento rozdílný účinek nevzniká jen jako důsledek plochého a širokého nebo úzkého a vysokého zobrazení — jde pouze o *poměr měřítek*. A ještě něco: nejdříve jsme stupnici začali na 6000, nyní zvolíme jako výchozí bod 7500 a dosáhneme optického účinku prudkého růstu od nuly do oblačných výšin.

Z určitého důvodu však chceme, aby tento růst se ukázal menší. K tomu máme mnoho možností. Jednou z nich







Plochý vzestup křivky (nahoře) jako by spíše naznačoval mírný růst — stejné údaje se však také dají vyjádřit optimisticky explozivně (uprostřed). Nesrovnávají-li se absolutní čísla, ale jen přírůstky dosažené v těchto dvou obdobích, je možno vyjádřit růst i pokles (dole).

Dvě rozdílné stupnice v jednom grafu naznačují srovnání, které ve skutečnosti je nesprávné.

je srovnávat navzájem procentní přírůstky, a nikoliv absolutní údaje. V našem příkladu máme tři údaje o růstu během čtyřletého období: 6,65, 8,75 a 8,05 %. V odpovídajícím grafickém zobrazení dostaneme zase smutně klesající křivku.

Druhou možností je vyhledat nějakou blízkou časovou řadu, která *má větší růst*. Předpokládejme, že srovnáme tyto „úspory" s indexem hrubých mezd, který v uvedených čtyřech letech má tyto hodnoty: 100, 108, 118 a 131. Na pravé straně nakreslím druhou stupnici, zakreslím tuto novou křivku jinou barvou do původního diagramu a pak se ukáže, jak růst úspor zaostává za přírůstky mezd. Zda se tím dokazuje nějaký nedostatek, je jiná otázka, rozhodně jsem však dosáhl sledovaného účinku — znevážit křivku růstu úspor.

Třetí možností je použít k *srovnání údajů o absolutním růstu* většího podniku — je tu např. koncern, který v uvedených čtyřech letech vyrobil 16 000, 16 600, 17 500 a 19 000 šicích strojů, čímž běžně zvyšoval svůj absolutní předstih.

Čtvrtou možností je provést *srovnání s malým předstihem*, který však dosáhl vyššího relativního růstu — např. ze 3000 přes 3400 a 3750 na 4100. V tomto případě porovnám přírůstky zase v procentech nebo nanesu absolutní údaje na logaritmický papír, čímž se dosáhne stejného výsledku.

Takových srovnání jsme se již krátce dotkli, když jsme různými grafy vyjádřili relativní a absolutní růst exportu automobilů tří značek do USA (srv. oddíl 4.4).

V mnoha případech vůbec nemusí jít o klamné manévry, někdy se vtěsnají dvě křivky do stejného grafu jen pro nedostatek místa. Nebo se šetří místem (úmyslně nebo neúmyslně) tak, že se počátek určité časové řady posune. Průběh kursů např. od roku 1965 do roku 1970 na většině mezinárodních burz nevypadá tak špatně — co se však stane, když se počátek křivky položí do roku 1962? (Viz obr. na str. 236.)

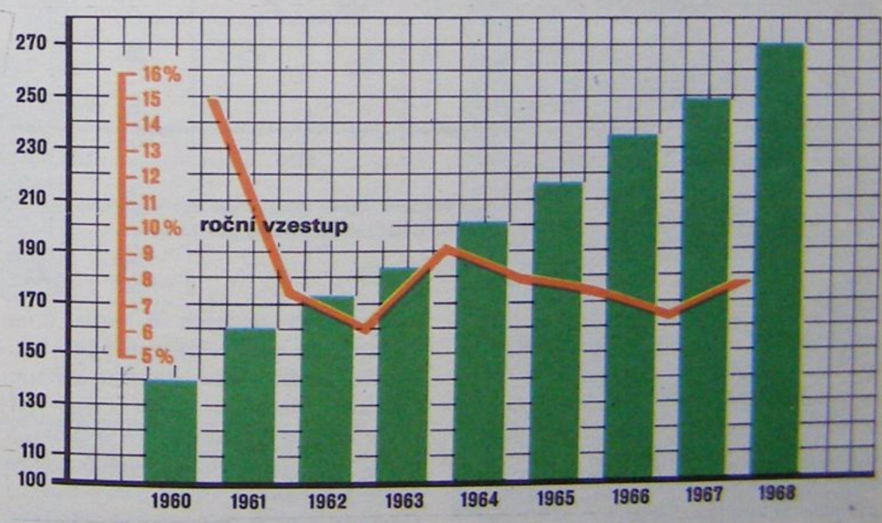
A co by se stalo, kdyby se k propočtu nepoužilo indexů absolutních čísel, ale kdyby se počítalo — např. s použitím cenového indexu — jen na základě indexu kupní síly? Jestliže jsem v roce 1962 investoval 1000 DM do určité akcie, kterou jsem v roce 1970 musel prodat

při indexu 600, pak již těchto 600 DM, které dostanu, není týchž 600 DM z roku 1962 (pro srovnání jejich kupní síly musím počítat i s poklesem hodnoty peněz). Prakticky to vypadá takto: Index spotřebních cen (1962 = 100) je pro čtyřčlennou rodinu v roce 1970 asi 124. Jinak řečeno: 100 DM v roce 1970 odpovídalo $\frac{100}{124} = 81$ DM v roce 1962. Na základě zvolené formulace jsem dostal buď jen 600 DM za 1240, nebo $6 \cdot 81 = 486$ DM za 1000. Jestliže tento výsledek vyjádřím graficky, bude se ztráta jevit ještě větší. V jiných případech, např. při průměrném hrubém přírůstku 3 %, mohu dostat celkem snadno klesající křivku, jestliže počítám „reálný" místo „nominálního" růstu.

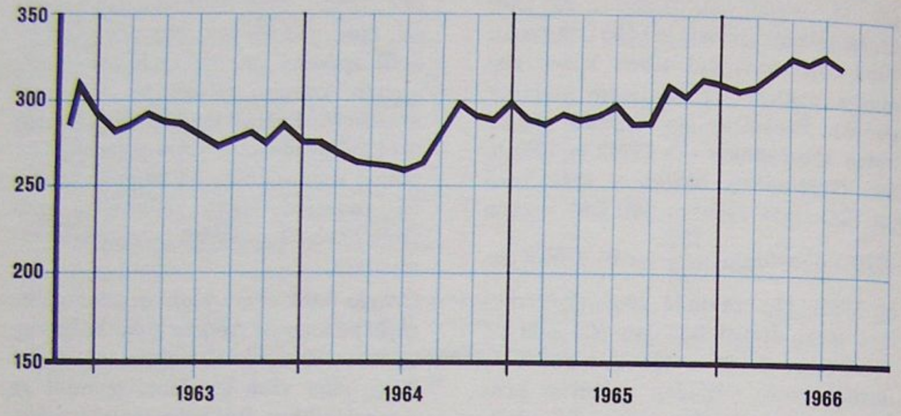
Přestože některá zobrazení vypadají tak „lživě", byla všechna zcela korektní. Platí pro ně, tak jako pro všechny další způsoby statistických zobrazení, zásada: *svévolné výklady jen na vlastní nebezpečí!* Pokud je pod diagramem správně uvedeno, o které jednotky jde, pokud jsou hodnoty na stupnici správně zaneseny, nelze mluvit o podvodu nebo lži, nanejvýš o pokusu o manipulaci.

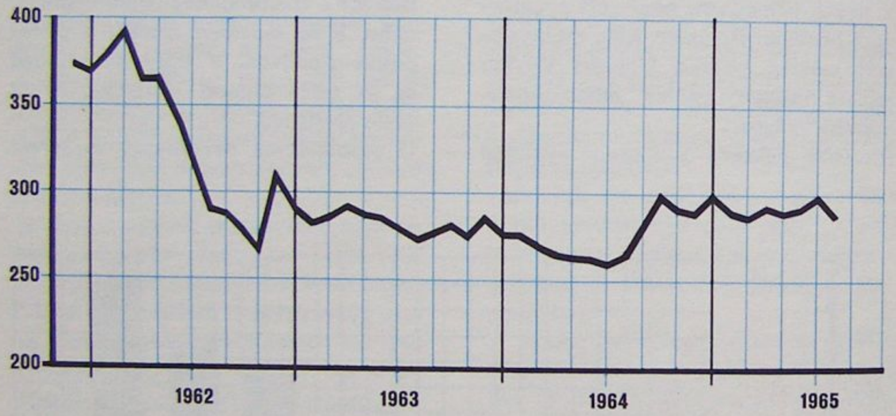
Docela jinak tomu však je, když náhle *chybí na stupnici měřítko* nebo když jsou na ní sice správně zanesena nějaká čísla, není však uvedeno, jedná-li se o marky, libry, liry nebo indické rupie. Buď jde o trestuhodnou lehkomyslnost, nebo, a to mnohem častěji, o přímý pokus o podvod. V takovém případě se již nelze pomocí statistiky, nýbrž zcela *prostě lže*.

U statistických obrazových vyjádření



Jestliže chci postavit do nepříznivého světla z roku na rok neustále stoupající růst, postačí vzájemně porovnat roční přírůstky. V našem příkladu dostáváme po růstu o 15 % prudký pokles a po krátkém znovuoživení o 9 % opět „pokles".





Kursovní diagram této akcie vykazuje klidnou a dokonce mírně vzestupnou tendenci. Uvede-li se však rok 1962, vykazuje diagram prudký a nikoli ani přibližně vyrovnaný pokles kursu.

je obtížné rozhodnout, *kde končí naivita a kde začíná podvádění.* Víme již, že obrázková statistika pracuje převážně se symbolickými figurkami, které jsou často uspořádány ve *skupinách po pěti* (podle zkušeností může lidské oko na první pohled zřetelně rozeznat nejvýš pět jednotek) a jsou seřazeny vedle sebe nebo nad sebou.

Samozřejmě je možno uvažovat také takto: Proč mám kreslit nejdřív 16 skupin po pěti zelených a vedle toho 8 skupin po pěti červených figurkách, jestliže totéž mohu vyjádřit jedinou figurkou? Nepotřebuji nic jiného než jednu zelenou figurku, která je dvakrát tak velká ako vedle stojící červená. Zcela jednc duché, že?

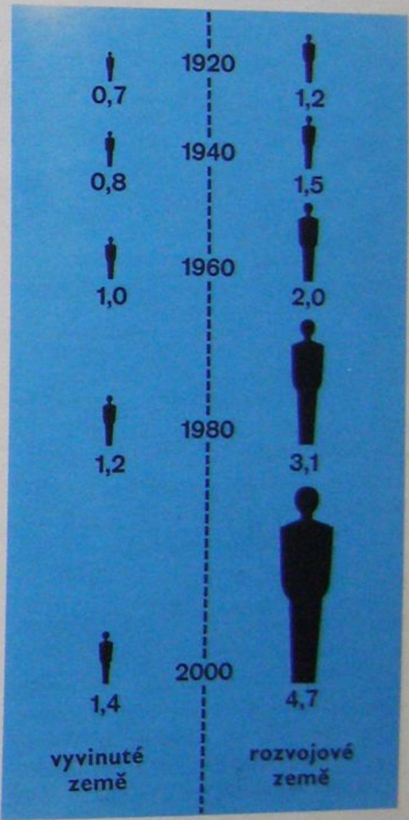
Otázkou však je, jak velké je „dvakrát tak velké“? Červenou figurku mohu nakreslit vysokou 3 cm a vedle toho zelenou vysokou 6 cm — jaký je však

účinek? Protože nejde o čáry (nebo případně o sloupky), ale o plochy, které symbolizují tři dimenze lidského těla, dvakrát tak vysoká zelená figurka nevypadá „dvakrát tak velká“, nýbrž čtyřikrát až pětkrát tak velká. Má tedy kreslíř přesně změřit plochu figurky a soustředit se na plošný poměr 2 : 1? Nyní však přechází účinek do jiného extrému: figurka plošně dvakrát tak velká vypadá jen „poněkud větší“.

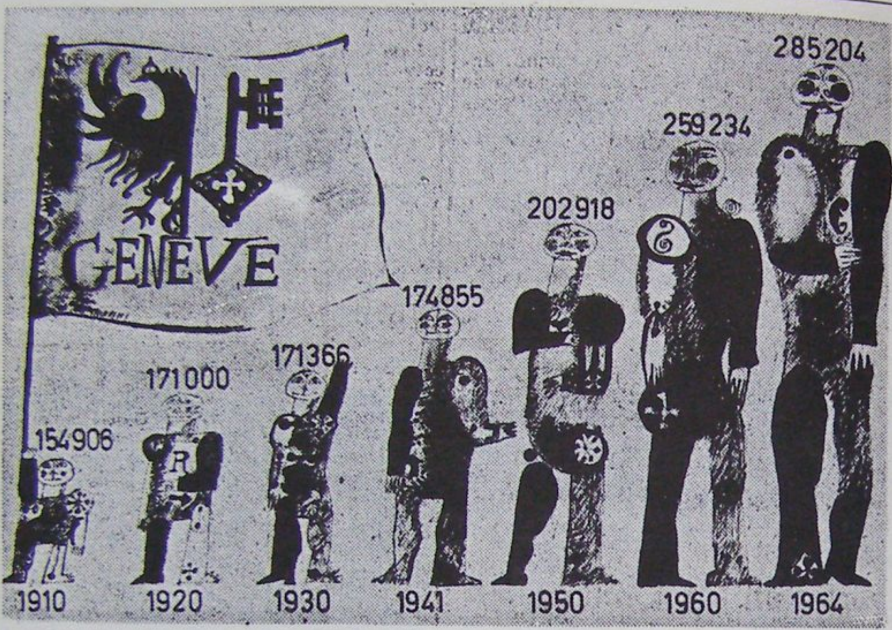
Problém však nevzniká jenom u statistických obrázkových figurek, ale i při srovnávání kruhů, které se často s oblibou používají k zobrazení mincí. Jak postupovat, chci-li zobrazit pokles kupní síly nějaké měny na polovinu její dřívější hodnoty? Poměr průměrů 1 : 2 zname-

Rok	Náklady na výzkum a vývoj v mil. US dol.	Počet nových registraci	
		jednotné účinné substance včetně biologicky aktivních ubstancí	zdvojené a kombinované preparáty; terapeutické formy
59	197	63	356
61	238	41	327
63	282	18	235
65	351	23	111
67	462	25	71
68	495	14	97
69	566	11	65

Statistická zobrazení mají přece jen svůj význam, jsou-li doplněna příslušnými údaji. Srovnání znaků dolarů nebo tablet (podle zobrazení CIBA), stejně jako figurek, které znázorňují růst obyvatelstva (podle „Sonntagsblattu“), dovoluje jen tušit symbolizované velikostní poměry; v případě figurek kresba dokonce klame, protože číselné poměry jsou vyjádřeny pouze jejich výškou: „málo rozvi-



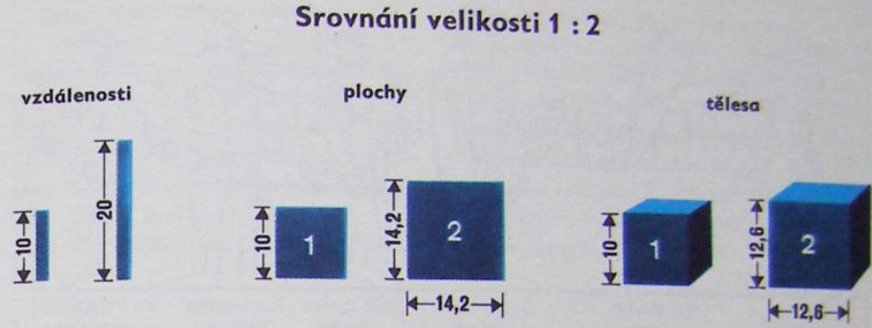
nutý“ se v roce 2000 zdá skoro desetkrát větší než „vyvinutý“ — ve skutečnosti je sotva třikrát větší.



Vyobrazení z „Tribune de Genève“ představuje drobné umělecké dílo, nikoliv však obrazovou statistiku. Figura vlevo by měla být vyšší, než je polovina velikosti obra vpravo; mezi symbolickými figurami roku 1920 a roku 1941 by neměl být téměř žádný rozdíl.

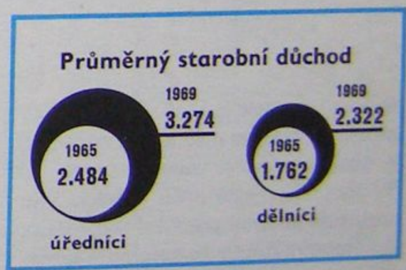
ná poměr ploch 1 : 4! Poměru ploch 1 : 2 odpovídá poměr průměrů $1 : \sqrt{2}$. Poměrně značný pokles kupní síly ve výši 30 %, zobrazený plošně, nebude zvlášť nápadný. (Viz obr. na str. 239.) Ještě obtížnější je, když jsou vedle sebe měšce peněz a navíc je v rukou drží kostýmované figury. Předpokládejme, že by se pomocí měšců mělo porovnat 900 miliónů, 400 miliónů a 100 miliónů. V plošném vyjádření by měly rozměry (protože jsou zhruba pravoúhlé) asi 9,4 a 1 cm². Čím přirozeněji se však kreslí, tím více se ukazuje potřebné vyjádřit třetí rozměr, a pak se již dosahuje poměru 81 : 16 : 1. Vzhledem k tomu je stejně nutné uvést u nich čísla — *a nakonec obrázek víc mýlí, než vysvětluje*. To platí bohužel pro velmi četná obrázková statistická vyjádření (statistická zobrazení). Usiluje se o originální účinky, hledají se nezvyklá působivá zobrazení — a přitom se ztrácí statistická výpověď a zůstane jen podivná zvláštnost, s jejímž pochopením má zájemce mnohem více práce, než kdyby dostal čtyři střízlivé údaje. Ale to je smutný osud čtenáře, který je většinou pokládán za naivnějšího, než ve skutečnosti je, a který je nakonec takto skutečně ohlupován.

Na druhé straně však existují také rafinovaná grafická znázornění, skutečná mistrovská díla statistické vynaléza-

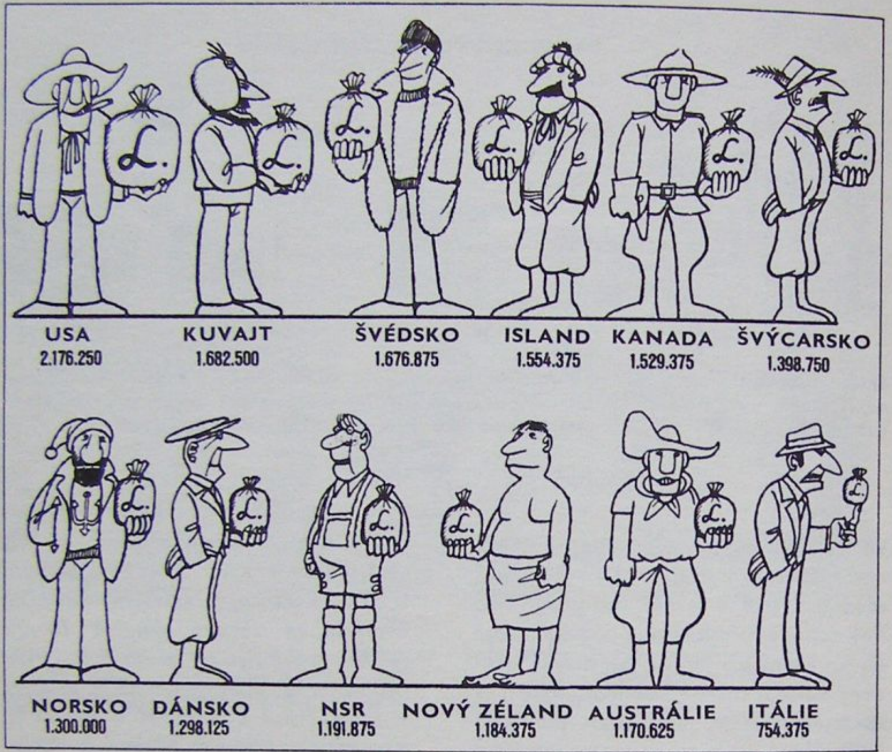


Srovnání velikostí pomocí přímek a sloupků nečiní zvláštní potíže. Plochy a zvláště tělesa mohou naproti tomu velmi klamat. Tak např. krychle s délkou hrany 10 jednotek má jen poloviční obsah než krychle s délkou hrany 12,6. Dvojnásobná délka hrany by dokonce dala osminásobný obsah.

vosti, která nakonec přece jen a mnohdy i na vyšší úrovni donutí zájemce pečlivě studovat graf a intenzívně přemýšlet. Tak tomu je především v případě, když na souřadnicích nejsou jen prostě uvedeny roky a měrné hodnoty, nýbrž již *zpracované údaje*, jako např. poměrná čísla. Nebo — a to se vyskytuje ještě častěji — když je v souřadnicovém systému uvedeno více *navzájem se doplňujících údajů*. Uvádíme zde graf Spolkového statistického úřadu NSR, který v souřadnicovém systému, na první pohled zcela jednoduchém (osa x — věkové skupiny; osa y — procentní stupnice) uvádí muže výdělečně činné podle tří hospodářských sektorů, a přitom zachycuje pro každou věkovou skupinu vývoj za posledních deset let na podkladě tří zjišťování. (Obr. str. 241.) Malá poznámka na okraj — doslova „na okraji“, totiž na pravém okraji tohoto grafu. Není to ostuda, když procento osob činných v zemědělství a lesnictví tak prudce roste v nejvyšších věkových skupinách?! Kolik starých lidí si ještě musí namáhavě vydělávat na živobytí? Snad je jich však pouze několik stovek. Graf nám totiž říká jen to, že ze všech výdělečně činných ve věkové skupině 70—75 (a těch může být na celém území NSR 200 nebo 200 000) celých 42 % pracuje v zemědělství a lesnictví a jen 21 % v ostatních výrobních odvětvích. Vzhledem k tomu, že ve výrobních odvětvích



Tento graf byl „střelen od boku“. Plochy kruhů se k sobě vůbec nemají jako 2484 : 3274 nebo jako 1762 : 2322. Plocha černých kruhů je téměř dvojnásobná. (Údaje jsou uvedeny v rakouských šilincích.)

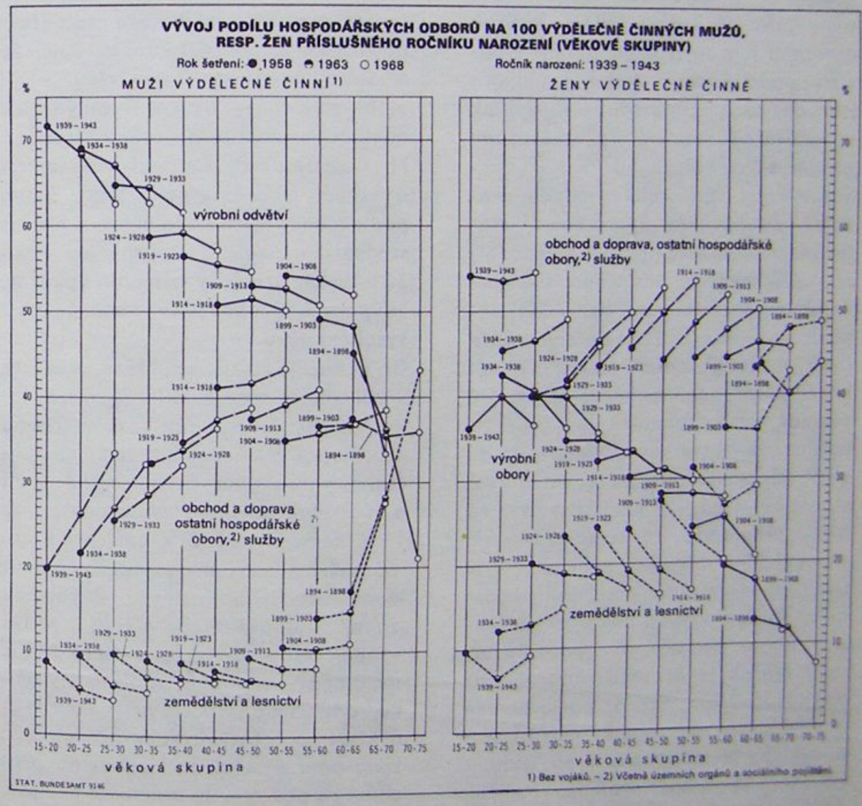


Obrázek z italského časopisu „Domenica del Corriere". Velikosti poměrů měšců peněz neposkytují správnou představu o číselných poměrech, které vyjadřují. Měšec Američana (vlevo nahoře) vypadá nepochybně deset až dvacetkrát větší než měšec Itala (vpravo dole). Skutečný číselný poměr je pouze 3 : 1, totiž 2176 : 754. Porovnání tří dimenzionálních objektů (v dvoudimenzionálním zobrazení) většinou víc klame, než objasňuje.

zdaleka největší počet zaměstnanců odchází po dosažení 65 let do důchodu, je absolutní počet zaměstnaných po 70. roce věku jistě příliš vysoký. Kdo prosazuje penzijní pojištění zemědělců, s nadšením uvítá tento absolutně správný a nanejvýš zajímavý graf.
Nikdy bychom neskončili, kdybychom chtěli uvést všechny možnosti zobrazování a klamání pomocí obrazové statistiky a statistických grafů. Na závěr se spokojíme prostou, ale důležitou radou: Neusuzujte jen podle obrázku. Vždyť je třeba všimnout si též čísel a stupnic! Křivka klame často víc než tisíc slov.
Je samozřejmé, že je možný také obrácený postup: několik slov může postavit bezpočet jasných statistických údajů do úplně jiného světla. Tímto problémem, interpretací, resp. výkladem se budeme zabývat v dalším oddílu.

7.5 Umění výkladu. Statistická interpretace

„Jeden duchaplný Francouz kdysi tvrdil: »Statistika je umění dokázat správnými čísly něco nesprávného.« Chtěl tím říci, že statistické údaje jsou sice vcelku správné, ale že jsou nesprávně vykládány. Wallis a Roberts naproti tomu ve své knize *Statistické metody* prohlašují: »To nejdůležitější, co je třeba o interpretaci říci, je, že statistické údaje se musí interpretovat.« Tím vzniká obtížná volba. Na jedné straně je výklad nezbytný, na druhé straně je však nebezpečný." Citovali jsme Hanse Kellerera, protože krátce ukázal, jak důmyslná statistická vyhodnocování údajů musí proplouvat mezi Scyllou nekomentovaných číselných řad a Cha-



Toto neobvyklé zobrazení představuje graf pro labužníky. Podrobnosti se musí pečlivě studovat, avšak již na první pohled je mimo jiné zřejmé, jak klesá podíl zaměstnaných ve výrobních oborech a současně vzrůstá podíl skupiny „obchod, doprava, služby atd."

rybdou hotových závěrů dodaných „až do domu". Čísla sama čtenáře hned od počátku buď odradí, nebo svádějí k vlastním závěrům, které mohou být ještě mylnější. Zpracovaná vysvětlení výsledků mohou naproti tomu velmi snadno vyjadřovat subjektivní názor, „*statistickou ideologii*", a tím ještě více klamat. V čistě vědeckých pracích je nebezpečí takového klamání poměrně malé, ačkoliv nezřídka bývají s použitím mnoha statistických metod složitým způsobem zpracovány nebo i vykouzleny z pochybného materiálu působivé závěry. Takové práce se však dostanou k průměrnému čtenáři v původní podobě jen zřídkakdy a mimoto jsou obvykle málo srozumitelné. Laik se převážně setkává se *statistikami z druhé a třetí ruky*.

Statistika z třetí ruky většinou znamená, že původně vypracovaný statistický materiál se nejdříve zjednoduší a pak komentuje. Při zjednodušování se škrtne to, čím nemá být čtenář zatěžován. Je možné, že se do komentáře (který může být sám z třetí ruky) vloudí nezamýšlené nebo záměrně nesprávné výklady všeho druhu.

Reichmann ve své knize „*Use and Abuse of Statistics*" (Používání a zneužívání statistiky) říká: „Denní tisk vybírá a zdůrazňuje vždy to, co se zdá zvlášť pozoruhodné, na úkor tabulek a poznámek, které často současně neotiskuje. To není kritika tisku, jehož úkolem je na skutečnosti upozorňovat, nikoliv však je do všech podrobností popisovat. Žádný čtenář novin nechce dlouhé statistické tabulky místo svého obvyklého čtení, ale nesmí podlehnout pokušení brát doslovně znázornění, které se stalo povrchním... Mnohé zprávy jsou ovšem dokonce výslovně nesprávnými výklady, dovozují nesprávné závěry a tím neprávem kazí pověst statistiky."

Čteme-li slova „statistika dokazuje, že...", můžeme být zpravidla přesvědčeni, že statistika právě to vůbec nedokazuje. „Statistika dokazuje, že *v Římě se žije déle*", znamená toto: V Římě je úmrtnost na 1000 obyvatel menší než např. ve Florencii, protože mnoho mladých lidí hledá práci v hlavním městě a zakládá tam rodinu. „Statistika dokazuje, že *obézní lidé jsou častěji nemocní*" — což může být pravda, je to však napadnutelné, jestliže se tohoto tvrzení používá jako důkazu pro odtučňovací kůru. „Statistika dokazuje, že *studenti umírají v mladém věku!*" Tak mohl znít tučný titulek v bulvárních novinách roku 1835; tehdy statistik H. C. Lombard zkoumal 8488 úmrtí a všímal si průměrného věku úmrtí podle jednotlivých povolání — a hrůza: studenti umírali průměrně ve věku 20 let! (Obráceně by se mohlo dokázat, že papežové umírají v nadprůměrně vysokém věku.)

Co dokazuje statistika, která uvádí, že v uplynulém roce bylo 32 % všech Rakušanů postiženo nemocí (z toho třetina s pobytem v nemocnici), zatímco současně podskupina „samostatně činní a svobodná povolání" vykazuje pouze 13 % nemocných (s 3 % pacientů v nemocnicích)? Jasný případ: to jsou bohatí lenoši, kteří mají pořád dovolenou a mají se dobře. Nebo nikoli, a právě naopak: jsou to chudáci, které žádné sociální pojištění nepřivádí na myšlenku odpočívat ve stavu nemocných a kteří se zaťatými zuby pracují dál, když páni zaměstnanci chodí na procházky s potvrzením o nemoci v kapse. Statistika „dokazuje" obojí.

„*Statistika dokazuje, že s bezpečnostními pásy je naděje na přežití 20 000 ku 1*" — načež při podrobnějším posouzení se ukáže: Odborníci jsou toho názoru, že pásy mohou být v jediném případě z 20 000 horší než vůbec žádné. S „nadějí na přežití 20 000 ku 1" nemá tento odhad (neboť nic jiného to není) ani nejmenší souvislost. „Statistika dokazuje, že *Eisenstadt je nejbohatší rakouské město*" — to znamená, že v městečku, které má několik tisíc obyvatel a je správním centrem rakouské východní provincie, Burgenlandu, jsou registrována všechna služební vozidla zemské vlády a většiny zemských úřadů, což dává „na jednoho obyvatele Eisenstadtu" takový počet automobilů, jaký je jen v Americe.

„*Statistika dokazuje: řítí se lavina lékařů.*" Znamená to, že stoupl počet posluchačů zapsaných na lékařských fakultách. Podkladem pro takové tvrzení může být jen stoupající počet absolventů, nikoliv však počet studentů (mezi nimiž mohou být ve větším počtu třeba dívky, které studium nedokončí). „Statistika dokazuje: *významné osobnosti žijí déle.*" Znamená to, že osoby uvedené v publikaci „Who's Who" z roku 1950/1951 měly nižší úmrtnost než průměr amerických úředníků, podnikatelů, vyšších úředníků a příslušníků svobodných povolání. Je otázka, zda bylo provedeno specifické pozorování podle věkových skupin. Jestliže nikoli, pak tu máme „papežský fenomén": kdo se dostane do Who's Who, udělal již kariéru a je starší než průměr srovnatelné skupiny zaměřené jen na určitá povolání. Soudci Nejvyššího soudního dvora nemohou zemřít ve 45 letech, protože v tomto věku takový úřad zastávat nemohou. „Soudci" umírají tedy v průměru dříve než „významní soudci". Formulaci „významné osobnosti žijí déle" je správné číst obráceně: *kdo žije déle, stane se spíše významnou osobností.* Kdo zemře ve 30 letech, nestane se spolkovým prezidentem. „Statistika dokazuje: *cizinci způsobují zmatek v dopravě.*" Znamená to, že během sezóny cizineckého ruchu se na 70 % dopravních nehod v Tyrolsku podílela zahraniční vozidla. Bylo by ale zcela dobře možné, že podíl zahraničních vozidel v celkovém provozu byl ještě větší — pak by ovšem cizinci jezdili bezpečněji než domácí.

„Statistika dokazuje: *opilci jsou za volantem jistější než jako chodci.*" Znamená to, že počet usmrcených opilých chodců je absolutně větší než usmrcených opilých řidičů.

Po tomto malém výběru charakteristických nesprávných výkladů se však také podívejme na několik příkladů, kde výpovědi a úkoly statistiky nebyly vůbec pochopeny. Pěkný příklad poskytuje tato polemika z jednoho časopisu nájemníků: „Statistika lže! Jestli si někdo myslí, že do statistiky patří také jednopokojové byty bez tekoucí vody..., pak si dovolujeme říci, že se započtením takových obydlí může snad statistika souhlasit, že však je ostuda, že takové byty v roce 1968 stále ještě existují." Tedy, přesné šetření o výši nájemného musí přihlížet ke všem typům bytů podle jejich počtu. Statistik nemůže vybírat a říkat: ne, malé, staré byty neuvažuji, protože tam by vlastně neměl nikdo bydlet. (Stejně dobře by bylo možno argumentovat i takto: ne, velké luxusní byty neuvažujeme, protože nikdo by neměl bydlet tak přepychově.)

Ve stejném smyslu píše jeden autor, když naráží na problém automatizace: „V tomto případě je ještě těžší sledovat hru čísel statistiků. Většina z nich se totiž shoduje v tom, že celkové obyvatelstvo poroste z různých důvodů rychleji než počet pracujících. Jestliže se z toho usuzuje, že v budoucnu bude k dispozici poměrně méně lidí pro zabezpečení potřeby zboží pro stále větší počet spotřebitelů..., zůstávají nám statistici dlužni odpověď na otázku, odkud získá kupní sílu stále rostoucí přebytek spotřebitelů." Aby bylo jasno: taková otázka se statistikům vůbec klást nemůže. Extrapolovali pouze z vývoje posledních desítiletí relativní úbytek pracujícího obyvatelstva — a basta! Sociální a ekonomické důsledky musí promyslet politici a hospodářští experti.

Jak špatně lze volky či nevolky chápat teorii výběrových souborů a počtu pravděpodobnosti, ukázala např. v Rakousku v létě 1970 činnost „komise pro prověřování výsad". Tato komise měla mimo jiné zkoumat, zda a jak by měly být zdaněny platy politiků. V průběhu prací komise rozeslala asi stovce politiků dotazník se žádostí, aby uvedli výši svých příjmů, služebních diet a služebních výdajů. Část těchto dotazníků se skutečně vrátila vyplněna a hle: politikové, kteří odpověděli, měli skutečně relativně nízké jiné příjmy a mimořádně vysoké služební výdaje (navzdory volné jízdence na železnici byly např. uváděny velmi vysoké výdaje na cestovné). Jeden novinář se zeptal předsedy komise, zda snad tyto dotazníky poněkud nezkreslují skutečnou situaci. Dotázaný se mírně usmál a pravil: „Nikoliv, statistická pravděpodobnost přece vyřadí extrémy a poskytuje správný obraz..."

„Statistická pravděpodobnost" může jedině zjistit střední hodnoty uvedených příjmů a doložených výdajů na základě dotazníků, které byly v průběhu akce zodpovězeny. Anketní šetření však jednak nebývá reprezentativní, jednak — z důvodu politické zdvořilosti — si nelze ověřit správnost zjištěných dílčích výsledků. Zkrátka: už samo zjišťování poskytlo nanejvýš pochybný prvotní materiál, a proto všechny matematické a statistické operace s těmito údaji mají, mírně řečeno, omezenou cenu.

Na závěr a jako přechod k další kapitole ještě několik slov o velmi častém zdroji nesprávných interpretací: Vypovídací schopnost malého šetření se cestou od statistika k poslednímu spotřebiteli stále více přeceňuje a zdůrazňuje. Co bylo původně uváděno s pravděpodobností jistoty a intervalem spolehlivosti jako zcela pravděpodobné, stává se nakonec zcela jisté, matematicky a statisticky dokázanou skutečností. Jediný vzorek se dotkl nulové hypotézy a titulek již hlásá: „Statistika dokazuje...!"

8 Jednotlivec a statistika

8.1 Statistika všedního dne

Zdrženlivost a nedůvěra, s nimiž „průměrný občan" přistupuje ke statistice, spočívá ve značné míře v neuvědomělém, zčásti však zcela správném přesvědčení, že se statistika vůbec nestará o jednotlivce, že předmětem jejího zájmu je pouze kolektiv, „obyvatelstvo", nanejvýš ještě výběrový soubor (rovněž sestavený z většího počtu). Víme, že o popisné statistice stále platí, že se zabývá jen hromadnými jevy.

Zcela jasně se však ukázalo již dříve, zvláště když jsme mluvili o technice zjišťování a anketních šetřeních, že statistika jednotlivcem nepohrdá. Jednotlivec se dokonce stává představitelem velkého počtu, je pro některý znak zachycen jako zástupce mnoha jiných. Jako velmi komplikovaný a jedinečný „svazek znaků" si však zachovává svou individualitu neomezenu, i když jednotlivé znaky jsou statisticky zjišťovány a vyhodnocovány.

Statistika nejenže jednotlivce tímto způsobem nepoškozuje, ale každý z ní může mít prospěch, i když v okamžiku uvažování nebo rozhodování často vůbec neví, že se zabývá statistikou. Každý z nás se totiž vždy znovu ocitá v situacích, které vyžadují posouzení a pak většinou jednání. Na čem se však má zakládat úsudek i jednání, ne-li na dobře uspořádaných znalostech o dosavadních událostech? Již v oddíle o extrapolaci na základě statistik jsme poukázali na to, že jsou v podstatě jen dva způsoby chování, jestliže se musí učinit rozhodnutí: *rozumové*, které vychází z informací, a *iracionální*, které vychází z intuice. Tyto dva způsoby jsou v různé míře obsaženy ve svých nesčíslných kombinacích. Většinou lze říci, že i rozhodnutí zdánlivě vycházející jen z intuice spočívá na rychlém propojení informací, které jsou třeba jen z poloviny vědomé, takže koneckonců rozhodnutí nevychází z intuice: „subjektivní pravděpodobnost" nebude téměř nikdy přímo odporovat statistické pravděpodobnosti.

Proč se však přesto laik téměř instinktivně statistice brání? Pokládá ji za *speciální matematický vědní obor*, kterému nerozumí a který se ho jen málo týká. K první z těchto bezděčných námitek je třeba říci toto: je samozřejmé, že je nutno vědeckou statistiku přenechat odborníkům, tak jako se jim přenechává srdeční chirurgie, policejní správa, daňová politika a jaderná fyzika. *K všeobecným znalostem* však dnes patří znalost alespoň základních pojmů z ekonomie, fyziky, medicíny a matematiky, z politiky a *statistiky*. Tyto požadavky nevyplývají z nějakého ideálu a teoretických úvah o politickém vzdělání, ale vycházejí docela prostě z praktických potřeb denního života.

Tím jsme se dostali k druhé námitce — *že totiž statistika může jen málo poskytnout jednotlivci*. Několik stránek nazpět — ale i v úvodní kapitole této knihy —

Výběr na ochutnání není často nic jiného než vzorek. Třešně nebo hroznové víno se ochutnávají proto, aby se mohlo posoudit, zda zboží odpovídá požadavkům.

jsme poukázali na základní význam statistiky pro jakékoliv rozhodování. Ještě jednou krátce a heslovitě shrneme, kde jsme se při řešení praktických otázek denního života setkali dosud se statistikou.

Každý, kdo se aktivně účastní hospodářského života, se nějakým způsobem zabývá výrobní kontrolou, vytvářením výběrových souborů nebo srovnáváním obratu. Každý čtenář novin, každý posluchač rozhlasu a televizní divák slyší a vidí statistický materiál, ale zároveň i propagaci (hospodářskou a politickou), která se odvolává na statistiku. Lidé zabývající se průzkumem potřeb a veřejného mínění chodí od domu k domu, a buď jsme sami dotazováni, nebo čteme výsledky zjištění aktuálních šetření. Život nás všech a především naše životní úroveň závisí rozhodujícím způsobem na *indexech*, především cenových a mzdových. Kdo se zabývá vědeckým výzkumem, neobejde se bez statistických metod — a ve společenských vědách patří mezi nejdůležitější pomůcky poznání vůbec. Elementární výběr se objevuje na trhu s ovocem, jestliže váhající hospodyňka vybírá třešni na ochutnání nebo prodavačka rozřízne jablko, aby dokázala, jak je pěkné a zralé. Statistické zkušenosti se dokonce zrcadlí i v příslovích a aforismech životních pravidel. Francouzský statistik Vessereau hned na prvních stránkách svého tenkého svazku „*La Statistique*" trefně poznamenává, že každý, kdo v jarním jitru uvažuje o tom, zda si má přece jen vzít plášť, a hlavou mu přitom proletí přísloví „En avril ne te découvre pas d'un fil" (asi: „V dubnu není žádný kus oděvu zbytečný"), používá pro své rozhodnutí koneckonců ve stejné míře statistiku jako ekonom-plánovač.

Není úkolem této knihy vypočítávat

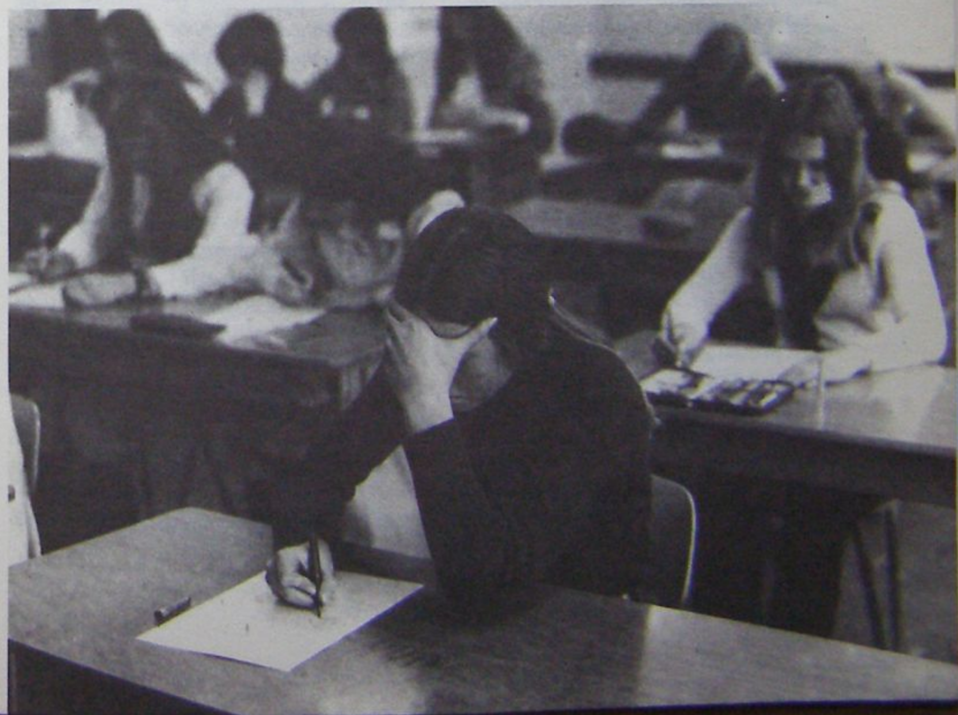
všechny oblasti života, kde má statistika nějakou úlohu. Vzhledem k tomu, že nelze dohlédnout konce, bylo by to velmi únavné. Již jsme ukázali nejdůležitější příklady všedního dne, ve kterých má statistika zvláštní význam, a v další části tohoto oddílu uvedeme ještě několik dalších příkladů.

V této souvislosti se seznámíme poněkud blíže s „klasickým" oborem statistiky — *s úřední statistikou, popisnou statistikou masových jevů*, která se ovšem dnes již vůbec neomezuje na obtížné zjišťování základních souborů, ale s velkým úspěchem používá nejmodernějších metod teorie výběru.

Část této kapitoly dále věnujeme *pojišťovnictví*, které je vedle počtu pravděpodobnosti a statistiky obyvatelstva třetím důležitým východiskem klasické statistiky a které má dnes pro jednotlivého občana větší význam než kdykoliv jindy. Pojištění mimoto s elementární jasností ukazuje úzké souvislosti mezi osudem jednotlivce, náhodou a vypočitatelnou četností jevů v kolektivu.

V jednom oddílu této kapitoly konečně nahlédneme do oboru vyšších statistických metod, který začíná mít v době hromadných sdělovacích prostředků převládající úlohu; je to tzv. *teorie hromadné obsluhy*, jinými slovy „*teorie front*", resp. „*teorie přepážek*".

Také zkoušky a školní úlohy jsou vzorkem — vzorkem celkových znalostí jednotlivých žáků.



8.2 Mikrocensus a hospodářská statistika

Číselné zachycení hromadných jevů ve velkých společenstvích, které bylo asi na přelomu našeho století pokládáno za statistiku, tvoří dnes — ani ne tak metodikou jako spíše předmětem — převážně pracovní okruh *úřední statistiky*. V jejím rámci převládající úlohu má *hospodářská statistika*. Ve Statistické ročence NSR připadají asi $^4/_5$ vnitrostátní části na údaje z oblasti hospodářské statistiky: obchod, zemědělství, ceny a mzdy, peníze a úvěr, dopravu, průmysl, zaměstnanost, státní rozpočet, národní důchod atd. Nejstarší forma statistiky, *statistika obyvatelstva*, má vedle toho stále ještě významnou úlohu, zatímco *statistika kultury*, školství, náboženství a soudnictví hrají spíše skromnou roli. (Ve Statistické ročence ČSSR je podíl hospodářské statistiky stejně velký.)

Dříve než řekneme několik slov o hospodářské statistice (a o obecných problémech úřední statistiky), musíme si ještě něco vyjasnit. Již několikrát jsme stavěli proti sobě „*klasickou*" *statistiku* (sčítání, příprava a uspořádání hromadných jevů) a „*moderní*" — analytickou nebo induktivní — *statistiku* (úsudky z výběrových souborů), takže mohl vzniknout dojem, jako by existovala jedna zastaralá a jedna moderní forma statistiky, které nemají navzájem nic společného.

Tedy, protiklady nejsou tak ostré. Také „*státovědná*" statistika používá ve stále větším rozsahu metod *induktivní* statistiky, a tím se jí otvírají nové, rozsáhlé obzory, které nebyly přístupné při vyčerpávajícím zjišťování. O dějinách sčítání lidu jsme již mluvili na jiném místě (oddíl 5.21) a víme, že vyčerpávající zjišťování takového rozsahu se i nadále důsledně provádí v desetiletých intervalech. Stále více jsou však doplňována novými druhy zjišťování, které si vynutila zvýšená potřeba informací v současné rychle se měnící světové a hospodářské struktuře a které umožňují používání techniky výběrů. Jednou z nich je metoda označená *CPS (Current Population Survey* — běžné sčítání lidu), nazývaná také „*mikrocensus*", jež se nejdříve objevila v Americe. První mikrocensus v NSR byl proveden v roce 1957. V podstatě jde o dotazování malého počtu domácností (1—2 %), které zvlášť svědomitě provádějí dobře školené osoby.

Tímto způsobem se zabijí dvě mouchy jednou ranou. Především je možno klást více a podrobnějších otázek než při vyčerpávajícím zjišťování, a za druhé vysoká kvalita výsledků zjišťování většinou vyloučí obávanou systematickou chybu, takže nevyhnutelná výběrová chyba je snesitelná. (Přitom se navíc získají cenné zkušenosti pro další sčítání). — V ČSSR se podobná šetření domácností provádějí zhruba ve dvou až tříletých intervalech.

Mikrocensus se většinou provádí ve čtvrtletních intervalech a je velmi aktuální. Kromě toho je (celkovou vynaloženou částkou) podstatně lacinější a dá odpověď na mnohem více otázek. Zvlášť dobře se osvědčuje při běžném pozorování osobní spotřeby a změn struktury zaměstnání a příjmů. Obsah kladených otázek je většinou výrazně aktuální, jak o tom svědčí tyto charakteristické formulace: „Kolik hodin jste v uplynulém týdnu skutečně pracoval?" nebo: „Jakou činností jste se v minulém týdnu zabýval?" Odpovědi poskytují přesný skutečný obraz současných událostí, nezkreslený nepochopenými otázkami dotazníků.

Někdy totiž není vůbec snadné nalézt správnou a jednoznačnou odpověď na všeobecnou otázku týkající se „pracovní doby" nebo „povolání". Stěží se také najde čtenář, který by si při studiu dotazníku nekladl otázku: Jaká je v tomto případě opravdu výstižná a poctivá odpověď? (Zda se potom napíše, to je už jiná věc.)

Takové potíže a pochybnosti nelze však zdůvodnit naivitou těch, kdo dotazník sestavují nebo jej pak vyplňují. Opravdu, některé otázky vypadající na první pohled docela prostě, působí při bližším zkoumání téměř zmateně.

Bylo by možné napsat celou knihu jen o jazykových a logických problémech statistických zjišťování. Postačí však, budeme-li se alespoň chvíli zabývat formulací otázek týkajících se „povolání" a „výdělečné činnosti". Co je povolání? Je to povolání, kterému se člověk vyučil, nebo které právě vykonává? Často jsou totožné, nikoliv však vždy.

Téměř „nesčetné" jsou pak údaje o povoláních, které se uvádějí, jestliže dotazník trvá na pokud možno přesném označení povolání: při posledním rakouském sčítání lidu jich bylo asi 10 000, a ta pak byla rozdělena do 550 skupin povolání, které dnes používá mezinárodní statistika.

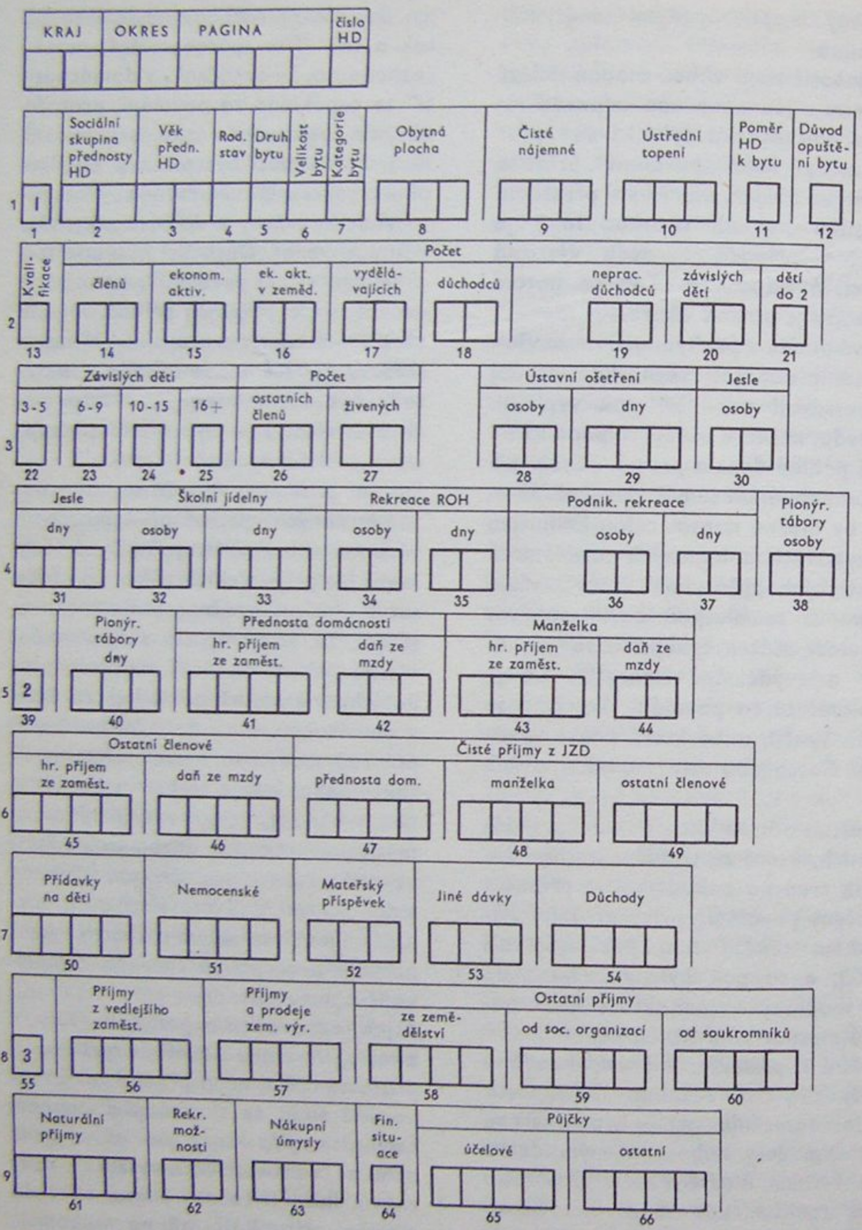
Povolání a postavení v povolání mohou někdy být dvě rozdílné skutečnosti: strojní zámečník může být mistrem, vedoucím čety nebo zaučeným dělníkem. *Povolání a odvětví* nelze vždy zcela jasně rozlišit: kdo může s určitostí rozlišit, ve kterém průmyslovém odvětví pracuje „lisař" nebo „formíř" anebo, chcete-li, „statistik"?

Některé problémy bylo možno řešit jen tak, že se rozhodlo: udělá se to tak a tak. Tím způsobem bylo např. rozhodnuto, že označení „v domácnosti" se nepokládá za povolání, protože obvykle používaná definice povolání říká, že to musí být placené zaměstnání trvající určitou dobu — s „platem" to však je u ženy v domácnosti většinou pochybené. Důchodci jsou pak tzv. „samostatní bez povolání", a tak patří do kategorie příjemců příjmů, nikoliv však do kategorie pracujícího obyvatelstva. Důchodce už ze zásady nemůže tedy být nezaměstnaný — tato pochybná výsada je vyhrazena pracujícím „osobám výdělečně činným".

Dalším problémem je, že se počty nezaměstnaných zjištěné při mikrocensu většinou velmi podstatně liší od těch, které lze při šetřeních získat od úřadu sociálního zabezpečení. Stále znovu se stává, že ženy opustí své povolání a jsou pak až na další v domácnosti. Subjektivně se už pokládají za ženy v domácnosti, ale z finančních důvodů pobírají podporu v nezaměstnanosti tak dlouho, jak je to jen možné. Jsou nezaměstnané, protože existují nezaměstnaní pracující, anebo jsou „vyživované osoby", protože jsou „ženami v domácnosti"? „Nezaměstnaní pracující" není mimochodem rozporný výraz, protože pracující se člení na zaměstnané a nezaměstnané.

A jak je to s *vedlejším povoláním?* Kdy se z vedlejšího stane hlavní? Je rozhodující pracovní doba, nebo příjem? Jaká různorodá směs se skrývá pod pojmem „samostatný", který se, ať chceme nebo ne, ve statistikách vyskytuje vždy znovu. Rozpětí tohoto pojmu sahá od malého příštipkáře až po velkoprůmyslníka, od uměleckého řemeslníka až po srdečního chirurga.

Spokojíme se s těmito stručnými po-


KRAJ | OKRES | PAGINA | číslo HD

1 — Sociální skupina přednosty HD | Věk předn. HD | Rod. stav | Druh bytu | Velikost bytu | Kategorie bytu | Obytná plocha | Čisté nájemné | Ústřední topení | Poměr HD k bytu | Důvod opuštění bytu — 1 2 3 4 5 6 7 8 9 10 11 12

2 — Kvalifikace | Počet: členů | ekonom. aktiv. | ek. akt. v zeměd. | vydělávajících | důchodců | neprac. důchodců | závislých dětí | dětí do 2 — 13 14 15 16 17 18 19 20 21

3 — Závislých dětí: 3-5 | 6-9 | 10-15 | 16+ | Počet: ostatních členů | živených | Ústavní ošetření: osoby | dny | Jesle: osoby — 22 23 24 25 26 27 28 29 30

4 — Jesle: dny | Školní jídelny: osoby | dny | Rekreace ROH: osoby | dny | Podnik. rekreace: osoby | dny | Pionýr. tábory osoby — 31 32 33 34 35 36 37 38

5 — 2 | Pionýr. tábory dny | Přednosta domácnosti: hr. příjem ze zaměst. | daň ze mzdy | Manželka: hr. příjem ze zaměst. | daň ze mzdy — 39 40 41 42 43 44

6 — Ostatní členové: hr. příjem ze zaměst. | daň ze mzdy | Čisté příjmy z JZD: přednosta dom. | manželka | ostatní členové — 45 46 47 48 49

7 — Přídavky na děti | Nemocenské | Mateřský příspěvek | Jiné dávky | Důchody — 50 51 52 53 54

8 — 3 | Příjmy z vedlejšího zaměst. | Příjmy a prodeje zem. výr. | Ostatní příjmy: ze zemědělství | od soc. organizací | od soukromníků — 55 56 57 58 59 60

9 — Naturální příjmy | Rekr. možnosti | Nákupní úmysly | Fin. situace | Půjčky: účelové | ostatní — 61 62 63 64 65 66

Federální statistický úřad
Český statistický úřad
Slovenský statistický úřad

MIKROCENSUS 1973
Vyznačovací list za HD

známkami. Z těchto některých — poměrně velmi prostých — otázek z relativně malé dílčí oblasti hospodářské statistiky bylo snad přece jen možno vidět, jak mohou být (a většinou jsou) obtížné i takové statistické problémy, které se obvykle pokládají za obyčejné početní postupy. Kdo vyplňuje dotazník nějakého statistického úřadu, může se s některými otázkami dost nazlobit nebo se jim divit, ale na jedno je třeba přitom myslet: většinou jsou výsledkem dlouhých, pečlivých a svědomitých úvah. Samostatnou otázku tvoří fakt, že pro laika jsou někdy přesto téměř nesrozumitelné nebo v něm přímo vyvolávají zmatek svou zdánlivě přemrštěnou přesností — ale s tím se ovšem nepotýkají jen statistikové. Je už prostě v podstatě mnoha složitých věcí, že je nelze formulovat zcela jednoduše.

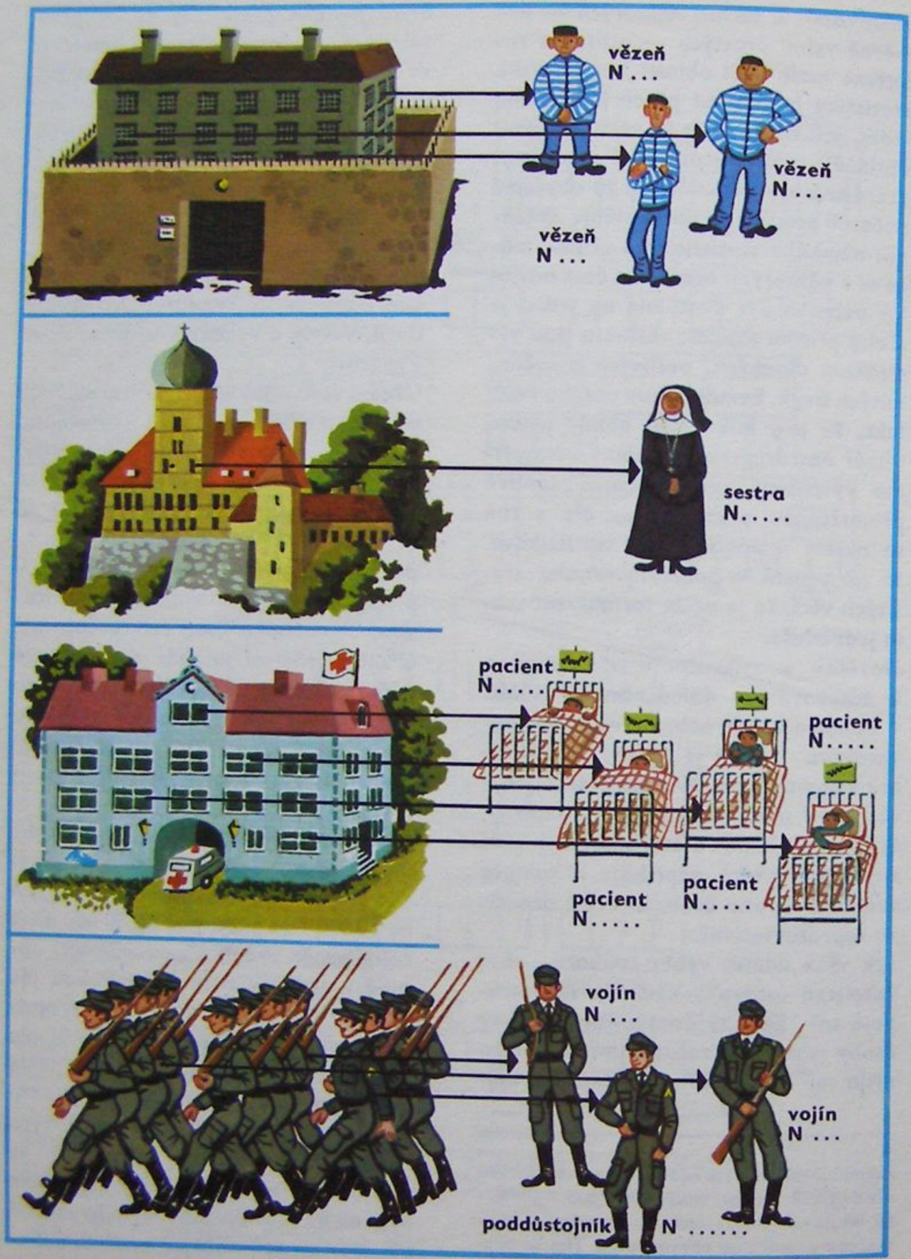
Dovětek k mikrocensu všeobecně a k dotazování v domácnostech zvlášť: v domácnostech nebydlí veškeré obyvatelstvo. V NSR je téměř 700 000 lůžek v nemocnicích. K tomu patří i „obyvatelstvo ústavů", jako jsou kasárny, kláštery, věznice a starobince. Každé zjišťování, které nepřihlíží k tomuto velkému okruhu osob, lze stěží označit za reprezentativní.

Jak však udělat výběr souboru „obyvatelstvo ústavu"? Většinou se postupuje tak, že jsou dotazovány všechny osoby se stejným začátečním písmenem příjmení (např. N), přičemž se nejdříve musíme přesvědčit, že, za prvé, relativní četnost takových jmen je ve Šlesvicku přibližně stejná jako v Bavorsku, jinak se nám do zjišťování nepozorovaně vplíží systematická chyba, a že, za druhé, toto písmeno není charakteristické pro určitý původ. (V jedné písni zpívá „roduvěrný Vídeňák", který listuje v telefonním seznamu: „Všichni moji přátelé jsou zde — a sice u písmene V: Vondrak, Vitek, Vytlacil, Vesely a Vykoupil, Vranek, Vrtil, Vypacil...")

Úřední statistika vůbec a hospodářská statistika zvlášť naráží na skutečnost, která je příčinou závažných potíží: není možno vždy očekávat, že přes veškeré zákonné předpisy (na něž se v dotaznících více nebo méně nenápadně upozorňuje) budou formuláře přesně a okamžitě vyplněny. Dotazovaný se nezřídka snaží čelit zvědavosti úřadů zběžnými odhady nebo v jisté míře nepřesnými údaji. Potom je třeba zvláštních výběrových šetření, aby bylo možno posoudit, do jaké míry a v jakém směru se od sebe liší „bajka" a pravda.

V podstatě některých jevů je však také to, že prostě *nejsou přesně zjistitelné:* kdo se odváží určit s přesností alespoň na 10 000 DM nebo jen na 50 000 DM, kolik peněz vydali západoněmečtí občané v cizině v roce 1970! Pokud jde o hrubý domácí produkt, z opatrnosti se uvádí jen na desítky miliónů, ale i toto zaokrouhlení je ještě dokladem překvapujícího optimismu. Hrubý domácí produkt NSR byl v roce 1969 („předběžný výsledek") vykázán ve výši 601 miliard DM. Kdyby toto číslo bylo na miliardu přesné, bylo by možné být šťastný, spokojený a vděčný. („Konečné údaje" za uplynulé tři až čtyři roky se vyznačují dalším místem za

◄ **Formulář mikrocensu Federálního statistického úřadu ČSSR. Pomocí mikrocensu jsou v poměrně krátkých intervalech dotazovány výběry tvořené z bydlícího obyvatelstva. Tím se překlenují desetiletá období mezi sčítáním lidu a vedle toho jsou zpracovávány četné další informace.**



K vytvoření vzorku z „osazenstva ústavu" se často používá příjmení. V našem případě byli vybráni všichni, jejichž jméno začíná písmenem N.

◄

desetinnou čárkou a přihlížejí k těm nepřesnostem, které se mezitím podařilo vystopovat, ale i pro ně stále ještě platí, že miliardu víc nebo méně nelze brát příliš vážně.)

Poznámkou o hrubém domácím produktu jsme se mimochodem již dotkli také otázky, *kde vlastně přechází hospodářská statistika do ekonometrie.* Je již v povaze přechodů, že nelze vždy přesně zjistit bod, ve kterém něco přestává a něco jiného začíná. Nejspíše je možno říci, že *ekonometrie je vyšší forma hospodářské statistiky*, obohacená o vědeckoekonomické modelování. Ragnar Frisch, jeden z nejvýznamnějších národohospodářských teoretiků naší doby, jednou řekl: „Ekonometrie se zabývá použitím matematických metod k řešení hospodářských a statistických problémů. Je v povaze věci,... že tyto problémy často vyžadují matematické postupy a metody relativně pokročilého druhu."

Zatímco statistika — a to platí zejména pro popisnou statistiku — zaznamenává a případně sestavuje údaje, skutečnosti nebo vývoj, ekonometrie se pokouší uvést tyto údaje do souladu s hypotetickými modely hospodářských procesů, analyzovat vzájemné působení různých faktorů a případně je pomocí matematických modelů napodobit a prověřit.

TVORBA A UŽITÍ NÁRODNÍHO DŮCHODU ČSSR ROKU 1971

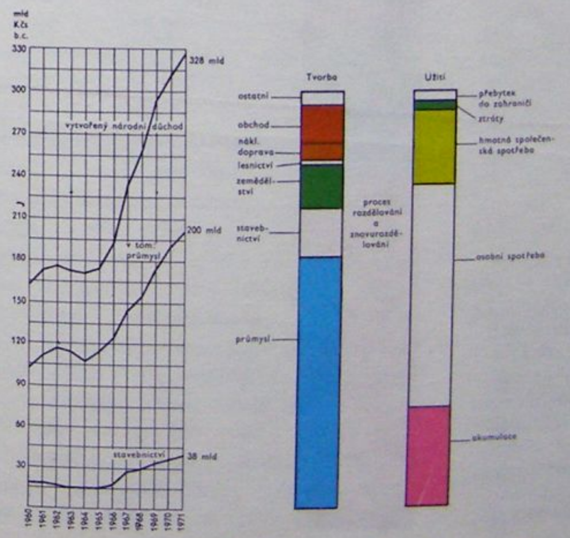


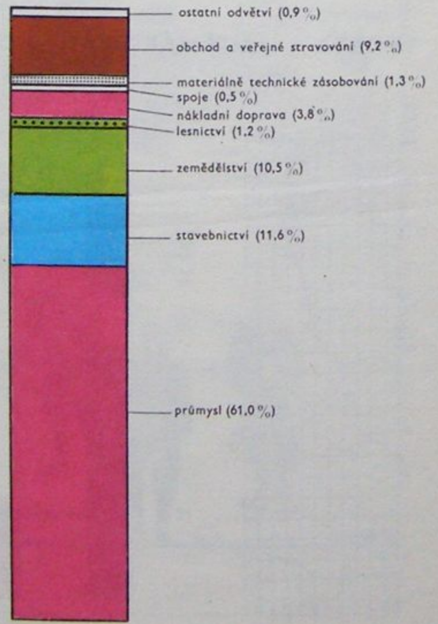
Diagram ukazuje růst národního důchodu v ČSSR, a to se zvláště vyjádřenou růstovou čarou dvou odvětví, která mají — jak je zřejmé z diagramu na str. 254 — největší podíl na tvorbě národního důchodu ČSSR: průmyslu a stavebnictví. Národní důchod je vyjádřen v běžných cenách. To znamená, že jeho růst je ovlivňován i pohybem cen. Proto se v některých případech národní důchod ČSSR vyjadřuje i ve stálých cenách: do roku 1966 šlo o ceny z roku 1960 a od roku 1967 se užívá cen, které platily k 1. 1. 1967.

Jestliže zásadám čisté statistické práce odporuje interpretace číselných závislostí (korelací) jako vztahu příčiny a následku, je ekonometrie od těchto omezení osvobozena a může vypracovat na základě statistických údajů, matematických metod a národohospodářských teorií ekonometrický model. Tím není určeno pořadí: Větší možnost použití hypotéz a teorií, které je ekonometrii dovoleno, se platí ztrátou vypočítané přesnosti; pravděpodobnostní jistota statistického rozboru platí jen na základě počtu pravděpodobnosti. Je však pozoruhodné, že obě první Nobelovy ceny za hospodářské vědy dostali dva ekonomové — Ragnar Frisch a Jan Tinbergen, a to „za vývoj a použití matematických modelů pro analýzu hospodářských procesů".

Ještě těžší je vymezení statistiky ve vztahu k těm vědním oborům, které zdomácněly v hraničním území ekonometrie, jako např. *operační výzkum nebo lineární programování*. Jsou to dva nejmodernější vědní obory, jejichž základ tvoří matematika, statistika, národohospodářská teorie a podniková ekonomika a které úzce navazují na zpracování dat.

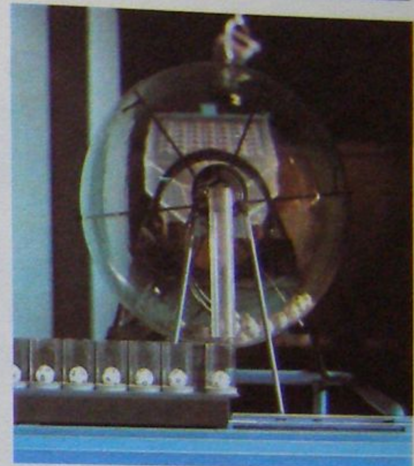
Konečně by bylo možno se ptát: *kde přechází statistika do teorie rozhodování?* To je zvlášť těžká otázka, protože lze plným právem tvrdit, že moderní statistika je podstatnou částí teorie rozhodování. Co se vlastně sleduje všemi těmi tabulkami, zjišťováním a porovnáváním, když ne toto: dát uživateli návod k racionálnímu chování! Od nejprimitivnějších prastarých mudrosloví, která jako by zachycovala rozdělení

PODÍL ODVĚTVÍ NA TVORBĚ NÁRODNÍHO DŮCHODU ČSSR ROKU 1971



Tento diagram znázorňuje strukturu vytvořeného národního důchodu v roce 1971 podle odvětví materiální výroby, která se podílejí na jeho tvorbě. Dále ukazuje strukturu užití vytvořeného národního důchodu. Za užitý národní důchod se považuje vytvořený národní důchod po odečtení přebytku platební bilance a ztrát na vytvořeném národním důchodu. V uvedeném roce činil užitý národní důchod 95,55 % vytvořeného. Nejvyšší podíl na užitém národním důchodu měla osobní spotřeba obyvatelstva (55,9 %). Akumulace činila 25,5 %.

četnosti povětrnostních úkazů, až k nejsložitějším faktorově analytickým ověřovacím metodám vyšší statistiky jde koneckonců vždy o přípravu rozhodování. *Učit se ze zkušeností* — to je zároveň metoda i konečný účel statistické práce: zkoumání skutečností, většinou jen zlomkovitě vysledovatelných, a jejich promítání do neznámého základního souboru, ale také, a to především — do neznámé budoucnosti. Statistika se nedělá pro poznávání včerejška, nýbrž proto, *aby bylo možno lépe poznat zítřek.*



Sportka je do jisté míry opakem pojištění: vyplácí se však nikoliv v případě škody, ale v případě výhry. Pojistné je zde sázkou.

8.3 Máme se dát pojistit?

Dávno před prvními elementárními úvahami o počtu pravděpodobnosti a dávno před prvními předchůdci politické aritmetiky a statistiky obyvatelstva byly známy dva jevy, které vlastně představují syntézu teorie pravděpodobnosti a popisné statistiky — totiž *sázky a pojištění*.

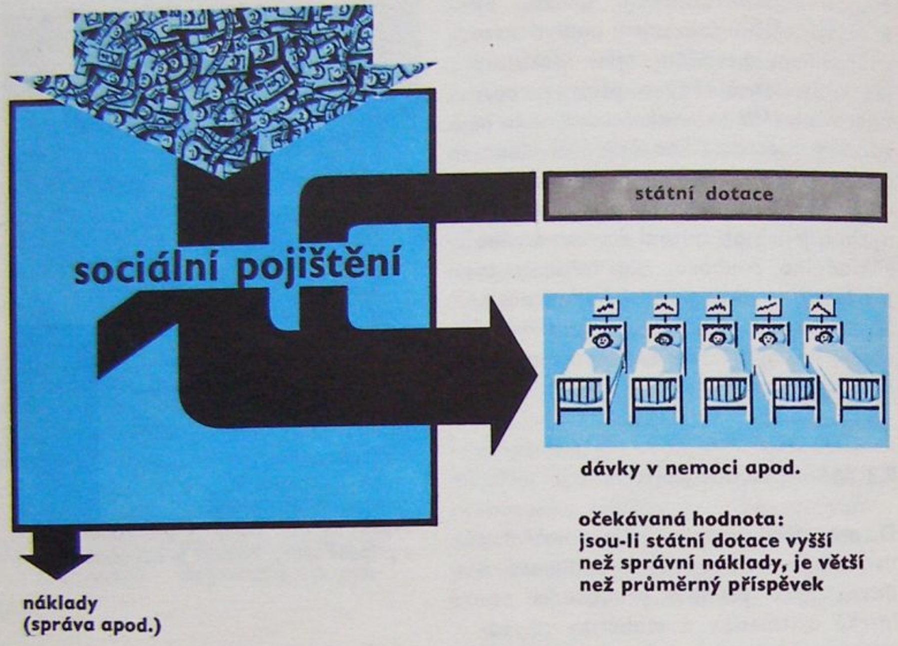
Pojištění je svou strukturou velmi podobné sázkám, ovšem s jedním podstatným rozdílem: zakládá se na velkém počtu událostí stejného druhu, zatímco při sázce stojí proti sobě dva jednotlivci rozdílných subjektivních názorů. Plynulý přechod tvoří sázkové kanceláře a velké loterie: v tomto případě — především u loterie — podnikatel (většinou stát) zprostředkuje bez vlastního rizika vyrovnání mezi množstvím sázejících. Početné malé dílčí příspěvky *(sázky)* jsou po srážce nákladů a daní odevzdány do rukou těch *několika málo*, kteří měli štěstí. (Viz obr. na str. 256—257.)

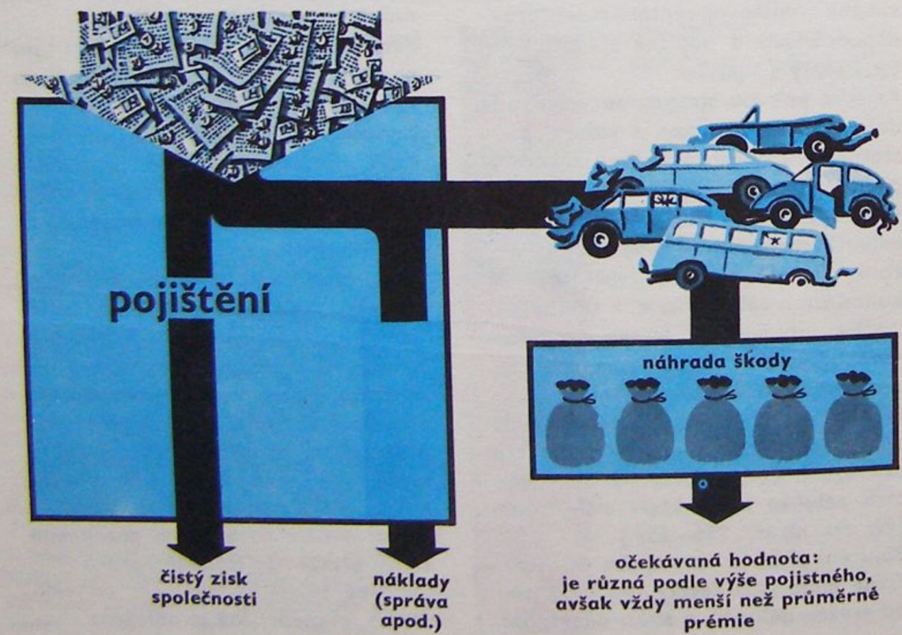
Pojištění pracuje na stejném principu: četné malé dílčí částky *(pojistné)* jsou po srážce nákladů a zisku odevzdány těm *několika málo*, kdo mají dostat náhradu za utrpěnou škodu.

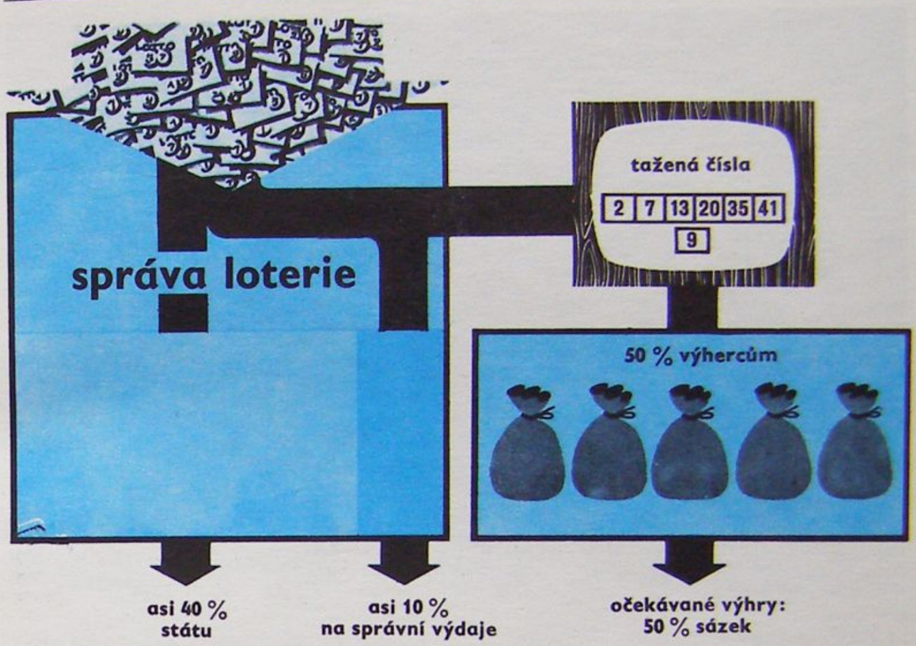
Předpokladem pro vznik pojištění bylo poznání, že jisté škodní události se vyskytují s přibližně odhadnutelnou četností. Pak přišel další logický krok. Když „každá desátá loď" ztroskotá, je možno škodu vyrovnat tak, že každý vlastník lodi zaplatí desetinu hodnoty jako pojistné.

Již ve čtvrtém století před naším letopočtem, když ostrov Rhodos ovládl lodní plavbu ve východním Středomoří a vytvořil počátky obchodního a námořního práva, vznikla první úprava rozdělení ztráty při vyhazování zboží přes palubu v případě nebezpečí na moři — úprava, která byla později jako „*lex Rhodia de iactu*" (rhodský zákon o odlehčování lodi potopením zboží) převzat do římského práva.

Uvedený zákon se zakládal na této situaci: obchodní loď je naložena zbo-







Pojištění a loterie se navzájem strukturálně podobají. Celková částka pojistného (případně cena prodaných losů) plyne nejdříve do „společného hrnce", a potom náhoda vybere jednotlivé šťastlivce (případně poškozené), kteří obdrží částky vyšší, než byl vklad, jako výhru (případně náhradu škody). Průměrná očekávaná hodnota je zásadně nižší než zaplacená částka vzhledem k tomu, že je nutno uhradit náklady apod. Pouze v případě, že státní dotace (např. na sociální pojištění) zvýší prostředky použitelné k vyplacení, může očekávaná hodnota dosáhnout výše příspěvků nebo ji překročit.

zím, které patří více obchodníkům. Dostane se do bouře a musí se zbavit nákladu, aby se nepotopila. Lodní posádka popadne, co jí právě přijde pod ruku a co se dá zvlášť snadno hodit přes palubu nebo co je zvlášť těžké — a doplujе konečně do ochrany přístavu. Záchrana lodi, mužstva a často většiny zboží byla možná jen za podmínky, že bylo obětováno zboží jednoho nebo více obchodníků — a měli by právě tito obchodníci sami být poškozeni? „*Lex Rhodia de iactu*" rozhodl tedy tak, že škoda se rozdělí na všechny, kdo měli zájem na záchraně lodi nebo nákladu. Podobným způsobem se tento problém řeší v námořním obchodním právu i dnes, jestliže se má po havárii lodi rozhodnout o rozdělení škody.

Od tohoto zákonem upraveného dělení škody po havárii je pouze malý krok k dobrovolnému *předchozímu* placení pojistného za dopravované zboží. Náklady přitom velmi podstatně klesnou, protože se pojistné platí i za ty lodní přepravy, které skončí beze ztrát.

Již v antické době podněcovalo ztroskotání lodí a podobné škodní případy vznik prvních smluv a zákonů o dělení škody. To posloužilo jako příklad pro pozdější havarijní pojištění.

Takový záměr ovšem předpokládá mnohem důkladnější organizaci systému placení, a proto první základy moderního pojištění škod nacházíme teprve koncem 13. století v Itálii, kde v souvislosti s *„námořními půjčkami"* *(foenus nauticum)* byla provedena úprava placení pojistného tak, že bohatí jednotlivci nebo peněžní ústavy pojišťovali lodní náklady.

Nemůžeme se zde podrobně zabývat dějinami pojišťovnictví, ale přece jen ještě malý výlet do oblasti vzniku slov: *„havárie"* pochází z arabského slova „avar", kterým se označovalo vadné nebo poškozené zboží. V angličtině se používá slova „average" — toto slovo známe v užívanějším významu „průměr". Tento průměr není nic jiného než *průměrná škoda*, kterou lze vypočítat na základě právě uvedených starých ustanovení námořního práva.

V počátcích pojišťovnictví bylo nutno si vypomáhat dost nespolehlivými údaji o četnosti nehod a škod. Dnes je k dispozici bohatý statistický materiál, který umožňuje výpočet *„férového"* pojistného. Jestliže mám jen určité představy o četnosti škod, mohu očekávat dva stejně nepříjemné omyly: Buď podcením četnost škod, požaduji nízké pojistné, musím moc vyplácet — a udělám úpadek. Nebo z opatrnosti nasadím pojistné příliš vysoko, zpočátku vydělávám víc než dost, ztratím však brzy zákazníky, kteří přejdou k správněji kalkulujícímu, a tím lacinějšímu konkurentovi — a vystavuji se úpadku v ještě větší míře, protože předpokladem existence pojištění je pokud možno velký počet pojištěných.

Jen tak je zaručeno, že náhodou ovlivňované škodní události jsou rozděleny tak, jak to přibližně odpovídá četnosti všech škod určitého druhu v celé zemi. Během tří let 1967—1969 se v NSR podílelo na nehodách se zraněním 399 116, 410 463 a 417 082 osobních automobilů. Celkový počet osobních automobilů činil 10,2 (1967), 10,8 (1968) a 11,7 (1969) miliónu. Z toho vyplývá průměrná četnost asi 5 případů nehod se zraněním osob na sto osobních vozidel.

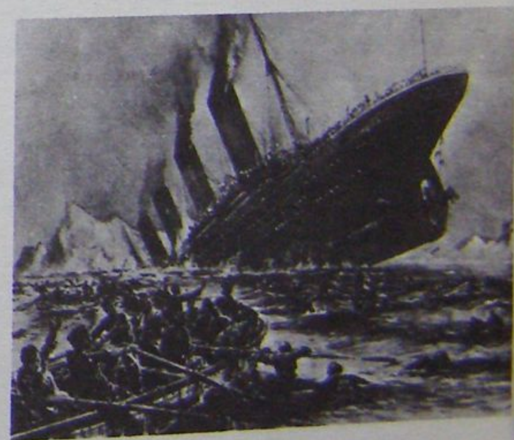
Pojišťovna s 10 000 pojištěných vozidel by tedy mohla s určitým oprávněním očekávat, že na těchto nehodách se bude podílet nejvýše 5 % jejích pojištěnců. (Přesně se to propočítá tak, že se z téměř nekonečného základního souboru s četností škod $p = 0{,}04$ usuzuje na četnost ve výběrovém souboru o rozsahu $n = 10\,000$; srovnej kap. 3 a 6.)

Docela jinak by tomu bylo, kdybych chtěl založit *minipojišťovnu*, našel 25 zákazníků a vypočítal si, že „podle počtu pravděpodobnosti a statistiky" se bude jen jeden podílet na nehodě se zraněním osoby, a podle toho bych kalkuloval pojistné. Výsledek by byl víc než zhoubný. Bylo by totiž docela dobře možné, že právě u mých pojištěnců by nastaly dvě, tři nebo ještě více nehod. Proti tomu bych se musel chránit zajištěním u některé velké pojišťovny — a vzniklé výdaje bych přitom bohužel musel přesunout na svých 25 pojištěnců. Ať už máme cokoliv proti koncernům, pojišťovny *musí* být velké.

Mimořádně veliká rizika proto málokdy nese jen jedna pojišťovna. Žádná pojišťovna není kapitálově tak silná, aby byla zabezpečena proti ztrátě 50 000tunové osobní lodi uzavřením dalšího tisíce podobných smluv. Proto dochází k celému řetězu zajištění, takže riziko je rozděleno různým dílem mezi půl tuctem velkých pojišťovacích ústavů.

Zájem pojišťoven na uzavírání pojistek je dobře znám přinejmenším těm, kterým pojišťovny neustále zasílají lákavé dopisy a k nimž posílají výřečné zástupce. Pojišťovny přece kalkulují pojistné tak, aby vždy dosáhly zisku. Ale proč by se jim mělo přitom pomáhat? Co je přece pro pojišťovnu ziskem, musí být pro pojištěnce nutně ztrátou — nebo ne?

V daném případě nutno rozlišovat dvě věci, které se velmi lehce směšují: *matematicky objektivní očekávanou hodnotu a subjektivní osobní riziko.* „Oče-



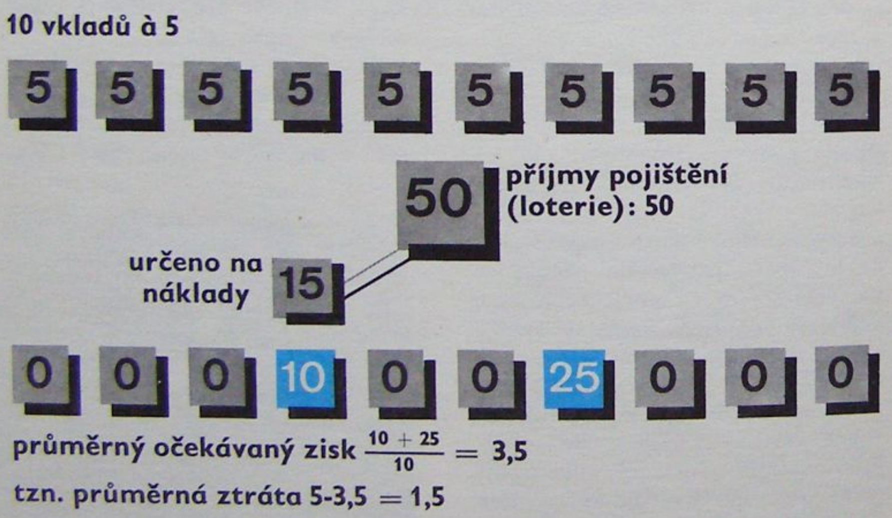
Katastrofy velkého rozsahu, jako např. zánik Titanicu, nemohou krýt ani nejbohatší pojišťovací společnosti samy. Systém zajišťování musí pomoci nést taková rizika.

kávanou hodnotu" je možno početně přesně stanovit: je to průměrná částka výplat vztažená na celkovou částku pojistného. Když tisíc lidí zaplatí do loterie po 100 DM a celkem je vylosována jedna hlavní výhra v částce 10 000 DM a deset výher po 5000 DM, pak proti celkové sázce ve výši 100 000 DM stojí součet výher 60 000 DM — očekávaná hodnota každého losu po 100 DM je tedy přesně 60 DM. Očekávaná výhra je vlastně „*očekávanou ztrátou*" ve výši 40 DM; na každého výherce 5000 marek připadá více než 80 prohrávajících.

Očekávaná ztráta v loterii není ani pro celkem rozumně uvažujícího člověka vždy odstrašujícím momentem, protože představuje i jakýsi *dráždivý a napínavý zážitek*. Rozumně myslící člověk se především ptá: *Mohu si dovolit takovou sázku*, kterou s největší pravděpodobností ztratím? Jestliže ano, bude tuto hru pokládat za to, čím ve skutečnosti je — za dárek státní pokladně s tím, že získává malou naději na velkou výhru. Nerozumně se chová ovšem ten, kdo má jen stěží zajištěno živobytí, a přesto nosí své peníze do sázkové kanceláře.

U pojištění musí naproti tomu znít první rozhodující otázka takto: *Mohu si dovolit škodu, ztrátu?* V tomto případě je tomu tak, že méně zámožný bývá více postižen.

Přesto se však bohatší chovají jako



Deset osob platí po 5 DM pojistné (nebo za losy). Z celkových 50 marek se — v našem příkladu — vydělí 15 marek na náklady, zbytek se vyplácí. Jeden účastník dostane 10 DM, druhý 25 DM jako výhru nebo odškodnění, zbývající vyjdou naprázdno. V této „hře" činí očekávaná hodnota 3,5 DM, sázka 5 DM.

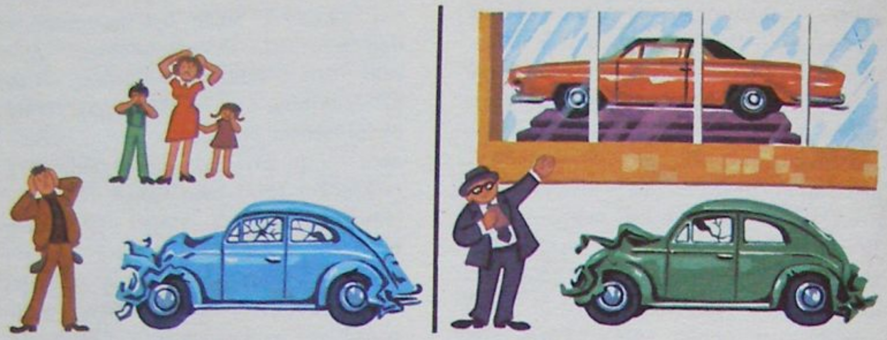
chytřejší obchodníci: ačkoliv je ztráta ve vážné chvíli ani zdaleka tak nepostihne, přece platí běžně pojistné (které je sotva zatěžuje) a jsou ušetřeni případné mrzutosti a neočekávané škody. „Malý člověk" naopak skrblí na pojistném, které se mu zdá příliš vysoké — a v případě škody je bez ochrany.

Je pojistné opravdu příliš vysoké? Samozřejmě že je příliš vysoké, a to nikoliv bez příčiny. Pojišťovatel musí vybrat dost peněz nejen proto, aby mohl nahradit škody vzniklé v běžném roce, ale i proto, že musí také krýt všechny své náklady — od nákladů na propagaci až po platy ředitelů, od výdajů za formuláře až po platy zaměstnanců, a kromě toho dosáhnout určitého zisku. Musí si také vytvořit rezervy, protože skutečná četnost škod někdy může být mnohem vyšší než odhadnutá. Na 100 DM vyplacených náhrad škod (nebo jiné pojistné dávky) musí tedy inkasovat 120, 140 a více DM pojistného — podle druhu pojištění, počtu pojištěnců, vlastních výdajů a dosahovaného zisku.

Takto posuzováno je každé pojištění pro pojištěného špatným obchodem a je spojeno jen s „očekáváním ztráty". Pouze sociální pojištění může mít kladnou očekávanou hodnotu — nikoliv však proto, že by zde byla pojistná matematika postavena na hlavu, nýbrž proto, že z příspěvků zaměstnavatelů, dotací apod. plynou pojišťovně prostředky, které pojištěnec nemusí vynakládat. I v tomto případě se musí za 100 DM „škody" zaplatit 120 nebo 140 DM — pojištěnec sám však z toho platí jen 60 nebo 80 DM a o zbytek se postarají dobří „skřítkové". Jestliže toto nevezmeme v úvahu, je každé pojištění „v průměru" a „pro pojištěnce jako celek" ztrátové. Pro jednotlivce však může být nevyhnutelnou nutností. Pravděpodobnost zničení vozu o hodnotě 4000 DM — za předpokladu, že oba řidiči mají stejné jízdní schopnosti, vozy stejný výkon atd. — je pro čtvrtý vůz milionáře (s kterým jezdí jeho dcera nebo tajemník) přibližně stejná jako pro obchodního cestujícího, který do vozu vložil celou svou hotovost a který jej naléhavě potřebuje pro výkon svého povolání. Je koneckonců pouze věcí náhody, zda ten nebo onen vůz je jedné noci úplně demolován cizím šíleným řidičem. Pro pojišťovnu to jsou dva naprosto stejné případy: vůz v hodnotě 4000 DM byl totálně demolován. Pro pojištěné jsou však důsledky velmi rozdílné. Pro „malého člověka" může mít taková škoda dalekosáhlé důsledky: výpůjčka peněz a lichvářské úroky, zadlužení, případně nervové zhroucení a sebevražda. Pro bohatého je to jen nemilá příhoda a nic víc.

Zásadně tedy platí pravidlo: *Čím méně je možno dovolit si ztrátu, tím spíše je nutno se pojistit.* (Obtížný hraniční případ se vyskytuje tehdy, když je četnost škod tak nepatrná, že i malé pojistné se zdá být zbytečným nákladem.)

To není žádné průkopnické poznání. Již v začátcích počtu pravděpodobnosti, před 250 lety, byla položena otázka, jak velký musí být majetek jednoho člověka, aby sám mohl nést určité riziko. Všeobecně platné matematické řešení této otázky nelze nalézt, ačkoliv bylo od té doby provedeno mnoho pokusů a vyskytla se řada podnětů. Velmi mnoho totiž přece záleží také na *osobním potěšení z podstupovaného rizika*. Úzkostlivý a opatrný člověk nebude tolik riskovat jako temperamentní a dynamický, a přitom oba mohou



Oba škodní případy jsou prakticky úplně stejné, pro jednoho to však znamená finanční katastrofu, pro druhého nový, lepší vůz. Poučení: kdo by byl škodou těžce postižen, musí se proti ní bezpodmínečně pojistit.

jednat subjektivně zcela „správně“. To samozřejmě platí i v otázce pojištění, které v podstatě riziko vyřazuje — za odpovídající úplatu. Kromě toho je nutno poukázat na jedno důležité hledisko: pojištění nemůže ovšem zabránit určité události, může jen poskytnout finanční pomoc v případě vzniku škody. Pojištění příslušníků rodiny může tedy vyplynout pouze z finančních, nikoli citových pohnutek.

Před uzavřením pojistné smlouvy zní proto otázka vždy takto: *Jak silně by mne* (případně příslušníky mé rodiny) *finančně zatížil „vznik pojistné události“? A jak pravděpodobný je vznik takové události?* Odpověď je nutno dlouho, pečlivě a kriticky zvažovat, protože nesprávný optimismus může být osudný. A konečně — *v jakém poměru je výše pojistného k mému odhadu rizika*, i s přihlédnutím k nematematickým okolnostem? Rozhodující úlohu často totiž má osoba pojištěnce. Stojím-li před otázkou, jestli uzavřít pojistku na cestovní zavazadlo proti krádeži, musím si položit otázku, zda zpravidla dávám na své věci pozor, zda jsem důvěřivý, zda se mi již vícekrát ztratily cenné věci apod. Existuje samozřejmě i jiné riziko, jestliže mám v úmyslu bydlet s příručním kufříkem v rodinném penziónu, anebo budu-li s drahým fotoaparátem a několika prvotřídními obleky bydlet každý druhý den v jiném hotelu.

Pojišťovnám je ovšem také známo, že o pojistnou ochranu se ucházejí právě ti lidé, kteří se cítí nadprůměrně ohroženi — a podle toho také propočítávají pojistné. Absolutně spolehlivé pravidlo, co a kdy se má pojistit, není možno stanovit.

Ještě několik slov *k vysokým rizikům s malou pravděpodobností*. Zdá se, že by mělo smysl platit jednu marku ročního pojistného za pojištění rodinného domku proti zřícení letadla? — Tak jako je malá pravděpodobnost takové katastrofy, je nízké i pojistné. Avšak před uzavřením takovéhoto pojištění by bylo nutné si ověřit, zda je pojišťovna schopna platit. Kdyby se totiž o zavedení podobného speciálního pojištění pokoušel nějaký čilý, ale malý podnikatel, musel by pojistit téměř všechny rodinné domky ve Spolkové republice, aby vůbec mohl z přijatého pojistného krýt jeden nebo dva případy vzniklé škody — a mimoto by jeho výlohy byly vyšší než celková částka pojistného. (Zpravidla jen jedna nebo dvě světoznámé mezinárodní pojišťovny pojišťují nohy filmových hvězd proti úrazu nebo cenné sbírky známek proti zničení v případě války. To proto, že jde o vysoké pojistné, které nemůže být vyrovnáno statistickými četnostmi. Vznik škody by mohl malou nebo střední pojišťovnu zcela zničit, zatímco pro velkou je ztráta snesitelná a bezplatná reklama je skutečným ziskem.)

Kromě toho pojistná ochrana přestává platit mimo smluvní podmínky a při změně situace, která existovala v době uzavření smlouvy. Proti válce a jejím následkům pojištění neexistuje. Dokonce ani proti pronikavé změně politického kursu a pouze ve zcela omezeném rozsahu proti světové hospodářské krizi. Vzhledem k tomu, že neexistuje absolutní jistota neomezeného trvání kteréhokoliv státu a hospodářského řádu, nemůže být jistota ani v tom, že pojištění bude platit na smluvní dobu. Nicméně k základnímu principu rozumného chování patří, aby se nejvýš nepravděpodobné pokládalo za nemožné. Tedy nikoliv pojištění pro případ zřícení letadla, ale přesné placení pojistného uzavřeného životního pojištění, sdruženého pojištění domácnosti, případně i havarijního pojištění.

8.4 Délka lidského života a statistika obyvatelstva

Již několikrát jsme hovořili o statistice obyvatelstva — např. když jsme psali o historických začátcích statistických šetření, která všechna byla sčítáním lidu, nebo také, když jsme v sedmé kapitole při výkladu možností a úskalí extrapolace na příkladu populační exploze a populačních pyramid ukazovali, jak velké je nebezpečí nesprávných odhadů i v této zdánlivě tak jednoduché oblasti. Konečně jsme se také zmínili o tom, že vedle převládajícího významu hospodářské statistiky má statistika obyvatelstva stále ještě i dnes významnou úlohu. Tři kapitoly: „Území a obyvatelstvo“, „Pohyb obyvatelstva“ a „Zdravotnictví“, nejsou na začátku statistických ročenek uváděny jen tradičně, nýbrž tvoří stále ještě základ pro všechny ostatní údaje, neboť pro koho jiného pracuje statistika, když ne pro lidi, pro obyvatelstvo?

V kapitole o statistice obyvatelstva můžeme samozřejmě tuto oblast vyčerpat pouze v neúplné míře, stejně jako tomu bylo i u ostatních dílčích odvětví statistické práce. Abychom doplnili, co již bylo dříve uvedeno, zastavíme se poněkud podrobněji u jedné dílčí oblasti, která stále znovu vyvolává nedorozumění a která má pro jednotlivce jistě velký význam: u problému *„délky lidského života“* nebo, řečeno negativně, *„pořadí vymírání“*.

I když někdo sotva ví něco o statistice obyvatelstva, zcela jistě slyšel o tom, že lidský život je dnes zásluhou pokroku v lékařství a sociálních vymožeností v průměru mnohem delší, než byl ještě na přelomu našeho století, nemluvě už vůbec o stoletích dřívěj-

věk úmrtí v letech: 2 14 50 60 65 70 75 80

dosažené průměrné stáří: $\frac{2+14+50+60+65+70+75+80}{8}=\frac{416}{8}=$ **52 let**

věk úmrtí v letech: 50 60 65 66 70 70 75 80

dosažené průměrné stáří: $\frac{50+60+65+66+70+70+75+80}{8}=\frac{536}{8}=$ **67 let**

Vyloučení úmrtnosti dětí vede k prudkému zvýšení průměrné délky života. V horním příkladu zemřou dvě osoby z osmi ve věku dvou, resp. čtrnácti let. V dolním příkladu se předpokládá, že obě tyto osoby dosáhly věku 66, resp. 70 let. V ostatních případech se věk úmrtí nemění. Dosažený průměrný věk se přitom zvyšuje z 52 na 67 let.

šich. Průměrná délka lidského života vzrostla od začátku našeho století v NSR u mužů ze sotva 45 let na více než 67 let, u žen dokonce ze 48 na více než 73 let. (Živě narozené dítě mužského pohlaví má v ČSSR očekávanou průměrnou délku života 67,3 roku, i když některé z dětí jednou oslaví své sté narozeniny a jiné se smrtelně zraní na školním výletě. 67,3 roku je průměr a jednotlivá dřívější úmrtí snižují aritmetický průměr oproti mediánu.)

Přitom jde samozřejmě o „*průměrnou délku lidského života*", tedy střední hodnotu jednotlivých délek života. Výpočet však není tak jednoduchý — vždyť údaje o skutečném trvání života jednotlivých ročníků jsou plně k dispozici teprve o sto let později; v daném případě však nechceme jít do podrobností. Zásadně jde o zjištění, jakého stáří dosáhne novorozenec v průměru, který je aritmetickým průměrem. Výsledkem tohoto propočtu je ona „průměrná doba života", která pak slouží např. pojišťovnám jako základ pro životní pojištění novorozence.

Tato průměrná délka lidského života se od přelomu století velmi zvýšila a ve srovnání s dobou před sto lety se dokonce téměř zdvojnásobila. Příčinou je ovšem především skutečnost, že se velmi podstatně snížila úmrtnost kojenců a dětí, jak jsme mohli zjistit již při diskusi o populační explozi.

Nebezpečí prvního roku života se však dodnes statisticky projevuje v tom, že *jednoroční dítě má delší očekávanou průměrnou délku života než novorozenec:* překonalo již kritický první rok a je tedy možno odečíst dosud existující vysoké riziko „smrti v prvním roce života" (pravděpodobná úmrtnost přes 3 %, hodnota, která se znovu dosahuje až po 65. roce života!). Dnes tento zisk

činí jen o něco více než půldruhého roku života, ale v letech 1870—1880 to bylo ještě tak, že živě narozené novorozeně mělo očekávanou průměrnou délku života jen 35,6, jednoleté dítě již 46,5 a tříleté dokonce 49,5 roku! Dnes je očekávaná průměrná délka života tříletých přibližně stejná jako novorozených. Kdybychom chtěli použít pesimistický výklad, mohli bychom říci, že se ztratil dřívější zisk 14 let života.

Ve skutečnosti je to tak, že prodloužení průměrné délky života bylo dosaženo především, nikoliv však výlučně, velkým poklesem kojenecké a dětské úmrtnosti. Také střední a starší ročníky získaly několik let života zlepšením hygieny a moderními léčebnými metodami, ovšem zdaleka ne tolik jako novorození. Padesátiletý muž mohl před sto lety žít ještě 18 let, na přelomu století 19,5 roku a dnes 23 let. Šedesátiletá žena dosahovala před sto lety „průměrně" 73 let věku, na přelomu století 74 a dnes téměř 79 let. Pro osmdesátileté a devadesátileté se nezměnilo téměř nic.

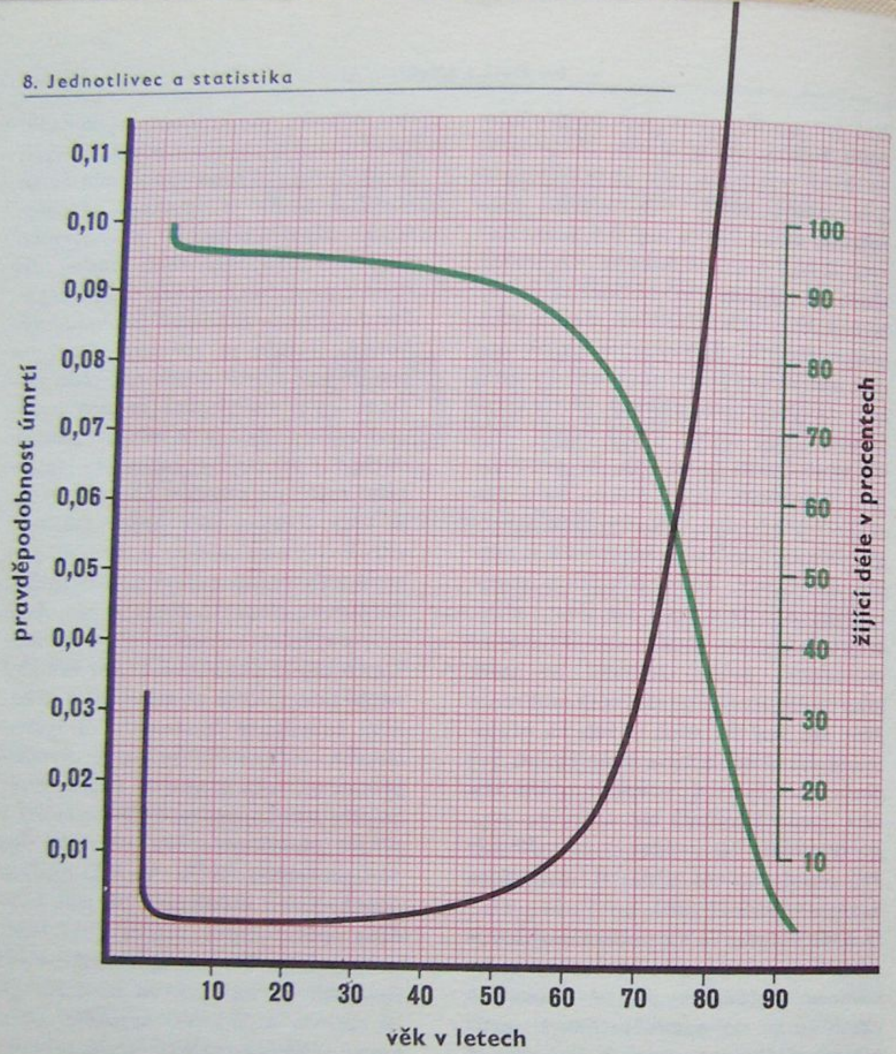
Je proto třeba velmi pečlivě rozlišovat mezi tím, co se nepřesně nazývá „*očekávanou průměrnou délkou života*" a vztahuje se na pravděpodobný „průměrný věk" novorozených, vypočítaný pojistnou matematikou, *a pro jednotlivce (a jeho pojištění) mnohem důležitější, věkem specifikovanou průměrnou délkou života.* Kdo např. uzavře životní pojistku ve věku 35 let, tomu pojišťovna přizná dalších asi 36 let života, pojistné vypočítá za předpokladu, že „tento zákazník se dožije věku 71 let".

Pojišťovna ovšem nekalkuluje na tomto základě. Většinou totiž vyžaduje lékařské vyšetření a vyšší rizika (např. nemocné těžkou tuberkulózou, po operaci žaludku atd.) přitom z pojištění vylučuje nebo pojišťuje pouze za vyšší pojistné. Kromě toho vypočítaná délka lidského života se vztahuje na celek *všech* pětatřicetiletých a normální pojištěnci stejného věku budou žít spíše o něco déle.

Pro pojišťovnu zde může být velmi výhodné, jestliže je smlouva uzavřena pro případ úmrtí: pojištěnec platí pojistné déle. Nevýhodné to pro pojišťovnu bude, bylo-li dojednáno placení důchodu; při pojištění důchodu, splatného ročně po dosažení 65. roku věku, je tedy třeba použít jiných měřítek: vážně nemocní budou v daném případě kalkulačně nejlepšími zákazníky. Pojišťovna naproti tomu může, je-li stav pojištěných dostatečně velký, dokonce běžně porovnávat, zda vznikají neobvyklé výkyvy ve srovnání s úředními tabulkami úmrtnosti. Ve statistických ročenkách se totiž uvádějí pro každý nebo pro každý pátý rok věku příslušné „pravděpodobnosti úmrtí". (Pravděpodobnost úmrtí 0,045 658 pro padesátiletého muže ve zkrácených úmrtnostních tabulkách, jak jsou uvedeny ve Statistické ročence ČSSR 1968, znamená, že ze 100 tisíc mužů, kteří se dožili 50 let, jich ve stáří 50—54 let zemře 4566.) Ve vysokém věku tato pravděpodobnost prudce vzrůstá: ze 100 000 devadesátiletých se jich 95. narozenin nedožije 86 332.

Z úmrtnostní tabulky je možno téměř na první pohled nalézt odpověď na tyto otázky: třicetiletý muž se ožení s dvacetiletou ženou. Jak velká je pravděpodobnost, že za 40 let oba budou ještě živi, že zůstane na živu jen jeden z nich, že oba zemřou?

Nejdříve najdeme v řádku „muž, 30" číslo 94 311. Říká, že ze 100 000 novorozených mužského pohlaví tohoto věku



Úmrtnostní tabulka v grafickém vyjádření. Černá křivka udává pravděpodobnost úmrtí pro každou věkovou skupinu. Zelená křivka ukazuje pozvolné vymírání jednotlivých ročníků. Zelená stupnice vpravo ukazuje, kolik procent z daného ročníku ještě žije. Nápadné na pravděpodobnosti úmrtí je stále ještě vysoké ohrožení nově narozených: teprve v sedmdesátém roce věku je pravděpodobnost úmrtí v průběhu roku opět tak vysoká jako pro nově narozeného. Pak však křivka stoupá ostře vzhůru: v osmdesátém roce věku dosáhne pravděpodobnosti 0,1, v osmdesátém pátém roce již 0,16, v devadesátém pak 0,24 (tzn. že jedna ze čtyř devadesátiletých se nedožije svých 91. narozenin). Graf se zakládá na údajích rakouské všeobecné úmrtnostní tabulky pro ženy 1959/1961, je však typický charakteristickými znaky i pro muže a též pro jiné evropské země. (Důležitý rozdíl: nově narození mužského pohlaví mají vysokou pravděpodobnost úmrtí: 0,042, protože však muži vůbec umírají dříve, dosáhne se této pravděpodobnosti opět již v šedesátém sedmém roce věku.)

dosáhne jen 94 311. Náš muž však tohoto věku již dosáhl a tak to, co již bylo, se nás netýká. Nyní se podívejme, jak to vypadá o 40 let později: v řádku „muž, 70" najdeme 54 683 — rovněž vztažených na 100 000 nově narozených. Naději na dožití udává kvocient 54 683 : 94 311 = 0,58. Pravděpodobnost úmrtí pro celé období 40 let představuje to, co chybí do jistoty (1), tedy 1 — 0,58 = 0,42.

Naděje na dožití u podstatně mladší ženy je pochopitelně mnohem větší. V řádku „žena, 20" najdeme 97 068, „žena, 60" stále ještě 86 788; kvocient je tedy 0,895 pro dožití a 0,105 pro úmrtí.

Ke zjištění pravděpodobnosti pro složené jevy postačí tato čísla odpovídajícím způsobem navzájem násobit, protože máme navzájem nezávislé pravděpodobnosti. Pro případ „oba se dožijí": 0,58 . 0,895 = 0,52. Případ „oba zemřou" bude velmi nepravděpodobný: 0,42 . 0,105 = 0,04. Naproti tomu lze zcela dobře počítat s tím, že žena zatím ovdoví: 0,42 . 0,895 = 0,38. Není však příliš pravděpodobné, že by muž přežil svou ženu: 0,58 . 0,105 = 0,06.

Z 1000 takových manželských párů by se podle toho mělo 520 dožít čtyřicátého výročí svého sňatku, a dále dožít by se jej mělo 380 vdov a 60 vdovců; ze 40 manželských párů by již nezůstal na živu ani jeden z manželů. V tomto případě jde o výpočty na základě nezávislých pravděpodobností. Protože však zkušenost ukazuje, že smrt jednoho z manželů často urychluje úmrtí druhého, je třeba se dívat na tato čísla jen jako na pojistně matematické ukazatele.

Nejen v těchto případech desetiletí společného života, nýbrž i zcela všeobecně totiž platí, že *taková čísla mají pro jednotlivce mnohem menší význam*

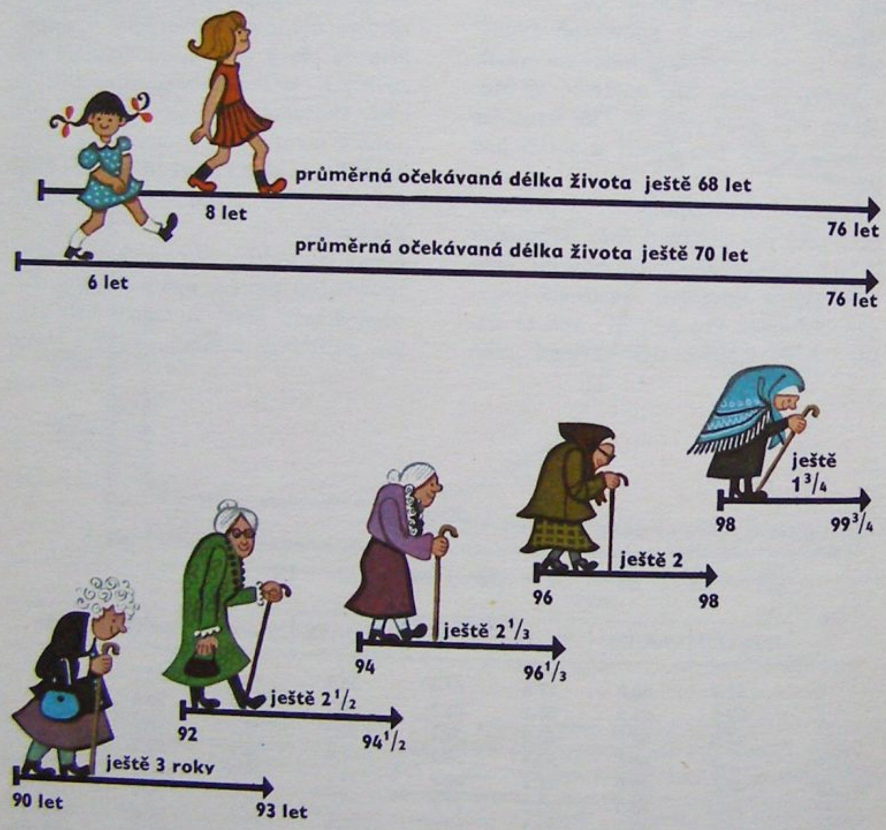
Očekávaná střední délka života v ČSSR

Věk	muži				ženy			
	1929/1932	1949/1951	1960/1961	1966	1929/1932	1949/1951	1960/1961	1966
0	51,9	60,9	67,6	67,3	55,2	65,5	73,1	73,6
1	59,9	65,5	68,4	68,2	62,0	69,4	73,6	74,1
10	54,0	57,7	60,0	59,7	56,1	61,5	65,0	65,5
20	45,3	48,5	50,4	50,2	47,4	52,1	55,2	55,7
30	37,1	39,6	41,2	40,9	39,2	43,0	45,6	46,0
40	29,0	30,8	32,0	31,7	31,0	33,9	36,0	36,4
50	21,2	22,4	23,2	23,0	22,8	25,1	26,9	27,2
60	14,3	15,2	15,5	15,3	15,3	17,0	18,4	18,6
70	8,7	9,4	9,6	9,5	9,2	10,2	11,0	11,3
80	4,7	5,4	5,2	5,3	5,1	5,6	5,8	5,9

Pramen: Statistická ročenka ČSSR 1966, 1968. Údaje za rok 1966 jsou vypočteny ze zkrácených úmrtnostních tabulek.

než pro pojištění. Ve velikém množství pojištěných se nakonec projevuje náhodné rozdělení úmrtí (od autonehody až po sešlost věkem) s překvapující přesností, pro jednotlivce však tato čísla mohou představovat jen určité odhady. Průměrně zdravý padesátník si může vypočítat a říci, že osmdesátky se asi nedožije, přesto však více než čtvrtina všech padesátiletých tohoto věku dosáhne.

Kdo ovšem dosáhl osmdesáti let, dožije se velmi pozoruhodného statistického úkazu: *s každým dalším rokem života roste zcela mimořádně očekávaná délka jeho života (nikoliv absolutně, ale relativně).* Zatímco šestileté děvče se má dožít dalších 68,9 roku, má sedmileté před sebou 67,9 roku života, devítileté 66 let. To znamená: tři prožité roky — od šesti do devíti — téměř nic nezměnily na celkové očekávané délce života, která činí přibližně 75 let.



V mládí a ve středním věku se prožité roky života plně „odpisují": zbývající délka života osmiletého dítěte je o dva roky kratší než šestiletého prostě proto, že již žilo o dva roky déle. Dosažené narozeniny ve vysokém věku jsou naproti tomu do jisté míry zhodnoceným dobropisem: očekávaná délka života 92leté činí ještě $2^1/_2$ roku; když dosáhne 94 let, může se dožít ještě více než dvou let; dosáhne-li 96, může se dožít ještě téměř dvou let.

Osmdesátiletá žena má naproti tomu ještě očekávanou délku života asi 6 let, jedenaosmdesátiletá 5,6 roku a triaosmdesátiletá 4,9 roku. Dožitím tří let se tedy zvýšila očekávaná celková doba života ze zhruba 86 let na téměř 88 let! Očekávaná délka života jedenadevadesátileté ženy je pouze 2,8 roku, osmadevadesátileté však stále ještě 1,8 roku. (Viz obr. na str. 268.)

K průměrné očekávané délce života novorozených je třeba ještě dodat, že ve skutečnosti je vyšší než vypočítaná nebo, jinak řečeno, uvedená střední hodnota klame. Pro jednotlivce je vzhledem k délce života nejzajímavější hodnota mediánu: v jakém věku zemřela polovina mého ročníku? Tento střední věk je však znatelně vyšší než aritmetický průměr věku úmrtí, kalkulovaný pojistnou matematikou, jak bylo již krátce uvedeno.

Vysvětlení je prosté. Předpokládejme, že např. z deseti osob stejného ročníku narození zemře jedna ve věku 12 let a dalších devět v těchto letech: 46 — 54 — 63 — 67 — 69 — 71 — 73 — 79 — 86. Celkový součet činí 620 let života. Průměr 62 let je „očekávaná průměrná délka života" v době narození. Polovina těchto osob však dosáhla vyššího věku než 67 let, a protože jde o sudý počet osob, činil by medián $\frac{67 + 69}{2} = 68$.

Dříve byl tento rozdíl mezi aritmetickým průměrem a mediánem často velmi výrazný. Dnes, kdy je velmi málo případů předčasné smrti (většinou jde o oběti nehod), není tak znatelný. Přesto však dnes činí „očekávaná průměrná délka lidského života" živě narozených chlapců v NSR 67,5 roku, ale polovina z nich dosáhne asi 71. roku věku. (Po prvním roce života se však tento rozdíl silně zmenšuje. Víme již totiž, že průměrná délka lidského života jednoletého je vyšší než novorozeného; medián však není tímto posunem dotčen: první část oné poloviny, která se nedožije více než jedenasedmdesáti let, již zemřela.)

Veškerá pořadí vymírání a tabulky úmrtnosti mají však vážnou vadu. Nemohou v dostatečné míře přihlížet k *hygienickým a zdravotním poměrům,* nyní tak odlišným ve srovnání s minulostí. Kdo má dnes 70 let, ten se narodil těsně po přelomu století, dětství prožil pod hrozbou životu nebezpečných dětských nemocí, zažil dvě světové války, protrpěl dvě poválečná období. Očekávanou délku jeho života lze však vypočítat jen na podkladě údajů úmrtnosti ještě dřívější generace. Stejně tak nelze s jistotou říci, že nastupující generace (třeba dnešních padesátníků) bude mít v sedmdesáti letech věku pro zbytek svého života stejné podmínky jako dnešní sedmdesátníci.

Kdyby se rakovina stala vyléčitelnou, ztratí všechny úmrtnostní tabulky během několika let svou vypovídací schopnost. Kdyby se však naopak třeba ukázalo, že znečistění ovzduší má trvalejší škodlivé účinky na organismus, než se dnes domníváme, musely by být tabulky opraveny v opačném směru. Jisté je jedno: v posledních dvaceti letech v NSR vzrostla délka života novorozeňat mužského pohlaví jen minimálně.

u patnáctiletých zůstala stejná, u 25 až 75letých se nepatrně snížila. Teprve za dalších deset až dvacet let bude možno posoudit, zda tato čísla nesignalizují neblahý vývoj. Velice pochybná se však zdá prognóza, podle níž je třeba během dvaceti let počítat v NSR s průměrnou délkou života 85 let. To by totiž vyžadovalo nejen zvítězit nad rakovinou, ale i podstatně snížit počet dopravních nehod, jejichž růst znatelně snižuje průměrný věk mužů. Pro ženy to zatím neplatí: délka jejich života vzrostla ve všech věkových skupinách, jejich předstih proti mužům se výrazně zvýšil.

Dovětek k populační explozi. Heslo „*nejdříve ženy a děti*", které platí pro všechny záchranné akce, nemá snad svůj původ ani tak ve vrozené galantnosti a rytířství mužů, jako spíše ve správném pocitu, že další život národa, rodiny, kmene zajišťují ženy a děti, nikoliv muži nebo staří lidé. Proto absolutní počet obyvatelstva není výchozím bodem pro očekávanou porodnost. Tu je možno vypočítat pouze podle počtu žen v plodném věku. (Dokonce tak těžké ztráty mužů, jaké byly ve druhé světové válce, se v poválečné době projevily ve statistice porodů jen nepatrně.) Někdy vyslovovaný názor, že s ohledem na explozívně rostoucí počet obyvatelstva světa by se neměl také ještě všemi prostředky geriatrie prodlužovat život starých lidí o měsíce nebo dokonce léta, je nesprávný ve dvou směrech: jednak má lékařství na tomto úseku jen relativně malé možnosti, jednak se tím dlouhodobě mnoho nedosáhne.

8.41 Morální statistika, aneb: „Utíkejte! Všechno je prozrazeno!"

Jedna humorná historka z Balkánu vypravuje o muži, který se chtěl pomstít hostinskému. Jednou večer, když byla hospoda jako obvykle plná, vrazil udýchaně dovnitř s významnými posuňky a hlasitě volal: „Utíkejte! Všechno je prozrazeno!" Všichni hosté se v divokém spěchu rozprchli a nezůstal nikdo, kdo by zaplatil útratu.

Na první pohled by nemuselo být jasné, co má tato anekdota společného se statistikou. Poslechněte si však k tomu druhý, velmi podobný příběh, který uvádí Arnold Schwarz ve své knize „*Zahlen beweisen*" (Čísla dokazují): „Před několika desítkami let byla celá Paříž znepokojena velkým počtem anonymních dopisů, které měly vzápětí za následek desítky rozvodů, a velkým počtem anonymních udání, která v mnoha případech končila sebevraždami. V činnosti byla jakási ďábelská moc, jež se zdála být zasvěcena do nejintimnějších tajností stovek osob... Policie zahájila rozsáhlé vyšetřování a nakonec se dostala k domovnici, která byla autorkou dopisů. Přiznala, že každé ráno zabodla do stránek pařížského adresáře pletací jehlici. Když přišla na nějakou firmu, napsala jí, že její účetní se již léta dopouští zpronevěry. Jestliže přišla na nějakou domácnost, psala manželce, že manžel je jí nevěrný. Byla to metoda náhodného výběru v oblasti morální statistiky podle všech pravidel."

Nabodávání obsáhlého seznamu nebo dlouhého textu jehlicí pro náhodný výběr je mimochodem prastará, dokonce velmi doslovná metoda náhodného výběru. Pochází z okrajové oblasti sta-

Také z adresáře a telefonního seznamu je možno vytvořit „vzorek". Není to sice většinou vědecky správné, ale lze tím dosáhnout pozoruhodných výsledků, jak ukazuje výběr jedné pařížské domovnice, o níž pojednáváme v oddílu o morální statistice.

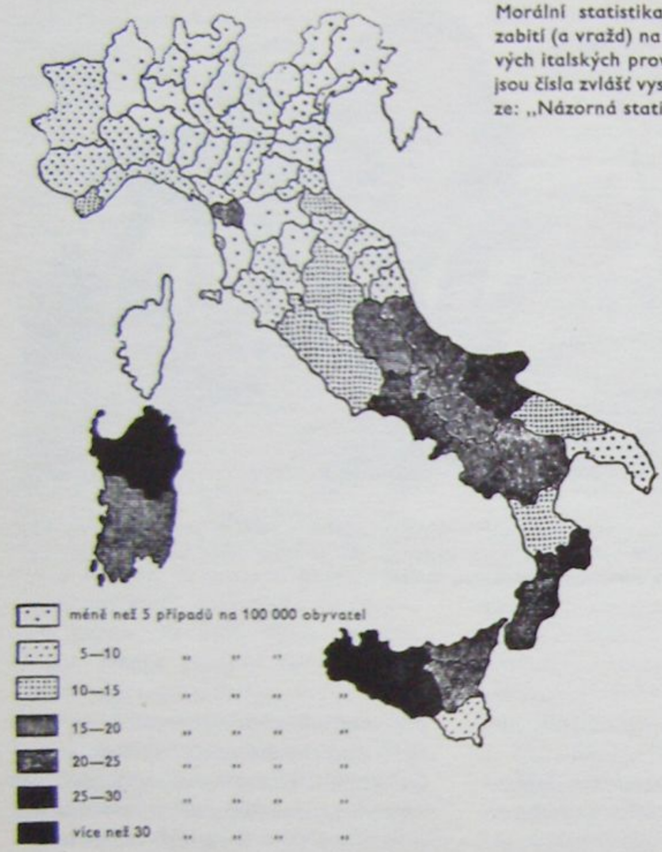
tistiky, v níž nelze přesně oddělit důvěru v osud od víry v moc čísel.

John Wesley, zakladatel církve metodistů, dokonce říká, že tímto způsobem hledal v bibli odpověď na otázku, zda je Boží vůle, aby se oženil. Odpověď byla záporná a Wesley zůstal svobodný.

Nechceme však zdůrazňovat uvedenou techniku náhodného výběru, nýbrž říci několik slov o „morální statistice". Zdá se, že tento výraz pochází z třicátých let 19. století od Francouze Gherryho, který napsal „*Essay sur la statistique morale de la France*" (Úvaha o francouzské morální statistice). Morální statistika se zabývá těmi údaji, které umožňují zpětný úsudek o mravech, právních názorech a morálce obyvatelstva. Doba jejího rozkvětu spadá do druhé poloviny minulého století. Dnes se tohoto výrazu užívá jen zřídka, ale údaje, které jsou pro něj charakteristické, se různými způsoby stále ještě zjišťují.

Do morální statistiky se mimo jiné zahrnují pozorování o sebevraždách, o zločinnosti, o porodech mimo manželství, o počtu odsouzených a vězňů, ale samozřejmě i anketní šetření, jako např. statisticky nečistě zpracovaná tzv. „*Kinsey-Report*" (Kinseyova zpráva) a jiná zjišťování o sexuálním chování lidí.

Potíže morální statistiky jsou zřejmé z již uvedených anekdot: *temné číslo* mnoha „morálních deliktů" je velmi vysoké. Reprezentativní dotazování o četnosti manželské nevěry musí skoro nevyhnutelně vést k rozporným a nespolehlivým údajům, z nichž závěry lze dělat jen s velkou opatrností. Naproti tomu existují určité údaje, které zcela dobře umožňují porovnání. K nim



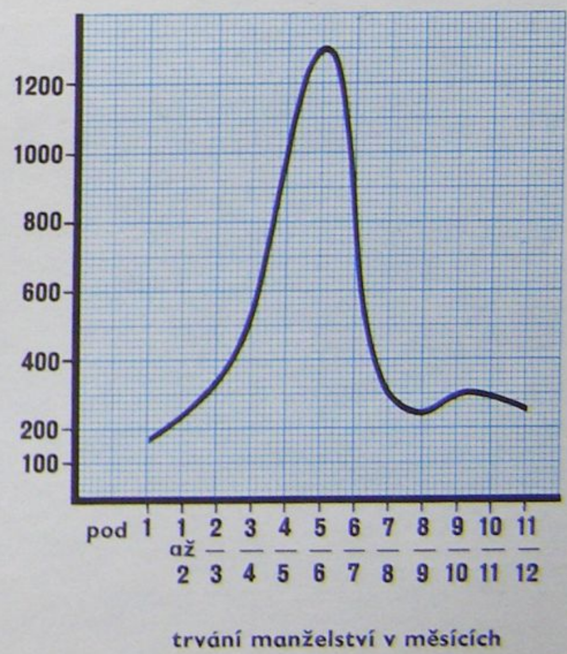
Morální statistika z přelomu století: počet zabití (a vražd) na 100 000 obyvatel v jednotlivých italských provinciích. Na Sicílii a Sardinii jsou čísla zvlášť vysoká. (Uvedeno u A. Schwarze: „Názorná statistika".)

především patří statistiky dat narození prvního dítěte. Na základě těchto údajů lze například velmi dobře dokázat, že manželství se ve velmi vysokém procentu uzavírají teprve tehdy, když je zcela jasně lékařsky prokázáno těhotenství nevěsty. Takto např. vypadá statistika *„narození během prvních dvanácti měsíců manželství"* v libovolně vzatém roce z nedávné minulosti ve Vídni — a Vídeň není o nic víc „nemorální" než jiná středoevropská velkoměsta:

Trvání manželství v měsících	Narození
(méně než 1)	168
1	233
2	340
3	578
4	965
5	1248
6	744
7	303
8	249
9 až 11	832

V prvním roce manželství se z celkového počtu 5660 dětí narodí jen o něco více než 800 dětí, které byly počaty v manželství. Velká kumulace porodů v 5. až 7. měsíci manželství (tedy po 4 a 6 měsících trvání manželství) jasně ukazuje, jak rychle po zjištění těhotenství byla zorganizována svatba. Vzhledem k tomu, že ve zvoleném roce bylo uzavřeno asi 14 500 sňatků (včetně nového sňatku ovdovělých apod.), ukazuje „morální statistika", že ve třetině manželství se dítě již čekalo. Na tom se však mimochodem, bez ohledu na vlnu sexu nebo existenci antikoncepčních prostředků, během doby změnilo jen velmi málo. Vcelku je však při těchto porovnáních třeba velké opatrnosti. Tak např. vysoký vzrůst zaznamenala jako důsledek rychlého rozvoje *automobilismu „kriminální statistika"*, vzhledem k tomu, že přibyla zcela nová skupina přečinů; výraz „národ potrestaných" pochází z vysoce rozvinutých zemí. Jestliže se dopravní trestní právo jediným škrtem pera „odkriminalizuje" například tím, že se trest bude ukládat jen správní cestou, dojde k prudkému poklesu kriminality. Stejně tak se snižuje i počet mravnostních deliktů, zejména když se některé již nepokládají za trestný čin (např. homosexualita apod.).

Proto lze nanejvýš víceméně porovnávat kategorie jako „loupež" nebo „vražda", ale dokonce i zde je možné postupovat nesprávně, např. proto, že trestní právo — vědomě nebo nevědomě — jinak vymezuje vraždu a zabití, loupež a krádež. Dříve než začneme porovnávat, je nutno pečlivě si ověřit,



„Morální statistika" ukazuje, že v době sňatku je čáp často již na cestě. Ve čtvrtém až sedmém měsíci manželství se narodí podstatně víc dětí než v devátém a dvanáctém.

zda právní pojmy skutečně obsahují stejnou skutkovou podstatu. Avšak i potom jde spíše o kulturně historický nebo sociologický popis prostředí než o matematickou statistiku.

Pro téměř celou oblast morální statistiky je však přesto charakteristické zjištění, které kdysi učinil Quételet o zločinech a Emile Durkheim o sebevraždách: pokud všeobecné životní podmínky zůstávají přibližně stejné, ukazuje jejich četnost překvapující rovnoměrnost. Zjevné změny nastávají až tehdy, když dojde k trvalejší změně životních poměrů. Přitom účinky jsou někdy překvapující: např. počet sebevražd, jak ukazují zkušenosti, poněkud klesá v době války.

To, co na morální statistice zasluhuje pozornosti, patří z převážné části do oblasti sociologie, a proto můžeme čtenáře odkázat na odborná sociologická díla. Kriminalita mládeže, delikty pod vlivem omamných látek, sdružování se do tlup, opilci za volantem a mnohé jiné dnes tak často diskutované aktuální problémy patří do „morální statistiky". Jak nepozorovatelný je přechod od morální statistiky ke statistice obyvatelstva, jasně ukazují např. antikoncepční prostředky.

Vraťme se však ke statistickým metodám, o kterých nebylo v 19. století ani potuchy, které však jsou pro hladký běh moderní civilizace často důležitější než morálně statistické výpočty. Tak *„teorie hromadné obsluhy"* se zabývá problémem, který je i matematicky velmi složitý, studiem optimální obsluhy zákazníků (v obchodech i úřadech, v dopravních prostředcích i lokálech).

8.5 Čekání na autobus — teorie hromadné obsluhy

Na zastávce autobusu čeká již pět lidí, další autobus tedy už musí přijet každou chvíli, neboť intervaly v tuto denní dobu činí, podle oznámení na vývěsce, 10 minut. Minuta za minutou utíká, zástup čekajících stále roste. Osm, deset, dvanáct minut uteklo, polohlasně se reptá, jeden netrpělivec dokonce již mezitím zamával na projíždějící taxík. Konečně se vzadu v zatáčce objevuje autobus, blíží se, zastavuje a je přirozeně nacpaný k prasknutí. Tlačenice u výstupu, tlačenice u vstupu, průvodčí netrpělivě gestikuluje a volá: „Za námi jede další!" Mezitím co se posledním pasažérům podařilo nacpat se dovnitř, ze zatáčky skutečně vyjíždí další, prázdný autobus stejné linky.

Podobných příkladů tohoto druhu je v moderním světě víc než dost. Jsou to fronty čekajících u hraničních závor v turistické sezóně, nebo zejména u úředních přepážek, kde si potřebují něco vyřídit. Dále je tu telefon, který je stále obsazen v důsledku nahromadění volajících, kteří stojí abstraktní frontu. Je to, jsou to — ale co to má všechno společného se statistikou? Víc než by se dalo na první pohled předpokládat, ačkoliv hned poznáme, že v daném případě narážíme na obtížnou speciální a okrajovou oblast statistiky a počtu pravděpodobnosti — na *„teorii hromadné obsluhy"* nebo *„teorii front"*.

Používané výpočetní postupy jsou obtížné, vlastní výpočty se provádějí většinou za pomoci výpočetní techniky. Nemůžeme popisovat výpočetní metody v jednotlivostech. Vzhledem k významu shlukování a front, který mají v našem životě, a vzhledem k tomu, že se v tomto oddíle zabýváme především rolí statistiky v každodenním životě, řekněme si přece jen několik slov o tom nejpodstatnějším z teorie hromadné obsluhy.

Jak dochází k shlukování? Velmi prostou příčinou vzniku shlukování je trvalý nedostatek kapacit, *převaha poptávky nad nabídkou*. Jestliže na jedné autobusové lince je k dispozici pouze 40 míst, zatímco potřeba činí 70, je trvalá tlačenice nevyhnutelná. Jestliže na poštovním úřadě je na počátku měsíce otevřena pouze jedna přepážka pro příjem peněz, narůstá tam fronta a takový stav trvá pár dní. Je-li telefonní centrála zařízena pouze na tisíc spojení, ozývá se ze sluchátka trvale obsazovací tón.

Na tento druh shlukování je těžké jít se statistikou. Musí se zařadit další autobus, otevřít další přepážka — a pak se všechno hladce vyřídí. Je nutno např. dbát o to, aby přepážka byla otevřena osm hodin, i když jednání se stranami by vyžadovalo jen 6 hodin (např. 180 zákazníků po 2 minutách). V daném případě má pak úředník ještě dost času, aby vyřídil i ostatní práce, snědl své obložené chlebíčky a popovídal si s kolegy — a nikdo nemusí čekat.

Opravdu nikdo nemusí čekat? Ponechme zatím stranou existenci „špiček" a optimisticky předpokládejme, že by dlouhé čekání ve „špičkách" vychovalo zákazníky tak, že by přicházeli k přepážce po celý den pravidelně —

◄ Klient vstupuje do přepážkové haly — půjde přímo k prázdnému okénku, nebo se zařadí do fronty čekajících? „Teorie obsluhy" se pokouší vypočítat optimální podmínky pro provoz u přepážek, v obchodech, při využití sedadel v kinech, hostincích a v dopravních prostředcích.

asi 20 až 25 lidí za hodinu. Přicházejí však náhodně rozptýleni — jednou tři krátce za sebou, jindy deset minut nikdo, jak už to bývá. Čekat však nemusí nikdo.

Ale copak, copak? Otevřete dveře a u „vaší" přepážky už zase stojí 4 lidé. Jak je to možné? Pomůže nám vzpomínka na Poissonovo rozdělení, s kterým jsme se seznámili již dříve (srovnej oddíl 3. 4) — jen se nesmí „vzácný jev" vykládat příliš doslovně. Otázka pak např. zní: Jaká je pravděpodobnost, že zákazník najde přepážku volnou, když k přepážce přijde za hodinu v průměru 22,5 zákazníka a každý z nich tam potřebuje k jednání 2 minuty? Myšlené východisko je jednodušší, než se zpočátku zdá. Z každých 60 minut doby, po kterou je přepážka otevřena, má úředník průměrně 15 minut volno, což již ostatně víme. To však neznamená, že během těchto 15 minut není nikdo odbaven, že jen 15 minut je přístup volný a bez čekání. To však znamená: kdykoli — náhodou — přijdu, nebudu moci ve třech ze čtyř případů (45 z 60 minut) ihned k přepážce, nýbrž nejméně jedna osoba bude přede mnou. Je možné, že přijdu právě v posledních vteřinách onoho úředního jednání, není však také tak nepravděpodobné, že čekají tři, čtyři nebo ještě více lidí. A jestliže nový přírůstek čekajících překročí dvouminutový rytmus vyřizování, fronta narůstá...

Jaký druh front můžeme u této přepážky očekávat? Celková doba otevření přepážky je 8 hodin, tzn. 480 minut; přijde 180 zákazníků a každý potřebuje 2 minuty, tj. celkem 360 minut. Doba vyřizování záležitostí jednoho zákazníka je tedy $360/480 = {}^3/_4 = 0{,}75$. Označme toto číslo λ, abychom znázornili strukturální souvislost s Poissonovým vzorcem. Nyní uvažujme: čtvrtina všech zákazníků najde přepážku volnou a v tom okamžiku, kdy je někdo obsluhován, tvoří potenciální čelo fronty, „fronty z jednoho člena". Pro výpočet této pravděpodobnosti použijeme následující varianty Poissonova vzorce:

$$p_x = e^{-\lambda x}\, \lambda^{x-1} \frac{x^{x-2}}{(x-1)\,!}$$

Za $x = 1$ dostaneme:

$$p_1 = e^{-0{,}75}\,(0{,}75)^0 \frac{1^{-1}}{(1-1)!} =$$

$$= \frac{1}{e^{0{,}75}} \cdot 1 \cdot \frac{1}{1} = \frac{1}{e^{0{,}75}} = 0{,}47.$$

Z celkové pravděpodobnosti (= jistoty; $p = 1$), že někdo utvoří čelo fronty a vznikne fronta — i když jen jednočlenná — připadá 0,47 na možnost „jen jeden člen". Z celku ($180/4 = 45$) zákazníků, kteří najdou přepážku volnou, odejde asi polovina (22), aniž by za nimi vznikla fronta.

Dále:

$$p_2 = e^{(-0{,}75)2}\,(0{,}75)^1 \frac{2^0}{(2-1)!} =$$

$$= \frac{1}{e^{1{,}5}} \cdot (0{,}75) \cdot 1 = \frac{0{,}75}{4{,}5} = 0{,}17.$$

Ze 17 % ze 45 zákazníků (7 až 8) vznikne dvoučlenná fronta: za ně se postaví 1 osoba a tou miniaturní fronta pak zanikne.

Za zbývajícími 15 až 16, tvořícími „čelo fronty", vznikne naproti tomu asi ve 3 případech trojčlenná, ve dvou případech čtyř nebo pětičlenná fronta. Stále ještě zbývá asi 8 „čel fronty", každá po 6, 7 a 8 členech (přibližně),

Nemusí to být vždy špatná organizace, jestliže se u jednoho okénka (vlevo) vytváří fronta, zatímco u jiných okének (vpravo) se přijde hned na řadu. Fronty vznikají často jen jako důsledek náhodného nepravidelného příchodu klientů.

ale během dne může vzniknout docela dobře také jedna dvanáctičlenná fronta:

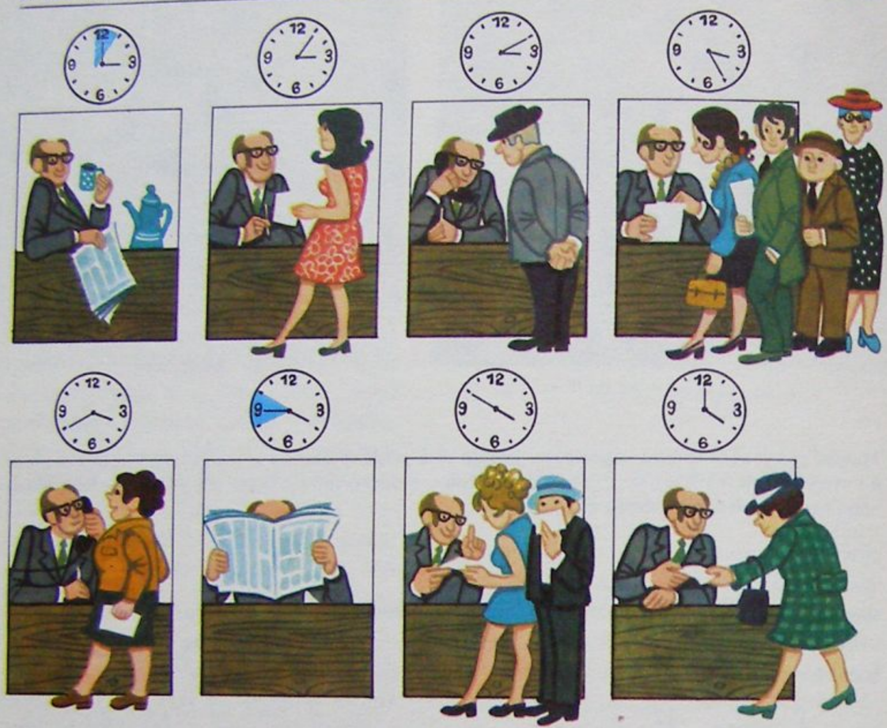
$$p_{12} = \frac{1}{e^9} \cdot (0{,}75)^{11} \cdot \frac{12^{10}}{11!} \approx$$

$$\approx \frac{1}{8100} \cdot (0{,}04) \cdot \frac{62 \text{ mld}}{40 \text{ mil.}} \approx 0{,}0075$$

Při tomto výpočtu dostáváme sice jen 0,75 % zákazníků přepážky, kteří tvoří čelo jedné dvanáctičlenné fronty, avšak pomyslíme-li, že pravděpodobnosti jedenácti nebo třináctičlenné fronty jsou přibližně stejné, je pro celkovou pravděpodobnost „jedné z těchto tří délek" přece jen P_{11}, P_{12} nebo $P_{13} \approx 0{,}02$ (= 2 % ze 45 osob, tedy jedna). Tímto způsobem mohou být již stejně velmi malé pravděpodobnosti pro ještě delší fronty sečteny a ukáže se, že některá fronta může mít až 24 osob. A to vše za situace, kdy přepážka je otevřena déle, než by bylo třeba k vyřízení záležitostí a ačkoliv se návštěvníci neshromažďují „ve špičkách", nýbrž přicházejí náhodně rozptýleni.

Tento výsledek se jeví tak nepravděpodobný, že jsme v pokušení provést jeho kontrolu. Můžeme sledovat osud našich 180 jednotlivých zákazníků a zjistíme (zčásti z dosavadních údajů, zčásti z údajů dalších, avšak zjištěných stejným způsobem):

		Celkem
22	zůstanou sami	22
7 až 8	vytvoří dvoučlennou frontu	14 až 16
asi 3	tříčlennou frontu	9
asi 2	čtyřčlennou	8
2	pětičlennou	10
1	šestičlennou	6
1	sedmičlennou	7
1	osmi nebo devítičlennou	8 až 9
1	deseti nebo jedenáctičlennou	10 až 11
1	12, 13 nebo 14členou	12 až 14



Fronty nevznikají jen tam, kde je kapacita vyřizovacího místa příliš malá, ale také tam, kde dochází k náhodnému rozdělení zákazníků, cestujících atd., tedy k nepravidelnému vytížení. U této přepážky nemá někdy úředník na práci nic, ale jindy se přesto vytvoří fronta — i když by se předpokládalo, že každý klient potřebuje k vyřízení své záležitosti stejně času.

1	15 až 18člennou	15 až 18
1	19 až 23člennou	19 až 23
1	24 až 30člennou	24 až 30
		164 až 183

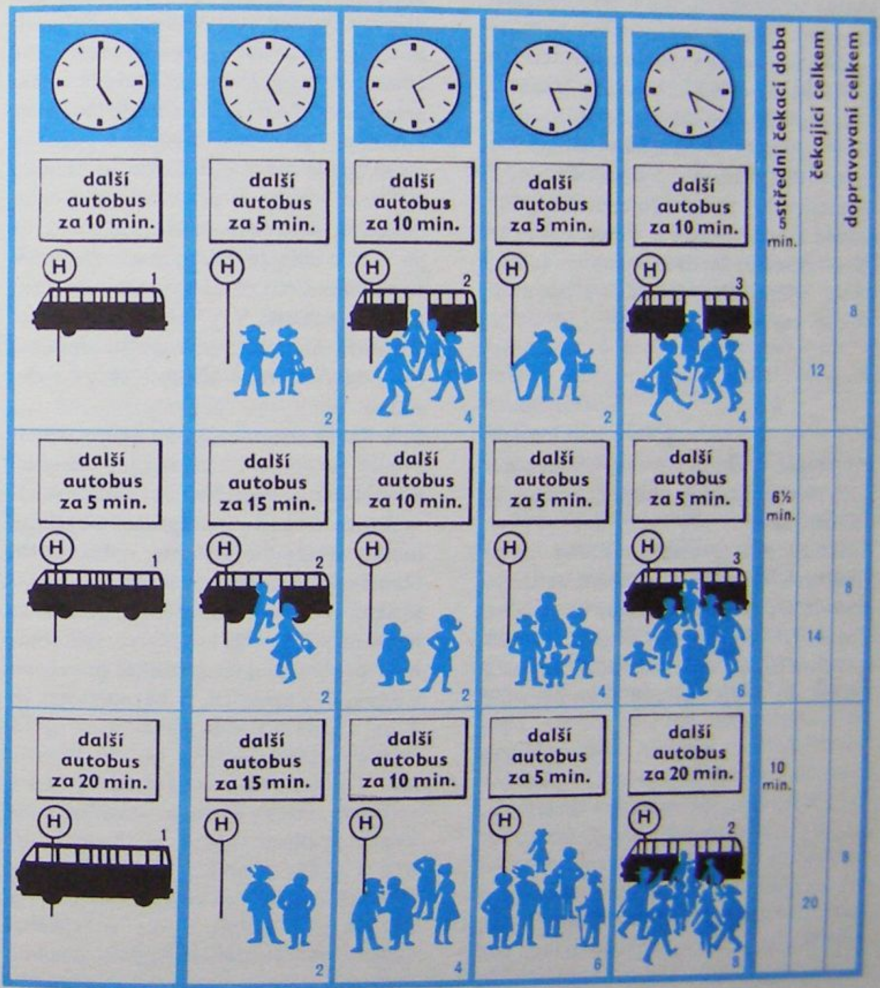
Tento model je samozřejmě možno ještě zlepšit, jestliže se zcela reálně předpokládá, že se nikdo nepostaví do 15členné fronty, nýbrž přijde raději ještě jednou. Již tento nejzákladnější výpočet je však, jak jsme viděli, dosti namáhavý. Vyšší postupy jsou ještě složitější.

Vraťme se od fronty na autobusovou zastávku a zamysleme se nad zdejším shlukováním. Autobus, který přijel opožděně, byl sice plný k prasknutí, ale další byl prázdný. Kapacita sedadel na lince byla tedy dostatečná, závadné

Stejnoměrné intervaly znamenají, že průměrná doba čekání činí polovinu délky intervalu. Nepravidelné intervaly všeobecně prodlužují čekací dobu. Jestliže je interval tak dlouhý, že zpožděný autobus přijede současně teprve až s druhým následujícím, který jede podle jízdního řádu, vzroste čekací doba na dvojnásobek. ▶

bylo jen rozdělení. Zde do sebe zapadají nebo se překrývají dvě pravděpodobnosti. Na jednotlivých zastávkách se zpočátku odehrává totéž co na poštovním úřadě: cestující přicházejí náhodně rozděleni, jednou krátce po sobě, pak zase vůbec nikdo. V tomto případě k tomu přistupuje ještě druhá série náhodných jevů: hustota provozu, světelné signály atd.

Nakonec se zpočátku nezávislé jevy navzájem vyrovnají: autobus zůstane stát na červenou, tím přibývá čekajících na nejbližší stanici. Jejich odbavení (nástup, později výstup) trvá déle, doby stání jsou nadprůměrné, jízdní řád již



nelze dodržet, na dalších zastávkách se nahromadí ještě více čekajících atd. Touto hrou na nervy se nebudeme vůbec zabývat, nýbrž jen zkoumat účinek nepravidelných intervalů na uživatele autobusu. Jestliže stanovený interval je 10 minut, může „náhodný cestující" (který např. denně jezdí určitou linkou, ale nemůže se řídit jízdním řádem) počítat s průměrnou čekací dobou 5 minut — poloviční dobou intervalu. Jestliže však naproti tomu autobusy z jakýchkoliv důvodů nejezdí v pravidelných intervalech, nýbrž střídavě v intervalech pěti a patnáctiminutových, vpadne ve třech ze čtyř případů do delšího intervalu (průměrná čekací doba 7,5 minuty) a jen jednou do kratšího (průměrná čekací doba 2,5 minuty). Tím se však vypočítá pravděpodobná čekací doba náhodného cestujícího za delší období takto:

$$\frac{3 \,.\, (7{,}5) + 1 \,.\, (2{,}5)}{4} = 6{,}25.$$

Při stejné hodinové frekvenci nabízených míst bude tedy muset čekat o $1^1/_4$ minuty déle.

Autobusy sice zpravidla vyjíždějí v pravidelných intervalech, nicméně neustále dochází k posunům intervalů v důsledku uvedených dopravních situací s jejich následnými účinky. Podívejme se ještě rychle na krajní případ, kterým jsme tento oddíl uvedli. Teoretický interval 10 minut se prodloužil tak, že přijel již další autobus — z dvakrát deseti minut je dvacet minut plus 0 minut. Dva jednotlivé autobusy prakticky splynuly v jeden velkoprostorový superautobus, který přijíždí jen každých 20 minut.

Podobné frekvenční problémy nejsou samozřejmě omezeny pouze na úřady, dopravní prostředky a jiná veřejná zařízení. V těchto případech se jen zvlášť nápadně projevují, protože „velké číslo" si přímo vynucuje statistické zpracování a výpočet na základě teorie pravděpodobnosti. V zásadě to platí dokonce i pro malý obchod s potravinami, jehož šéf se ptá, kolik asi banánů má mít v sobotu pro své zákazníky: náhodná poptávka může způsobit, že již v 11 hodin je vyprodáno, ale také, že třetina zboží zůstane ležet a do pondělka se zkazí. V případě restaurací, biografů atd. je použitelnost modelů hromadné obsluhy zřejmá ještě daleko více.

Bohužel se zde nemůžeme dále zabývat těmito otázkami, které již zasahují do oblasti operačního výzkumu, optimalizace a teorie her, protože bychom museli nadlouho opustit naši vlastní oblast — statistiku v užším slova smyslu. To však nemůžeme, neboť se musíme ještě blíže seznámit s ústředním problémem statistické práce — s naukou o vztazích, o korelacích.

9 Post hoc ergo propter hoc

9.1 Vztahy, příčiny, účinky

V oddílu 3.1 jsme mluvili o diskrétních a spojitých znacích a o relativních četnostech. Mezitím jsme se vždy znovu zabývali rozdělením takových znaků ve výběrových nebo základních souborech. Pravděpodobně se již také objevil výraz „*proměnná*", kterým se nyní budeme stručně zabývat, dříve než přejdeme ke zkoumání vztahů mezi takovými „*proměnnými*".

„*Proměnnou*" je to, co může nabývat různých tvarů nebo hodnot. Je to „proměnná", avšak nikoliv s ohledem na jednotlivé nositele znaku, protože pak by „proměnnou" byla např. barva vlasů určité ženy, která může být bezmála každý týden jiná. Nikoliv, „proměnná" se vztahuje na statistická zjišťování a označuje předmět pozorování, nebo přesněji: hodnoty, jichž zkoumaný znak nabývá. Je to *náhodná proměnná*, která se nemění přesně předvídatelným způsobem, nýbrž která nabývá určitých hodnot jen v závislosti na náhodě.

Zkoumáme-li 100 osob z hlediska jejich tělesné výšky, příjmu, váhy, pohlaví a barvy vlasů, dostaneme údaje o pěti proměnlivých znacích — tři *kvantitativní, které se týkají spojitých znaků* (tělesná výška, příjmy a váha) a dva *kvalitativní, týkající se diskrétních znaků* (pohlaví a barvy vlasů). Zde se nebudeme podrobně zabývat problematikou často poněkud nepřesného rozlišování mezi diskrétními a spojitými znaky, mezi kvantitativním a kvalitativním, mezi homográdním a heterográdním. V tomto případě jde jen o vyjasnění pojmu „proměnlivý znak". Doposud jsme se téměř výlučně zabývali jediným proměnlivým znakem: počítali jsme podíl dobrých ořechů v pytli (proměnlivý znak: jakost ořechů), diskutovalo se o příjmech obyvatelstva ve Zbohatlíkově (proměnlivý znak: příjem), analyzovali jsme pyramidu obyvatelstva (proměnlivý znak: stáří, případně přebytek porodů). Někdy jsme časově propojili i časově odlehlé měrné hodnoty navzájem a vytvořili časovou řadu.

Veškeré dění však je — a tím neříkáme nic nového — *nejrůznějším způsobem spojeno*, přičemž některé souvislosti se vyznačují elementární jednoduchostí (např. souvislost tělesné výšky a váhy), jiné jsou naopak složité, těžko prokazatelné, dokonce velmi sporné — např. otázka, zda datum narození má vliv na charakter a osud člověka. Nebo, abychom vybrali aktuálnější a vědečtější téma: existuje souvislost mezi proměnlivým znakem „znečišťování ovzduší" a proměnlivým znakem „úmrtnost na rakovinu", a jestliže ano, jak těsná je tato souvislost?

Dokonce jsme již vyjádřili *vztahy mezi proměnnými*: přebytek porodů je vypočitatelný ze dvou proměnlivých znaků: „porodnost" a „úmrtnost". Reálný příjem udává poměr mezi proměnnými „zvýšení příjmů" a „zvýšení cen". V této kapitole se však chceme vypořádat s těmi významnými a důležitými problémy, kdy se klade otázka po *vztahu mezi dvěma nebo více proměnnými* a u nichž jde o číselné vyjádření tohoto vztahu, aby se mohl vyznačit jeho význam a řádová velikost. Takové vztahy mají důležitou úlohu i v průzkumu trhu a veřejného mínění, protože v daném případě často jde o zjištění zájmu o určitý výrobek v závislosti na *věku*, *příjmu* nebo *úrovni vzdělání* dotazovaných. Tak např. počet knih, které někdo má, asi souvisí se všemi třemi uvedenými znaky, přičemž k tomu ovšem mohou přistupovat ještě další, těžko měřitelné rozhodující faktory, jako např. *inteligence*.

V případě otázek týkajících se průzkumu trhu většinou postačí hrubé roztřídění podle příjmové, věkové a zaměstnanecké skupiny: starší lidé kupují spíše A, mladší spíše B; do přesného kvantitativního výpočtu souvislostí (např. na 10 DM měsíčního příjmu je v knihovně o 0,2 knihy více) se raději nepouštíme.

Zjišťování např. příjmu se týká jen *jednoho* proměnlivého znaku. Zajímám-li se současně o věk příjemců, dostávám *dvě* proměnné. Vytvářím *dvourozměrnou sérii pozorování (angl. bivariate)*. Jestliže se dále ptám na počet m² bytu, na měsíční spotřebu elektrického proudu atd., jde o *vícerozměrná pozorování* (angl. *multivariate*).

Uvést do vzájemného spojení vícerozměrné (multivariantní) řady pozorování, včetně zpracovaných údajů, v jejich vzájemných souvislostech a při jednotlivých vlivech je úkolem tzv. *vícerozměrné analýzy*. Za tímto působivým módním slovem je skryt obtížný početní problém, se kterým, nebo alespoň s jeho podstatou, se seznámíme, až budeme mluvit o analýze rozptylu (oddíl 10.3).

Nejdříve je však třeba pochopit podstatu korelačního počtu: *dvě nebo více proměnných má být prozkoumáno* z toho hlediska, *zda mezi nimi není prokazatelná souvislost*. Přitom jde nejdříve — a to je třeba neustále zdůrazňovat — jen a jen o *matematické souvislosti*, které nelze zaměnit za *vztah příčiny a následku*. Korelace může poskytnout důležité poznatky o příčinách nebo účincích, může však jít také o náhodnou korelaci, zdánlivou korelaci nebo prakticky zcela bezvýznamnou korelaci. Dokonce je to tak, že nelze očekávat úplnou matematickou nezávislost dvou proměnných znaků ani tehdy, jestliže je zcela jisté, že neexistují vůbec žádné souvislosti. Zdánlivé a náhodné korelace se mohou velmi lehce vyskytnout zejména u výběrových souborů poměrně nepatrného rozsahu.

Korelační počet je matematický postup, který činí číselná zjištění o souvislostech dvou měrných hodnot. Druh těchto měrných hodnot je pro korelační počet zcela bezvýznamný. Pokud neexistuje nějaká souvislost, bude to jasně vidět z výsledku výpočtu. Více nebo méně zřetelná korelace ještě však neříká, že existuje vnitřní přímá souvislost mezi zjištěnými hodnotami — a už vůbec ne, že existuje příčinná souvislost.

Korelaci je možno hledat mezi vším a v různé míře také mezi vším existuje. Tělesná výška a váha jsou nepochybně v těsné souvislosti, velikost bot a objem hlavy asi také, vnější teplota a spotřeba paliva také, avšak nekonečnou řadu mohou tvořit měrné veličiny, u nichž je pochybné, nikoliv však vyloučené, že mezi nimi existuje nějaká měřitelná souvislost. Je nějaká souvislost mezi manželstvím a rakovinou plic? Mezi vlastnictvím rodinného domku a zájezdy do Itálie? Mezi dobrými známkami z matematiky a tělesnou vahou? Mezi počtem přenocujících cizinců a cenou piva? Mezi vývozem magnezitu a dovozem pomerančů? Mezi počtem opic v městské zoologické zahradě a časem vyhrazeným reklamám na prací prostředky v televizi?

Pokud by byla v těchto pozoruhodných pojmových, případně číselných dvojicích zjistitelná (pozitivní nebo negativní) korelace, zůstává vhodný výklad vyhrazen především ostrovtipu interpreta. Ke korelaci nepatří zásadně nic jiného než dva proměnné znaky, dvě veličiny proměnné ve svých hodnotách. Stále znovu se však vyskytují statisticky jasně prokazatelné souvislosti, které nepochybně nejsou zdánlivými korelacemi, zdají se zcela logické, a přesto mají své skryté *záludnosti*. Mezi ně patří *korelace mezi vzděláním a příjmem*: kdo déle studoval, má větší příjem.

Tak před několika lety bylo zjištěno, že v NSR 45 % zaměstnanců s vysokoškolským vzděláním mělo měsíční čistý příjem 1200 DM; u zaměstnanců s odborným školním vzděláním činil tento podíl jen 6 % a u zaměstnanců se základním vzděláním pouze 2 %. V nízkých příjmových skupinách tomu bylo přesně naopak: jen 9 % vysokoškoláků mělo příjem nižší než 600 DM, naproti tomu spadalo do této skupiny 54 % osob s odborným vzděláním a 76 % se základním vzděláním. Tato sestava má několik drobných závad. Např. se můžeme ptát, kde zůstali abiturienti bez vysokoškolského vzdělání nebo zda do zjišťování byli nebo nebyli zahrnuti i samostatně činní (mluví se poněkud nejasně o „výdělečně činných osobách").

Tyto podrobnosti lze však pominout, protože zatím jde o zásadní věc — všeobecný jev, že vyšší vzdělání má zřejmě za následek vyšší příjem. Tak v USA bylo vypočítáno, že pracovní síla bez školního vzdělání („illiterate", což znamená analfabet) vydělá za 40 let pracovní činnosti 40 000 dolarů, pracovní síla se středním vzděláním již 150 000 dolarů a pracovník s ukončeným vysokoškolským vzděláním dokonce 750 000 dolarů.

Zjistit tuto očividnou korelaci není nikterak obtížné. Již z výběrového souboru několika set osob zcela zřetelně vyplyne, že oba proměnné znaky nejsou na sobě nezávislé, nýbrž jejich vzájemná souvislost je zjistitelná již „pouhým okem". Sestaví se matice s řádky a sloupci, v řádcích se uvede několik příjmových skupin, ve sloupcích tři nebo čtyři stupně vzdělání a již se hromadně objevují dvojice „vyšší vzdělání — vyšší příjem", „střední vzdělání — střední příjem" a „nižší vzdělání — nízký příjem".

V daném případě jde dokonce o velmi zjednodušený případ korelace, protože jedna proměnná může sice mít libovolné množství hodnot (příjem může kolísat od nuly až do miliónů, proto musí být uměle vymezen třídami a středy tříd), pro druhou proměnnou jsou však prakticky možné jen tři až čtyři hodnoty: základní škola, maturita, odborná škola, vysoká škola.

Statistická souvislost — první krok — *je početně jasně prokázána* (aniž bychom prováděli vlastní korelační výpočet) a jeví se — druhý krok — také zcela logická a pochopitelná. Jistě, někdy vydělává pomocný stavební dělník víc než mladý absolvent vysoké školy, ale všeobecně mají vysokoškolsky vzdělaní zaměstnanci (nebo „výdělečně činné osoby") nesporně vyšší výdělkové možnosti než lidé s nižším vzděláním nebo dokonce bez vzdělání. Tedy—třetí krok: *kdo navštěvoval vysokou školu, vydělává proto podstatně více než ten, kdo vysokoškolské vzdělání nemá.*

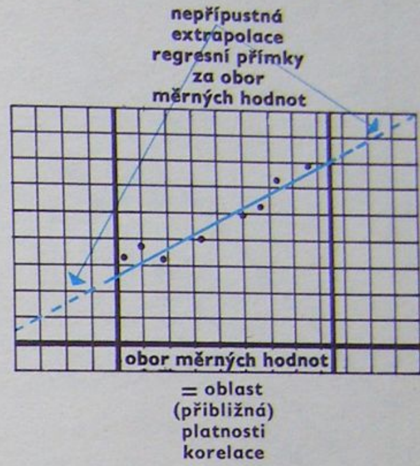
I když tento úsudek vypadá jako nesporný, musíme zvolat: „Pozor!" Co jsme zjistili, je „*post hoc*" — *po* dosažení odpovídajícího školního vzdělání jsou příjmy vyšší. Je to nezbytně také „*propter hoc*" (v důsledku toho)? Znamená to též, že proto a jen proto jsou příjmy vyšší? Například co když často studují mladí, dynamičtí lidé, kteří mají úspěch nikoliv v důsledku studia, ale v důsledku svých všeobecných schopností? Nebo jak to je, když v životě něčeho dosáhnou lidé, o které bylo již od kolébky vždy předem dobře postaráno: zámožný otec financuje studium a také později se pečlivě stará, aby synáček dostal teplé místo! To jsou úvahy, které je nutno brát se vší vážností a jejich oprávněnost by bylo možno objasnit až v další sérii zkoumání.

Tímto tak jednoduše vypadajícím příkladem jsme chtěli ukázat, jaké léčky číhají dokonce i v domněle nejjasnějších případech korelací. Dost často se stává, že další společný, v pořadí už třetí faktor je příčinou dvou jevů nebo úkazů, které nemají přímo nic společného. Když v Maďarsku mrzne, kupují Švýcaři palivo — kde je závislost, kterou by bylo možno korelaci lehce prokázat? Nuže právě v tom, že klimatické podmínky jsou přibližně stejné: ve středních šířkách severní polokoule je od prosince do března zpravidla zvlášť chladno, v Maďarsku jako ve Švýcarsku.

Často se uvádí jako nesmyslná korelace, která úplně nesmyslnou není: po dlouhou řadu let lze prokázat, že je zřejmá souvislost mezi platy pastorů a cenami alkoholu. Je možno z toho soudit, že větší příjmy pastorů se obratem přemění v pálenku? Nikoliv, tou třetí společnou věcí je obyčejná světská inflace — ponenáhlu je všechno dražší, stoupají všechny příjmy a všechny ceny. Někdy jsou to právě takové pošetile vypadající korelace, které umožňují mnohdy získat pozoruhodné poznatky, jako např. tato: Jednou hrál mladý hudební nadšenec na Floridě svým pokojovým květinám ve třech různých místnostech různé hudební skladby, až nakonec objevil, že rostlinám v pokoji, kde hrál beatovou hudbu, se dařilo nejlépe. To snad mohlo být náhodným výsledkem ledabyle uspořádaného pokusu, nebo snad byla ona místnost slunnější či vlhčí. Přece však se zdá, jak konečně vyplývá i z jiných pokusů, že krávy, kterým se pouští hudba, dávají více mléka.

Tak se při korelacích neustále pohybujeme mezi Scyllou a Charybdou — *mezi nebezpečím, že je nedoceníme nebo přeceníme.* Z těchto důvodů jsou korelace oblíbeným rejdištěm demagogů: jasné závislosti se mohou číselně dokázat a k tomu se připojí vlastní výklad: „statistika dokazuje, že...". Statistika však pouze zjistila korelaci, ale neřekla nic o příčině a následku.

V předcházejícím oddílu (viz str. 208) jsme mluvili o matoucím titulku „Šílenství chrání před rakovinou" a ukazovali, že statistika to vůbec nedokazuje. Nyní však musíme říci: jestliže je prokazatelná negativní korelace mezi duševní chorobou a úmrtím na rakovinu (i kdyby byla zcela nepatrná), bylo by nutno dalšími sériemi pokusů zjišťovat vzájemné vztahy mezi různými proměnlivými znaky. Bio-



Za oborem měrných hodnot není již statisticky zjištěná korelace. Regresní přímka nesmí být zpravidla extrapolována — pravidlo, které se často nezachovává a vede pak lehce k omylům.

chemické procesy v organismu, rozdíly ve stravování, uklidňující prostředky běžně používané v léčebných ústavech pro nervově choré, snížená (nebo zvýšená) vzrušitelnost — to by bylo několik proměnlivých znaků, které by měly být ověřeny. Není vůbec vyloučeno, že právě rakovina bude jednou překonána oklikou přes nesčetné korelační výpočty.

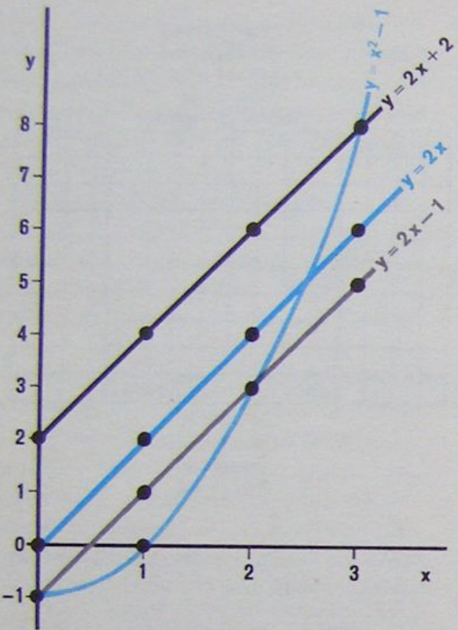
Korelační počet prokazuje svůj význam převážně v oblasti vědeckých experimentů. Jestliže se totiž při pokusu daří udržet v přibližné rovnováze většinu všech účinných faktorů a vždy jen jeden nebo několik málo postupně měnit, je někdy možno získat z korelací tak jasné poznatky, že přece jen lze mluvit o důkazu účinku. Když např. výnos sklizně daného roku a stejného osiva na půdách stejného druhu vykazuje vzestup odpovídající použitému množství hnojiva, může být za statisticky zjištěné pokládáno tvrzení, že větší použití hnojiva přineslo větší výnosy.

Právě na tomto příkladu je však možno hned také ukázat, že se korelace často *nedají extrapolovat donekonečna* a že *nemusí vždy probíhat „lineárně"*. Budu-li neustále zvyšovat dávky hnojiva, nebude se docilovaný výnos ve stále větší míře vyplácet (problematika mezní hodnoty v ekonomické teorii) a jednou dokonce dosáhnu bodu, v němž se výnos začne snižovat, protože dobrého bylo příliš mnoho. Od tohoto bodu se *korelace* stane dokonce *negativní:* čím více hnojiva, tím menší výnosy.

K provedení takového pokusu se pokusné pole se stejně kvalitní půdou rozdělí na jednotlivé ověřovací úseky a pak se na jednotlivé úseky použije různé množství hnojiva.

O tomto uspořádání „latinských čtverců", které se často vyskytuje při plánovaných pokusech, ještě uslyšíme, až se budeme zabývat analýzou rozptylu (viz str. 334), která této metody také často používá. To samozřejmě není s oblibou používá. V obou případech mají pouhá náhoda. V obou případech mají být objasněny závislosti mezi dvěma nebo více proměnnými znaky, variacemi v uspořádání pokusu. Analýza rozptylu přitom převyšuje korelační počet tím, že se snaží rozlišovat mezi náhodnými výběrovými chybami a významnými odchylkami, zatímco korelační počet přejímá zkoumání dvojice měřených hodnot jako dané a zatím se nezajímá o to, do jaké míry mají být tyto údaje pokládány jen za náhodný výsledek určitého rozdělení.

Tím jsme se již dostali k ústřední otázce tohoto oddílu: *co to jsou korelace, jaký mají význam a jak se počítají.*

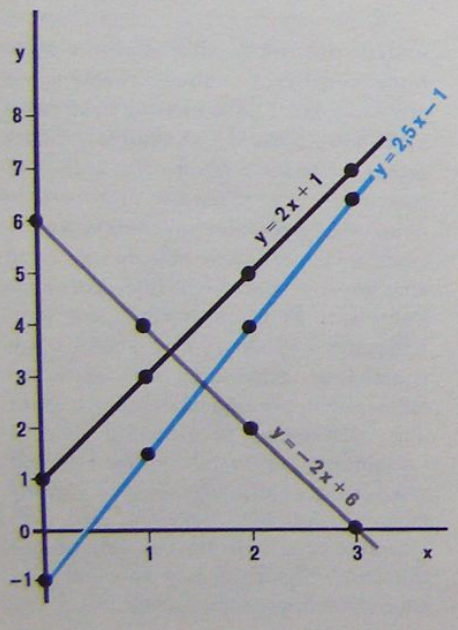


9.2 Korelace

Co je matematická funkce, si mnozí ze školy pamatují už jen tak přibližně — tedy jedna hodnota zde závisí na jiné, je jí vyjádřena. Nejjednodušší způsob je např.: $y = 2x$, o něco obtížnější je $y = 2x + 2$, sotva obtížnější je $y = x^2 - 1$. Přesto se rychle podívejme, které hodnoty závislá proměnná y přijímá, jestliže přidělíme nezávisle proměnné označené x různé hodnoty:

	$y = 2x$	$y = 2x + 2$	$y = x^2 - 1$
$x = 0$	$y = 0$	$y = 2$	$y = -1$
$x = 1$	$y = 2$	$y = 4$	$y = 0$
$x = 2$	$y = 4$	$y = 6$	$y = 3$
$x = 3$	$y = 6$	$y = 8$	$y = 8$

Zaneseme-li tyto hodnoty do souřadnicové sítě, dostaneme dvě souběžné přímky a jednu křivku (viz vedlejší obr. nahoře). Křivkou se zatím nebudeme zabývat, protože nám jde jen o lineární regresi. Podívejme se proto ještě jednou na obě funkční rovnice: $y = 2x$, $y = 2x + 2$. Je nápadné, že probíhají úplně souběžně; závisí to snad právě na těch „2 x"? Zkusíme to s $y = 2x - 1$



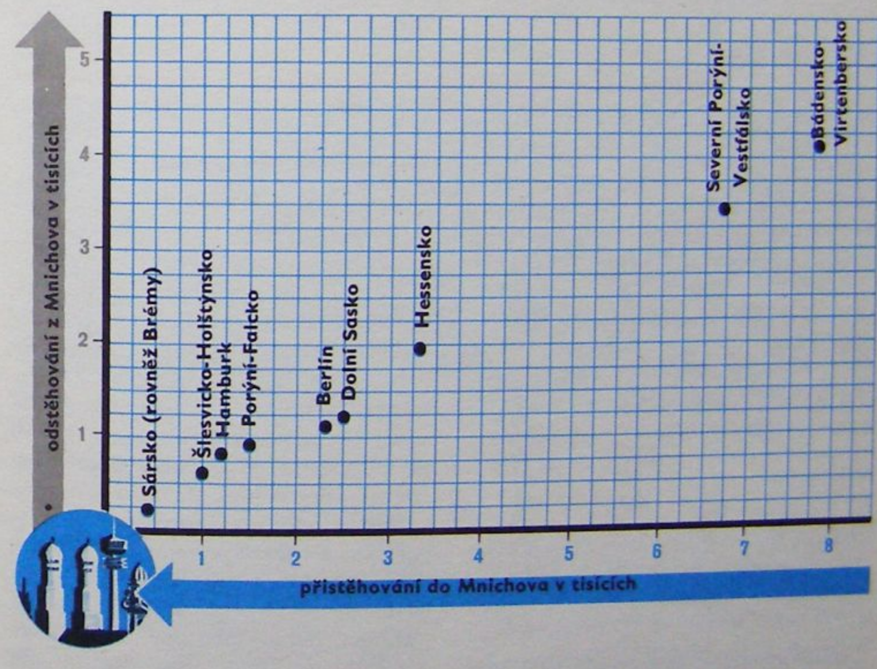
Nahoře:
Rovnice $y = 2x - 1$, $y = 2x$ a $y = 2x + 2$ mají stejné směrnice přímky; rozdílná konstanta na pravé straně rovnice udává pouze „úsek osy". Světle modrá křivka vyjadřuje kvadratickou rovnici.

Dole:
Větší strmost stoupání přibližuje $y = 2{,}5x - 1$ velmi rychle přímce $y = 2x + 1$, která začíná výš. U $x = 4$ by se přímky protínaly. Záporné znaménko $y = -2x + 6$ znamená „negativní směrnici", „záporné stoupání".

— a zase dostáváme rovnoběžku, tentokrát body se souřadnicemi (0, —1), (1,1), (2,3) a (3,5). Vzestup přímky závisí tedy jen na čísle před x, konstanta vzadu odpovídá pouze na otázku, kde přímka přetíná osu y. Tak má $y = 2{,}5x - 1$ strmější směrnici než $y = 2x + 1$, ačkoliv tato přímka začíná „výše" (viz obr. na str. 286 dole).

Směr ostatně nemusí vždy být vzestupem, může být také *poklesem a negativní korelací*. Je to např. případ funkce $y = -2x + 6$. Zde dostáváme jako první body v souřadnicovém systému (0, 6), (1, 4), (2, 2), (3, 0). Obecně zní lineární regresní rovnice takto: $y = ax + b$, přičemž *a* znamená *směrnici*, *b úsek osy*. V matematice jsou tyto strnulé závislosti podobného typu dosti časté, v běžné statistice, která je odrazem všedních dnů, se prakticky nevyskytují. Někdy se jim ovšem značně přibližují.

Podívejme se např. na stěhování z německých spolkových zemí (mimo Bavorsko) do Mnichova, resp. z Mnichova. V mnichovské statistice za rok 1969 jsou tyto hodnoty (stěhování obyvatelstva v tisících):



Téměř dokonalou korelaci ukazují údaje „vystěhování z a přistěhování do", v našem případě Mnichova. Ze všech spolkových zemí NSR se sice více lidí do Mnichova přistěhovává než obráceně, ale číselný poměr obojích je téměř stejný.

13 evropských států, vnitrostátní údaje

Velmi zřetelnou zápornou korelaci vykazují odeslané dopisy a telegramy (vždy vnitrostátní údaje) ve 13 evropských státech: kde se píše mnoho dopisů, telegrafuje se poměrně málo, a naopak.

Země	Odstěhování do Mnichova	Přistěhování z Mnichova
Bádensko-Württembersko	7,8	4,1
Severní Porýní-Vestfálsko	6,7	3,4
Hessensko	3,3	1,9
Dolní Sasko	2,5	1,2
Berlín	2,3	1,1
Porýní-Falcko	1,5	0,9
Hamburk	1,2	0,8
Šlesvicko-Holštýnsko	1,0	0,6
Sársko	0,4	0,2
Brémy	0,4	0,2

Jestliže zaneseme tyto hodnoty do našeho souřadnicového systému, dostaneme řadu bodů, které nás téměř nutí k tomu, abychom jimi protáhli zcela přímou linii. I ve statistice se tedy vyskytují skoro *dokonalé korelace.* (Viz obr. na str. 287.)

Ostatně i *záporné.* V souřadnicovém systému jsme sestavili údaje ze 13 evropských zemí o počtu odeslaných telegramů a dopisů. Na ose x je uveden počet odeslaných vnitrostátních telegramů na 1000 obyvatel, na ose y počet dopisů na 1 obyvatele. Čísla jsou z roku 1965. V tomto případě není sice vidět přímku, avšak zápornou korelaci lze zřetelně rozeznat. Čím více se píše dopisů, tím méně se posílá telegramů. Je možno uvést i psychologické důvody: kde se málo píše, hlásí se lidé, jen když je to životně důležité — a potom se posílá telegram. Nebo lze usuzovat i takto: kde není možno spolehnout se na doručení dopisu a na spolehlivou práci pošty, píše se málo a skutečně důležité věci se telegrafují, aby zpráva došla včas. (Viz obr. na str. 288.)

Jestliže je k dispozici velký počet dvojic měřených hodnot, doporučuje se seskupit výsledky do skupin (intervalů) obdobným způsobem, jaký byl použit v oddíle 2.6. K objasnění tohoto postupu uvádíme příklad z biometrie. Opět jde o korelaci mezi tělesnou výškou a váhou a čísla uvádíme z díla E. Webera „*Grundriss der Biologischen Statistik*" (Rukověť biologické statistiky).

Zde se uvádí, že při měření a vážení 648 mladých mužů byly zjištěny tyto výsledky, které byly podle četnosti uspořádány do skupin:

značily např. normální rozdělení tělesných výšek v příslušné váhové skupině. Již z číselné tabulky lze poznat zleva dole doprava nahoru stoupající charakteristickou přímku kladné korelace, zvláště tehdy, když se zvýrazní zvlášť často zastoupené dvojice měřených hodnot. Ještě výraznější však tato korelace bude, jestliže proti jednotlivým velikostním skupinám postavíme průměrnou váhu. Potom např. muži vysocí 157,5 cm váží v průměru:

$$\bar{y}_{\text{skup. }157,5} =$$

$$= \frac{3 \,.\, 64 + 10 \,.\, 59 + 9 \,.\, 54}{3 + 10 + 9} =$$

$$= \frac{1268}{22} = 57{,}6 \text{ kg.}$$

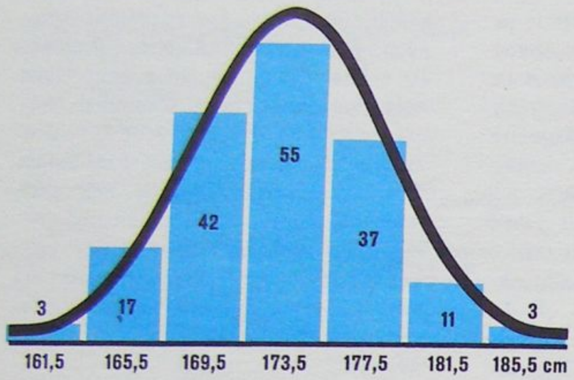
Zaneseme-li tyto hodnoty do našeho

Váha v kg	157½	161½	165½	169½	173½	177½	181½	185½
89							1	
84					3	1	2	2
79				1	8	5	3	2
74		2	4	10	21	18	12	1
69		3	17	42	55	37	11	3
64	3	12	51	72	44	9	3	
59	10	31	51	40	7	1		
54	9	23	13	4	1			

Výška v cm

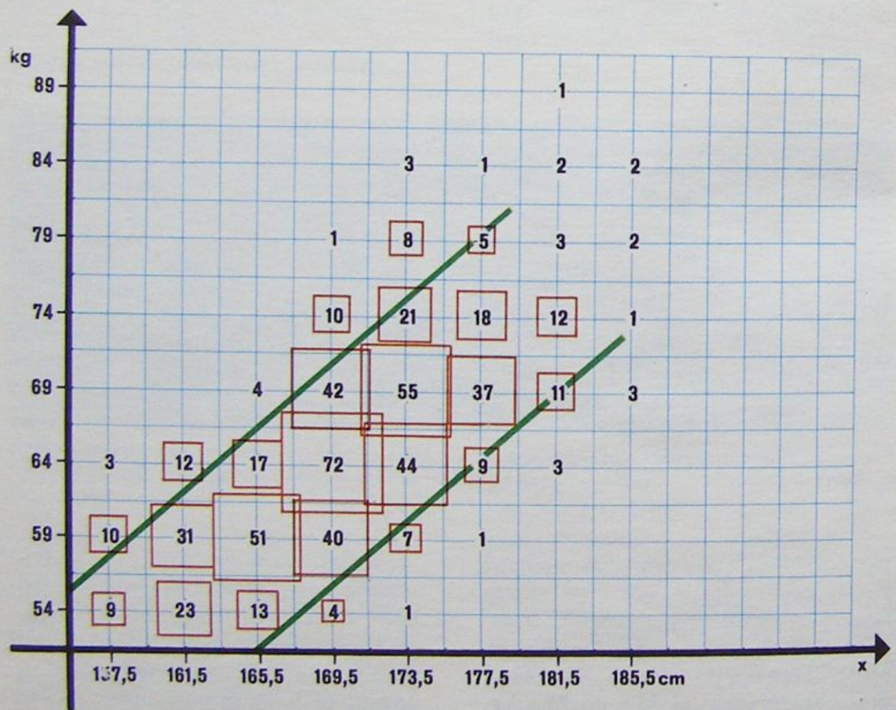
Přehled o absolutní četnosti měřených hodnot z našeho příkladu o korelaci mezi tělesnou vahou a tělesnou výškou (čísla vždy udávají střed intervalu; šířka intervalu činí 4 kg, resp. 5 cm).

Hodnoty znaků vypadají tak podivně jen proto, že jde o příslušné středy intervalů: 169,5 cm znamená „mezi 167,5 a 171,5 cm". Rovněž tělesná váha 69 kg udává střed intervalu „širokého" 5 kg, která sahá od 66,5 do 71,5 kg. Četnosti u silně zastoupených intervalů jsou ostatně již dost velké, aby jasně vy- souřadnicového systému, bude tato přímka ještě zřetelnější „*regresní přímkou*". Omezením původních 648 dvojic měřených hodnot na několik kombinací skupin s průměrnými hodnotami se podstatně zvýší přehlednost. Početní úloha se ovšem nezjednodušila: chceme-li totiž zjistit přesnou číselnou kore-



řádové rozdělení „váhové třídy 69 kg"

Tělesná výška v jednotlivých váhových intervalech (v tomto případě středu intervalu 69 kg) ukazuje již velmi zřetelně přiblížení k normálnímu rozdělení.



Jestliže se vyjádří četnosti shody jednotlivých váhových a výškových intervalů graficky (na obrázku pomocí čtvercových ploch), je vidět velmi pozitivní korelaci zvlášť zřetelně.

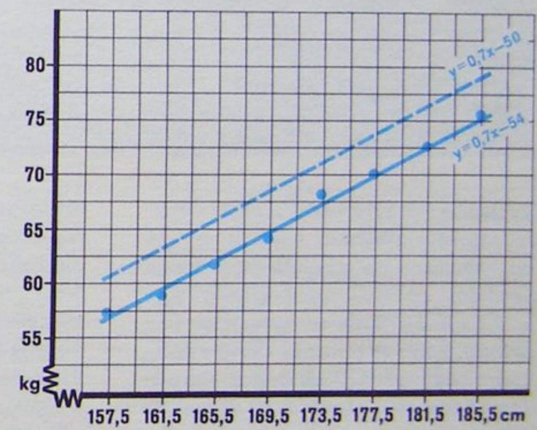
laci, nemohly by se již nynější dvojice měřených hodnot pokládat za stejně hodnotné jako jednotlivá měření (jako v případě 648 původních měření). Mnohé z těchto bodů zastupují totiž nyní několik desítek (v jednom případě dokonce 72) měření.

Před dalším početním „zpracováním" musely by proto být tyto dvojice průměrných hodnot vhodně „váženy": ojedinělí, zvlášť lehcí nebo těžcí, zvlášť velcí nebo malí muži by padali mnohem méně „na váhu" než ty četné stovky „ve středu pole". Do tohoto počítání se pouštět nebudeme nikoliv proto, že by bylo zvlášť obtížné, ale spíše proto, že vyžaduje mnoho času. Především však proto, že jsme se ještě neseznámili ani se základním principem této početní metody.

Jestliže si blíže všimneme grafu, můžeme dostat přijatelnou přibližnou hodnotu dokonce bez výpočtů. Jak tomu bude například, když $y = 0{,}7\,x - 50$? Činí tedy váha přibližně $^7/_{10}$ tělesné výšky minus 50? Tato rovnice se velmi dobře (ještě lépe ve tvaru $y = 0{,}7\,x - 54$) přizpůsobí bodům v souřadnicovém systému, ale je nutno vyvarovat se řešení rovnice podle x a usuzovat: když $y = 0{,}7\,x - 54$, pak $0{,}7\,x = y + 54$ a $x = 1{,}43\,y + 77$.

V tomto případě nejde o matematickou funkci, naše rovnice vyjadřuje pouze přibližnou závislost, a to především jen závislost tělesné váhy (y) na tělesné

Zběžně odhadnutá přímka $y = 0{,}7x - 50$ zobrazuje poměr váhy k tělesné výšce již docela dobře. Přesněji vypočítaná je regresní přímka $y = 0{,}7x - 54$. Z ní lze vyčíst odpověď na otázku: jak asi těžký bude muž vysoký x cm? Není to však nejlepší rovnice pro zjištění výšky při dané váze.



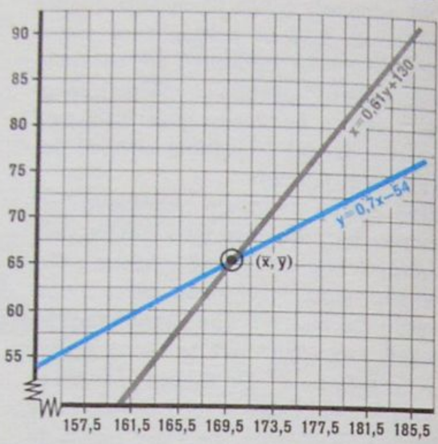
tělesná váha (zatím neznámá)

57,6	59,1	62,1	64,5	68,5	70,4	72,9	75,9
157,5	161,5	165,5	169,5	173,5	177,5	181,5	185,5

výška (střed intervalu) (známá)

Tato tabulka dává odpověď na otázku „jak asi těžký bude muž určité výškové třídy". Tělesná výška (x) je dána, hledá se váha (y). — Je-li naopak známa váha, výška však nikoliv, musí se vypočítat tabulka uvedená na str. 292.

výšce (x). Rovnice $y = 0{,}7\,x - 54$ nám poskytuje pouze odhad hodnoty, kterou potřebujeme pro odpověď na otázku: „Kolik asi bude vážit muž, který je x cm vysoký?“
K tomuto odhadu je třeba přistupovat opatrně; v příští kapitole se dovíme víc o tom, jak se tato regresní křivka vypočítává a jaké může mít chyby. Bez ohledu na to nelze x a y beztrestně zaměnit (mimo případ funkční závislosti) a pak očekávat od vzorce $y = 0{,}7\,x - 54$ odpověď na otázku: „Jak vysoký bude muž, který váží y kg?“ Výpočet se provede snadno, ale je nesprávný.
To uvidíme zcela jasně, jestliže ke každému středu skupiny tělesné váhy zjistíme průměrnou výšku:



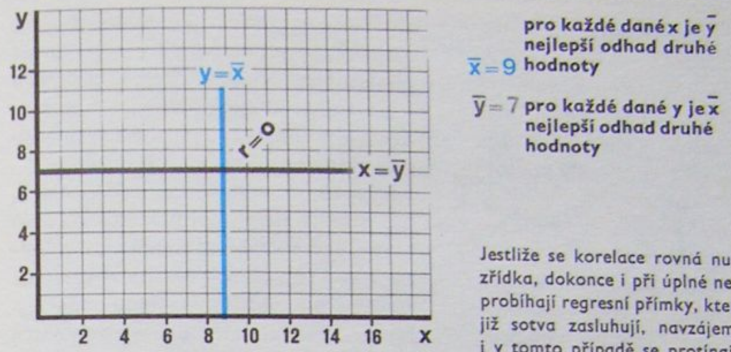
Regresní přímky se navzájem protínají právě v bodě $(\bar{x}, \bar{y})$.

Tělesná váha (střed skupiny) (známá)							
54	59	64	69	74	79	84	89
162,7	165,7	169,2	173,1	174,7	176,9	179,0	181,5
Průměrná výška (zatím neznámá)							

Vyhledáme-li tyto body v našem souřadnicovém systému, zjistíme, že až na několik málo výjimek jsou hodně vzdáleny od naší staré přímky $y = 0{,}7\,x - 54$. Spíše se formují do jiné přímky, která se protíná s první v důležitém bodě: při 170 cm a 65,2 kg, přičemž dodatečný propočet by ukázal, že tyto hodnoty jsou celkovými průměry všech tělesných výšek, případně tělesných vah. Krátce a statisticky formulováno:

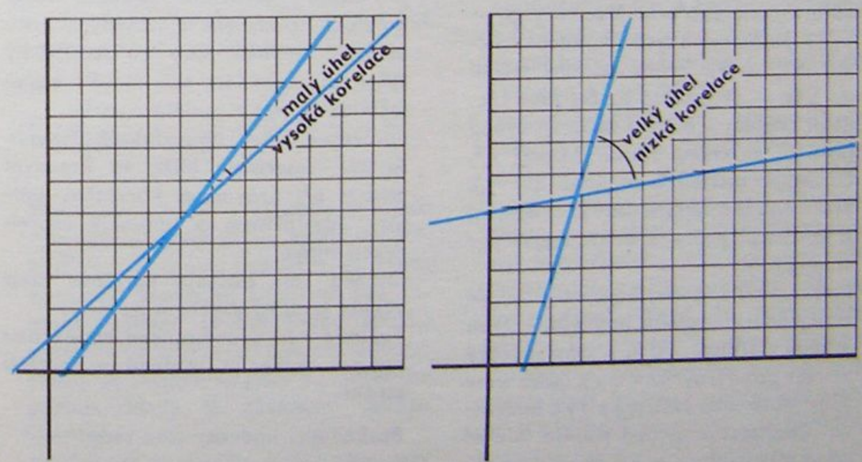
$\bar{x} = 170$ cm, $\bar{y} = 65{,}2$ kg. *Regresní přímky se navzájem protínají vždy v bodě* $(\bar{x}, \bar{y})$.

Než budeme pokračovat, ještě malá hra myšlenek. Co se stane s oběma regresními přímkami, není-li mezi oběma řadami znaků vůbec žádná souvislost, takže jsou v systému souřadnic rozděleny jako kulatý mrak bodů?
Tedy je-li y zcela nezávislé na x, mění se y naprosto „svébytně“, orientace na situaci podél osy x mizí úplně, regresní přímka se stane paralelou s hodnotou $y = \bar{y}$. Na otázku: „Jaké y asi dostaneme, máme-li tu nebo onu hodnotu x?“, je stále stejná odpověď: „Zkuste to s průměrnou hodnotou y, s $\bar{y}$.“ — Stejně tak se regresní přímka pro x stane paralelou osy y s konstantní hodnotou $x = \bar{x}$. Obě přímky se protínají v bodě $(\bar{x}, \bar{y})$ jako všechny regresní přímky,



Jestliže se korelace rovná nule (což se stává zřídka, dokonce i při úplné nezávislosti x a y), probíhají regresní přímky, které toto označení již sotva zasluhují, navzájem kolmo, avšak i v tomto případě se protínají v bodě $(\bar{x}, \bar{y})$.

svírají však spolu pravý úhel; jedná se o nulovou korelaci.
Druhý extrém již známe. Je-li y funkcí x a každé hodnotě x odpovídá přesně vypočitatelná hodnota y, regresní přímky splývají a tvoří grafický výraz funkce — např. $y = 2\,x$.
Ve statistické praxi se oba extrémy téměř nevyskytují, regresní přímky se tedy ani navzájem neslučují, ani nestojí kolmo na sobě, nýbrž svírají spolu více nebo méně ostrý úhel. Tento úhel bude tím menší, čím těsnější je korelace mezi x a y. *Velikost úhlu vyjadřuje tedy míru „závislosti“* — závislosti v uvozovkách, protože se nechceme vázat určením, která proměnná je opravdu závislá na druhé.



Míru korelace lze vyčíst z úhlu, který tvoří obě regresní přímky: čím ostřejší úhel, tím vyšší korelace.

9.21 Korelační koeficient

Závislost lze vyčíst nejen z velikosti úhlu, který svírají obě regresní přímky v systému souřadnic, ale je také možno *výpočtem zjistit* (což je obvyklý způsob) tzv. „korelační koeficient“.
Jak lze takový korelační koeficient vytvořit? Má vyjadřovat, jak je navzájem závislé chování obou proměnných x a y; toto chování je však charakterizováno rozptylem a střední hodnotou — v případě normálního rozdělení tedy pomocí $\bar{x}$ a s_x, případně y a s_y. (Nejde-li o normální rozdělení, je možno použít mimo jiné testů pořadí nezávislých na rozdělení.) Většinou však jde o normální rozdělení a věda považuje společné rozdělení obou proměnných za *„dvourozměrné normální rozdělení“*.
Mimo tyto charakteristické hodnoty jednotlivých proměnných musí být také nějak vyjádřeno jejich spojení. To se stane pomocí tzv. *kovariance x a y*. Při výpočtu rozptylů jsme vždy počítali rozdíly jednotlivých měrných hodnot od aritmetického průměru, tedy $(x_i - \bar{x})$; potom jsme tyto diference umocnili na $(x_i - \bar{x})^2$, druhé mocniny sečetli na $\Sigma (x_i - \bar{x})^2$ a tento součet zase rozdělili počtem n měření nebo správněji $(n - 1)$, abychom dostali rozptyl s_x^2. U dvojic měřených hodnot (x_i, y_i), které jsou charakteristické pro regresní a korelační počty, probíhá stejný postup i pro y.
Indexem i u x_i a y_i vyjadřujeme, že jde o nějakou libovolnou hodnotu, kterou x nebo y přijímá. To je přesný vědecký způsob psaní (rovněž x_i a y_i), často jsme však místo toho psali jen x a y a budeme to příležitostně dělat i nadále. Ale ať už se píše např. $(x - \bar{x})^2$ nebo $(x_i - \bar{x})^2$, je v obou případech míněno totéž.
Abychom dostali kovarianci znaků x a y, vezmeme u každé jednotlivé dvojice měřených hodnot (x_i, y_i) rozdíly proti příslušnému aritmetickému průměru a navzájem je vynásobíme; tak dostaneme pro každou dvojici měřených hodnot výraz $(x_i - x)(y_i - y)$.
Součet těchto výsledků se nakonec dělí, jak jsme se shodli při výkladu Studentova t, počtem měření zmenšeným o jednu $(n - 1)$ a pak dostaneme kovarianci znaků x a y, psanou většinou jako s_{xy}.
Nyní dosadíme naše hodnoty do vzorce pro korelační koeficient:

$$\boxed{r = \frac{s_{xy}}{s_x \cdot s_y}}$$

Jde-li o základní soubory, mění se s na σ a r na ϱ (ró). Zpravidla však půjde o výběry, korelační koeficient r je proto — jako vše, co pochází z výběrů — postižen nejistotou, je jen přibližně správný: r je pouze *„funkce odhadu“* pro ϱ a tím méně spolehlivý, čím menší je rozsah výběrového souboru. Nízký korelační koeficient z velkého vzorku ukazuje mnohem více na statisticky zjištěnou souvislost než vysoký korelační koeficient z malého vzorku.
Podívejme se nyní na podkladě příkladu André Vessereaua blíže na pracovní postup při provádění korelační analýzy. Jde přitom o náhodný vzorek deseti znaků.
Otázka zní: existuje korelace mezi věkem ženicha a věkem nevěsty?
Tabulka s již vypočítanými hodnotami pro náš vzorec je uvedena na protější straně.

Rozhodující hodnoty jsou tedy:

$$n = 10; \quad \Sigma(x - \bar{x})(y - \bar{y}) = 179;$$
$$s_x = \sqrt{30}; \quad s_y = \sqrt{18{,}9}; \quad s_{xy} = 19{,}9.$$

	Věk ženicha $\bar{x}$	Věk nevěsty $\bar{y}$	$x - \bar{x}$	$(x - \bar{x})^2$	$y - \bar{y}$	$(y - \bar{y})^2$	$(x - \bar{x})$ krát $(y - \bar{y})$
	22	19	— 7	49	— 6	36	+ 42
	25	22	— 4	16	— 3	9	+ 12
	26	21	— 3	9	— 4	16	+ 12
	26	23	— 3	9	— 2	4	+ 6
	27	23	— 2	4	— 2	4	+ 4
	28	24	— 1	1	— 1	1	+ 1
	30	29	+ 1	1	+ 4	16	+ 4
	30	27	+ 1	1	+ 2	4	+ 2
	35	33	× 6	36	+ 8	64	+ 48
	41	29	× 12	144	+ 4	16	+ 48
Celkem	290	250	—	270	—	170	+ 179
Průměr	$\bar{x} = 29$	$\bar{y} = 25$	—	—	—	—	—

Zobrazení výpočetního postupu příkladu uvedeného v textu o korelaci mezi věkem ženicha a nevěsty.

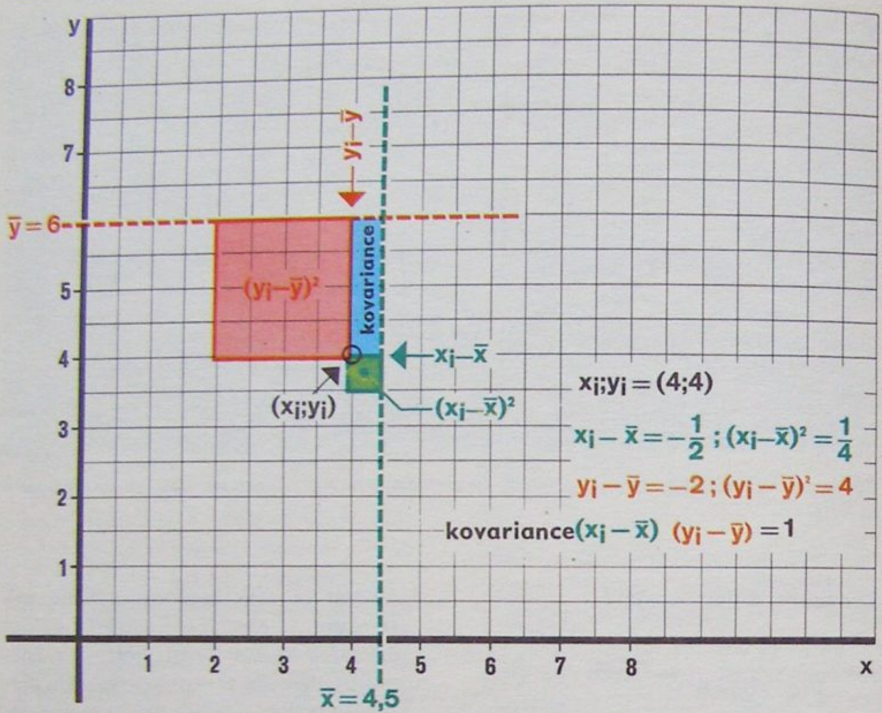
Dosadíme a dostaneme:

$$r = \frac{19{,}9}{\sqrt{30}\ \sqrt{18{,}9}} = 0{,}835.$$

Vessereau ostatně počítá zjednodušeně přímo s tabulkovými součty a dostává stejný výsledek tímto způsobem:

$$r = \frac{179}{\sqrt{270}\ \sqrt{170}} = 0{,}835.$$

Korelace je vždy — alespoň v tomto výběrovém souboru — velmi vysoká. Než však začneme uvažovat, co tento korelační koeficient znamená pro základní soubor sňatků (nebo sňatků uzavřených v Paříži během jednoho roku), podívejme se ještě pozorněji na postup, který je základem tohoto výpočtu.
Zjišťují se: především *rozptyl obou proměnných* — jaký je rozptyl x kolem $\bar{x}$, jaký je rozptyl y kolem $\bar{y}$? Při každém měření se však současně zjišťuje, jak se *navzájem chovají obě odchylky dvojice měřených hodnot* (proti vlastnímu průměru). Protože při tomto násobení platí znaménka, je zcela dobře možné, že součet těchto výsledků $(x - \bar{x})(y - \bar{y})$ bude velmi nízký, a protože je uveden v čitateli vzorec pro r, povede k tomu, že též korelační koeficient bude nízký. (Viz obr. na str. 296.)
Vzhledem k základnímu významu korelačního počtu na dalších dvou příkladech ukážeme, jak se tato výpočetní metoda projevuje. Již jsme předvedli vysokou korelaci a nyní představme nepatrný vzorek jiného složení. K tomu je třeba důrazně připomenout, že vzorek s $n = 5$, který pro jednoduchost volíme, je při použití této výpočetní metody zásadně nepřípustný, protože vypovídací hodnota takto zjištěného korelačního koeficientu se prakticky rovná nule.
Přesto se však podívejme na takový vzorek s dále uvedenými hodnotami:



Kovariance měří obě odchylky bodu (x_i, y_i) proti oběma středním hodnotám $(\bar{x}, \bar{y})$. Zatímco umocněné odchylky $(x_i - \bar{x})^2$ a $(x_i - \bar{y})^2$ jsou skutečně vždy kvadratické (čtvercové), jeví se kovariance jako obdélník. Jen když $(x_i - \bar{x}) = (y_i - \bar{y})$, je také kovariance kvadratická (čtvercová).

přitom opět volíme zjednodušený výpočetní postup:

Ženich	Nevěsta	$x - \bar{x}$	$(x - \bar{x})^2$	$y - \bar{y}$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
22	24	− 7	49	− 1	1	7
27	25	− 2	4	0	0	0
29	26	0	0	+ 1	1	0
31	29	+ 2	4	+ 4	16	8
36	21	+ 7	49	− 4	16	− 28
$\bar{x} = 29$, $\bar{y} = 25$			106		34	− 13

$$r = \frac{-13}{\sqrt{106}\ \sqrt{34}} = \frac{-13}{(10{,}3)\ (5{,}8)} = -0{,}22$$

To je nepatrný, bezvýznamný záporný korelační koeficient, který však celkem dobře vyjadřuje situaci: jeden muž se oženil se starší ženou, jeden s podstatně mladší, tři zbývající manželství jsou blízko středních hodnot; že je ženich přitom vždy starší než nevěsta, vyplývá z vyššího průměrného věku všech ženichů. Nechme nyní tytéž osoby uzavřít sňatek s jinými partnery:

Ženich	Nevěsta	$x - \bar{x}$	$(x - \bar{x})^2$	$y - \bar{y}$	$(y - \bar{y})^2$	$(x - \bar{x})(y - \bar{y})$
22	29	− 7	49	+ 4	16	− 28
27	26	− 2	4	+ 1	1	− 2
29	25	0	0	0	0	0
31	24	+ 2	4	− 1	1	− 2
36	21	+ 7	49	− 4	16	− 28
$\bar{x} = 29$, $\bar{y} = 25$			106		34	− 60

$$r = \frac{-60}{\sqrt{106}\ \sqrt{34}} = \frac{-60}{(10{,}3)\ (5{,}8)} = -1{,}0$$

Tím jsme vypočítali maximum záporné korelace — správně, protože čím mladší je ženich, tím starší byla (poměrně k průměru) nevěsta. Umocněné odchylky mezi muži (x), případně ženami (y), zůstaly stejné, protože jsme počítali se stejnými osobami. Změna korelačních koeficientů je proto vyvolána výlučně vysokou zápornou kovariancí s_{xy}!

Pearsonův korelační koeficient ϱ, případně r, nelze ostatně použít, jestliže regresní přímky nejsou přímkami (a to znamená, že proměnlivé znaky nemají normální rozdělení). V takovém případě je nutno použít buď přeformovaný koeficient, nebo jeden z ostatních korelačních koeficientů, jako např. *Spearmanův koeficient korelace pořadí*, který si ještě v oddílu o testech pořadí stručně ukážeme. (Viz str. 337.)

Mimo jiné je třeba také uvážit, že i nejlépe vypadající, na několik desetinných míst přesně vypočítaný korelační koeficient vzorku je jen odhadem z hlediska celku, a to většinou dosti nespolehlivým. Pomocí srovnávací tabulky lze určit jeho interval spolehlivosti a většinou se ukáže, že jeho vypovídací schopnost je značně omezena. Příklad je možno uvést z Pfanzaglovy „*Všeobecné nauky o metodách statistiky*“:

„...Z měření $n = 20$ byl vypočítán korelační koeficient $r = 0{,}73$. S 99% pravděpodobností platí: ϱ je větší než 0,29, ale menší než 0,92.“ Méně závazně to snad již nelze říci: nějaká pozitivní korelace tu jistě je, to je však též vše, co se nakonec dovíme.

Jestliže se v daném případě při dvaceti dvojicích měřených hodnot docílilo tak neurčité odpovědi o korelaci v základním souboru, je nejistota při ještě menších výběrech mnohem větší. Např. z našeho vzorku o rozsahu $n = 10$ u sňatků jsme dostali působivý korelační koeficient $r = 0{,}835$, který svědčí o velmi těsné korelaci. Při kontrole pomocí vhodných způsobů zjišťování intervalů spolehlivosti tohoto r by se však ukázalo, že je určitá, i když nepatrná možnost, že v základním souboru není vůbec žádná korelace! S pravděpodobností 95 % platí jen to, že ϱ je mezi 0,52 a 0,97.

9.3 Regresní počet

Regrese — kouzelné slovo. Znají a používají je psychologové, objevuje se v biologii a v právu (které zná regresní pohledávky) a vždy znamená „sáhnout dozadu", „ustoupit zpět", „vrátit se nazpět". Podíváme-li se do anglického slovníku, je hned na prvním místě doslovný význam: „*vstup nazpět*", „*krok nazpět*". Co toto slovo znamená ve statistice a proč jsme kvůli němu nahlíželi do anglického slovníku?

Byl to totiž anglický biolog a statistik sir Francis Galton, který zavedl „regresi" do statistiky a o němž jsme již slyšeli při výkladu historie normálního rozdělení. Tímto pojmem z biologie označoval korelaci, pro kterou našel zajímavý biometrický příklad: děti vykazují svou tělesnou výškou zřetelnou korelaci s tělesnou výškou otce; jsou-li však otcové velmi vysocí nebo velmi malí, objevuje se znatelná „regrese" k průměrné tělesné výšce celého obyvatelstva: „*ustoupily zpět*" k průměru, od kterého se jejich otcové příliš vzdálili.

Dnes je regrese ve statistice od těchto biometrických vazeb úplně osvobozena a regresní analýza přechází téměř nepozorovaně do korelační analýzy. Když jsme mluvili o korelacích, zmínili jsme se také o regresních přímkách a nyní, když mluvíme hlavně o regresi, zabýváme se opět zásadně korelacemi.

Vyžaduje-li se přesné rozlišení, je nutno říci, že *korelační počet pokládá obě proměnné za spojené, ale nezávislé,* zatímco regresní analýza se zabývá především otázkou, jak je možno na základě zjištěné souvislosti mezi proměnnými předpovídat chování *závislé* proměnné podle chování *nezávislé* proměnné. Francis Galton formuloval problém takto: výška dětí není zřejmě nezávislým proměnlivým znakem, nýbrž znakem závislým na výšce otců. Regresní analýza by tedy měla po zjištění souvislostí obecně předpovídat, jaká asi bude výška dětí.

Korelační analýza neřeší úplně otázku, která z proměnných ovlivňuje druhou (a zda ji vůbec ovlivňuje): určuje věk nevěsty ženichův věk, nebo naopak? Ovlivňuje váha člověka jeho výšku, nebo výška jeho váhu? Má počet odeslaných dopisů vliv na počet odeslaných telegramů, nebo naopak? Nevíme, a pro měření závislosti je to i bezvýznamné.

Regresní analýza používá v podstatě stejné výpočetní metody jako korelační analýza, zdůrazňuje však závislost: samozřejmě že výška dětí závisí na rodičích, a ne naopak. — Samozřejmě že krevní tlak závisí na věku, a nikoliv věk na krevním tlaku. — Samozřejmě že přírůstek váhy krmného vepře závisí na množství krmiva, a nikoli množství krmiva na přírůstku váhy.

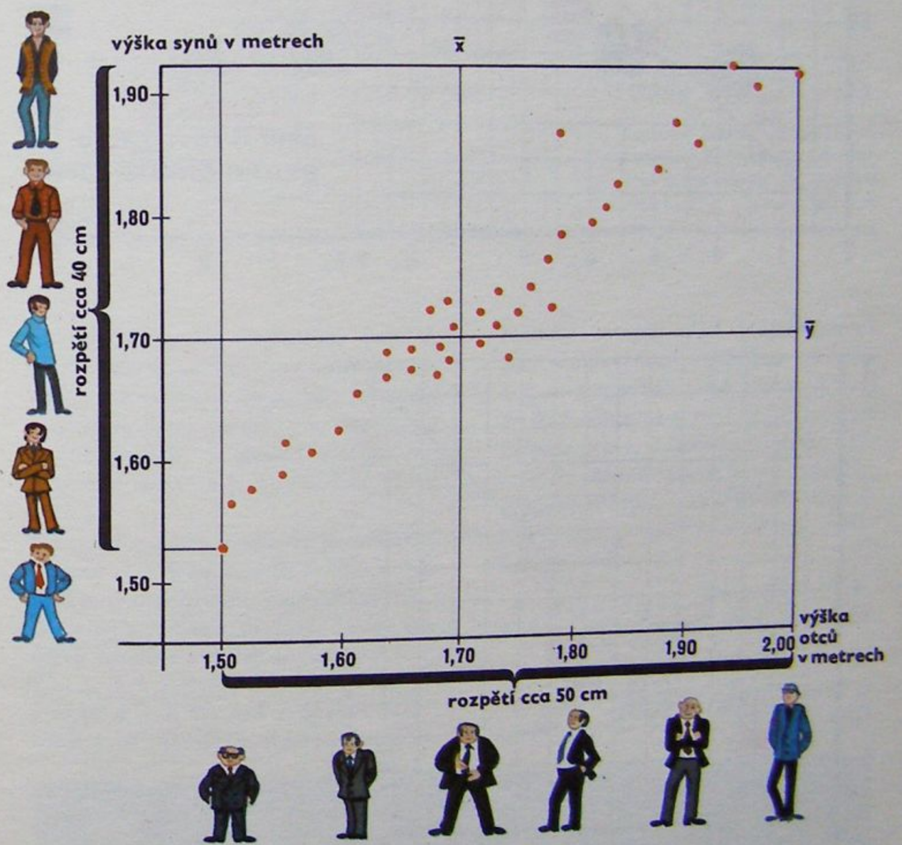
Jak jsme zjistili už při našich obecných úvahách o korelacích, jde o to nalézt matematický výraz pro závislost mezi dvěma proměnnými. Více nebo méně rozptýleným mrakem bodů dvojic měřených údajů máme vést přímku, která vyjádří situaci pokud možno nezkresleně. Takové regresní rovnice jsme si již ukázali, aniž jsme se zabývali způsobem jejich výpočtu. Nyní uvedeme alespoň základní principy.

Co můžeme očekávat od přímky, která má probíhat tak, aby alespoň poněkud věrně zobrazovala polohu více bodů v systému souřadnic a „trend" těchto dvojic měřených hodnot? Taková přímka zřejmě musí probíhat tak, aby co nejvíce přibližovala velkému počtu bodů. Jinými slovy: rozptyl bodů po

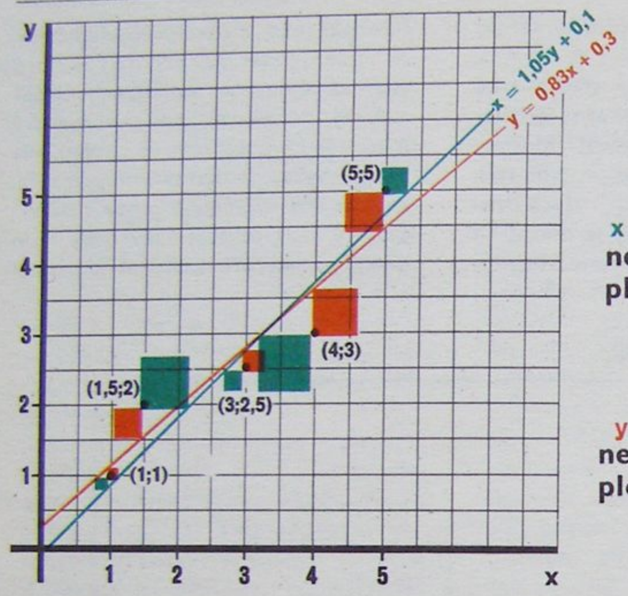
obou stranách regresní přímky se má omezit na nejmenší míru.

Jak víme již z výpočtu směrodatné odchylky, pro rozptyl je charakteristická druhá mocnina vzdálenosti. Musíme tedy hledat přímku, která probíhá tak, že součet všech rozptylů (tj. všech druhých mocnin vzdáleností) se omezí na minimum. Proto se tato metoda označuje také jako „*metoda nejmenších čtverců*". (Obr. na str. 300 nahoře.) Je stará více než půldruhého století, neboť francouzský matematik Adrien Legendre ji zavedl již počátkem 19. století, ale přesto vyžaduje ke svému uplatnění znalost diferenciálního počtu.

Existují však také metody zjištění regresní přímky, při nichž není nutné znát

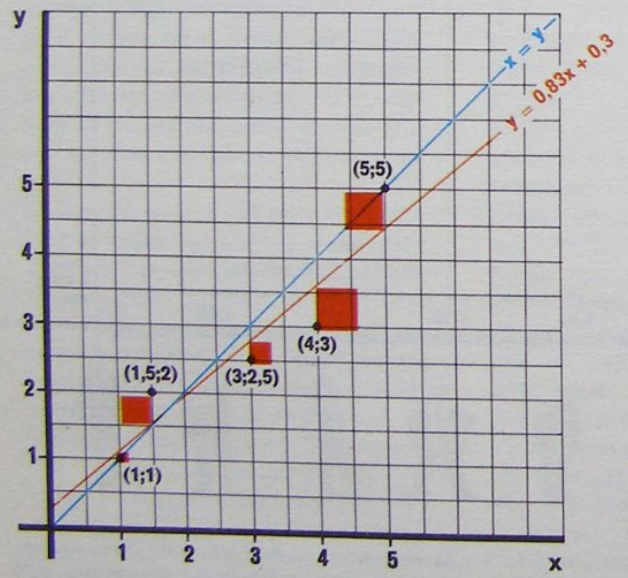


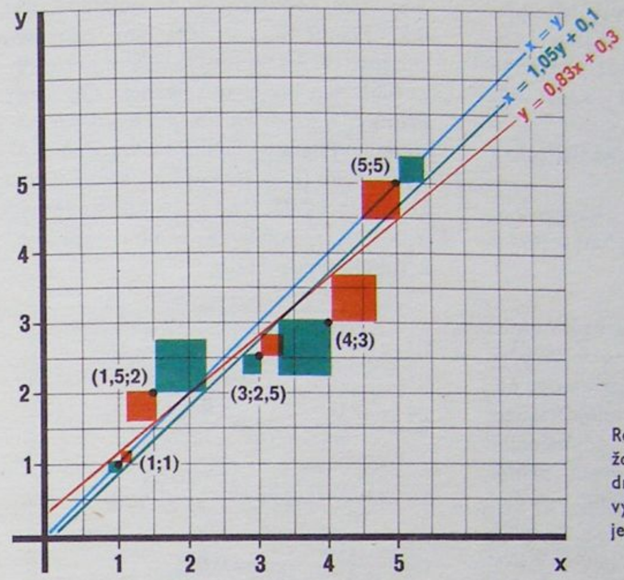
„Regrese" znamená „krok zpět", tedy krok zpět směrem k průměru. Jak např. ukázal Galton, jsou synové extrémně vysokých otců spíše o něco menší, avšak synové velmi malých otců spíše větší — příroda se jaksi snaží přiblížit se k průměru. Přesto je tu zřejmá korelace: synové vysokých otců jsou většinou vyšší než synové malých otců. (Náš obrázek se nezakládá na skutečných měřeních, vyjadřuje jen zásadu.)



x = 1,05y + 0,1
není-li rovnost, je plocha čtverce menší

y = 0,83x + 0,3
není-li rovnost, je plocha čtverce menší





Regresní přímka odpovídá požadavku, že součet všech kvadratických odchylek jednotlivých bodů je tak malý, jak je to jen možné.

diferenciální počet. Ovšem i ty jsou tak zdlouhavé, že se omezíme na zcela elementární početní úkol. Do systému souřadnic jsme zanesli pět dvojic výsledků měření s těmito hodnotami: ($x = 1$, $y = 1$), (1,5, 2), (3, 2,5), (4, 3) a (5, 5).

Na první pohled se zdá, že jednoduchá rovnice $y = x$ by mohla docela dobře poskytnout vyhovující regresní přímku. Je to však nejlepší možná? Zatím nejdříve zjistíme střední hodnoty $\bar{x}$ a $\bar{y}$.

$\bar{x} = (1 + 1{,}5 + 3 + 4 + 5)\ /5 = 2{,}9;$

$\bar{y} = (1 + 2 + 2{,}5 + 3 + 5)\ /5 = 2{,}7.$

Nyní zase sestavíme známou tabulku:

Několik malých chyb jako důsledek zaokrouhlení neovlivní podstatně výsledek — zanedbali jsme setinové body do 0,05. Chceme-li nyní určit regresní přímku pro y, dosadíme nejdříve hodnoty do tohoto vzorce, který vyjadřuje směrnici přímky, tedy a v onom $ax + b$, které jsme poznali jako obecnou lineární rovnici:

$$a = \frac{\Sigma\,(x - \bar{x})\,(x - \bar{y})}{\Sigma\,(x - \bar{x})^2} = \frac{9{,}3}{11{,}2} = 0{,}83.$$

Rovnice $y = 0{,}83\,x$ udává tedy již správný směr přímky, my však ještě

x	y	$x - \bar{x}$	$y - \bar{y}$	$(x - \bar{x})^2$	$(y - \bar{y})^2$	$(x - \bar{x})\,(y - \bar{y})$
1	1	— 1,9	— 1,7	3,6	2,9	3,2
1,5	2	— 1,4	— 0,7	2	0,5	1
3	2,5	0,1	— 0,2	0	0	0
4	3	1,1	0,3	1,2	0,1	0,3
5	5	2,1	2,3	4,4	5,3	4,8
				11,2	8,8	9,3

přezkoumáme, zda nemusí být připočtena nebo odečtena nějaká konstanta b. Také pro tento účel existuje vzorec, a to dokonce velmi jednoduchý: $b = \bar{y} - a\bar{x}$.
Všechny tyto hodnoty již známe, takže jen dosazujeme:
$b = 2{,}7 - (0{,}83)(2{,}9) = 2{,}7 - 2{,}4 = 0{,}3$.
Jako konečnou regresní přímku pro y jsme nyní dostali:
$y = 0{,}83\,x + 0{,}3$.
Po začlenění do systému souřadnic vykazuje tato přímka ještě lepší přizpůsobení než dřívější „střelená od boku“ $y = x$ (obr. str. 300 dole). Jestliže se skutečné hodnoty spojí odpovídajícími svislými liniemi s regresní přímkou, dostaneme ony nejmenší vzdálenosti, které — umocněny — odpovídají hledaným „nejmenším čtvercům“. Každá jiná přímka by byla horší z hlediska rozptylu y.
Vzpomínáme si však, že x je možno posuzovat také jako závislý proměnlivý znak ypsilonu a že pak, pokud tu není funkční závislost, musí regresní přímka x mít jinou hodnotu. Pro určení této regresní přímky postupujeme úplně stejně jako předtím a zaměníme pouze ve vzorcích vždy x a y, případně $\bar{x}$ a $\bar{y}$. Pokud jde o míru směrnice, čitatel se nemění, ve jmenovateli se však nyní objeví $\Sigma\,(y - \bar{y})^2$, takže na závěr dostáváme: $a = \dfrac{9{,}3}{8{,}8} = 1{,}05$.

A pro „úsek osy“:
$b = 2{,}9 - (1{,}05)(2{,}7) = 0{,}1$.
To dává:
$x = 1{,}05\,y + 0{,}1$.
Zaneseme-li nyní tuto druhou regresní přímku do systému souřadnic, ukáže se, že svírá s první velmi ostrý úhel, což svědčí o velmi těsné korelaci mezi x a y (obr. str. 301). Již nyní můžeme říci, že korelační koeficient bude nejméně asi 0,9. Přesnější propočet lze provést snadno. Kovariance je již v čitateli našeho zlomku pro směrnici regresní přímky, i když zatím ještě ve formě součtů výsledků. Dělíme tedy ono 9,3 hodnotou $(n - 1 = 4)$ a dostáváme kovarianci 2,32. Obě potřebné hodnoty pro s_x a s_y můžeme snadno upravit z tabulkových hodnot:

$$s_x^2 = \frac{\Sigma\,(x - \bar{x})^2}{n - 1} = \frac{11{,}2}{4}$$

$$s_y^2 = \frac{\Sigma\,(y - \bar{y})^2}{n - 1} = \frac{8{,}8}{4}$$

To dává $s_x = 1{,}67$; $s_y = 1{,}48$ a tím

$$r = \frac{2{,}32}{(1{,}67)(1{,}48)} = 0{,}935.$$

Pro výpočet korelačních koeficientů je ostatně bezvýznamné, zda se součet mocnin dělí n nebo $(n - 1)$ anebo vůbec nedělí; je však nutno být důsledný a počítat jmenovatele i čitatele stejným způsobem. Dělením n dostáváme pak $s_{xy} = 1{,}86$; $s_x = 1{,}5$; $s_y = 1{,}33$ a znovu:

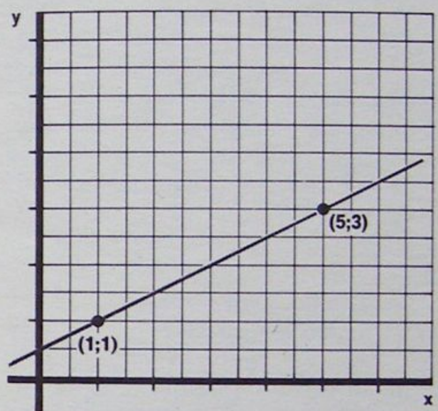
$$r = \frac{1{,}86}{(1{,}5)(1{,}33)} = 0{,}935.$$

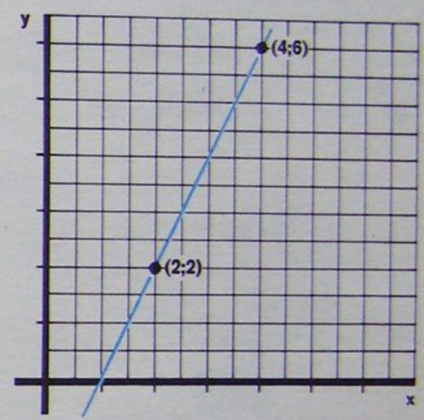
Nebo počítáme:

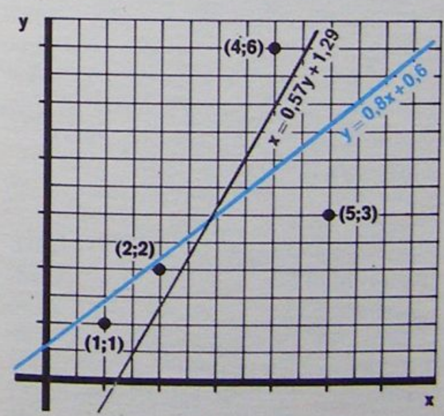
$$r = \frac{9{,}3}{\sqrt{11{,}2}\,\sqrt{8{,}8}} = \frac{9{,}3}{(3{,}35)(2{,}97)} = 0{,}935.$$

Zcela se potvrzuje naše domněnka, že korelace v tomto případě je velmi vysoká. Chápeme-li však tuto korelační analýzu jako průzkum vzorku, je nutno naléhavě varovat před spolehlivostí tohoto r. Čím méně dvojic měřených hodnot je k dispozici, tím lehčí je určit

regresní přímku, ale tím nejistější je také její hodnota. Jako varovný příklad nechť poslouží extrémní případ: výběrový soubor o rozsahu n dává po zanesení do systému souřadnic ve formě dvou bodů nádhernou regresní přímku, která přesně probíhá oběma těmito body; korelační koeficient $r = 1$ lze vyčíst na první pohled. Mimochodem není vůbec řečeno, že se výsledek regresní rovnice musí graficky znázorňovat vždy jen přímkou. Mezi proměnnými x a y může být zcela dobře nejtěsnější závislost, aniž by však regresní přímka musela plně odpovídat této skutečnosti. Uvažme jen třeba







$\bar{x} = \dfrac{12}{4} = 3 \qquad \bar{y} = \dfrac{12}{4} = 3$

	$(x-\bar{x})^2$	$(y-\bar{y})^2$	$(x-\bar{x})(y-\bar{y})$
(1;1)	4	4	4
(2;2)	1	1	1
(4;6)	1	9	3
(5;3)	4	0	0
	10	14	8

$a = \dfrac{8}{10} = 0{,}8 \quad b = \bar{y} - a\bar{x} = 3 - 2{,}4 = 0{,}6$

$y = 0{,}8x + 0{,}6$

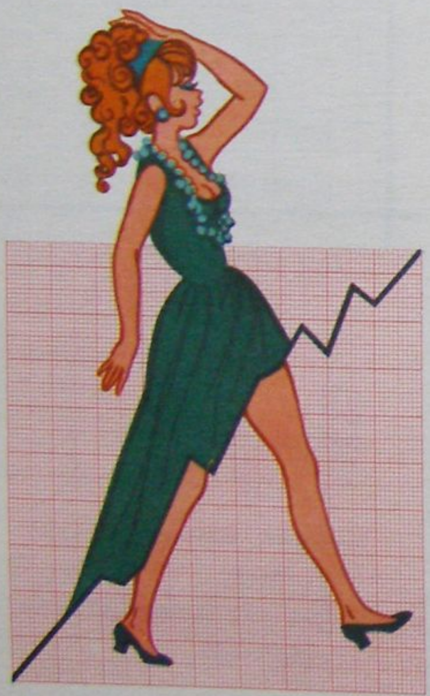
popř. $a = \dfrac{8}{14} = 0{,}57 \quad b = \bar{x} - a\bar{y} = 3 - 1{,}71 = 1{,}29$

$x = 0{,}57y + 1{,}29$

Dvěma dvojicemi měřených hodnot může vždy probíhat přímka, která předstírá dokonalou korelaci — tak, jak je tomu nahoře u dvojic měřených hodnot (1; 1) a (5; 3), případně (2; 2) a (4; 6). Jestliže však dáme tyto čtyři dvojice měřených hodnot do jednoho diagramu, tato dokonalá korelace zmizí; zůstanou jen dvě regresní přímky, z nichž žádná se nedotýká ani jediného měřeného bodu.

funkci $y = x^2 - 1$, o které jsme se zmínili na začátku oddílu 9.2. Nelineární regrese ovšem znamená, že aspoň jedna z obou proměnných nemá normální rozdělení.

S trochou fantazie a velkou matematickou erudicí je možno k téměř každé kombinaci měřených hodnot obstojně vypočítat odpovídající regresní křivku — od sinusové křivky jako $y = \sin x$ až po funkci třetího stupně, jako např. $y = x^3 - 6x^2 + 11x - 5$. Opovážlivý matematik může přitom až příliš lehce upadnout do pokušení hledat tvrdohlavě regresní křivku tak dlouho, až nakonec nějakou najde. Ta může sice stavu výběrového souboru vcelku odpovídat, pro posouzení závislostí v základním souboru ji však bude možno použít jen s nejvyšší opatrností.



Jedna z proslulých zdánlivých korelací: když se zkracuje délka sukní, kursy akcií stoupají. Nehledě k tomu, že to vždy nesouhlasí, šlo by určitě o zdánlivou nebo nesmyslnou korelaci.

Konečně budiž ještě zdůrazněno, že jsme se zabývali jen případem s dvěma proměnnými (x a y). Samozřejmě jsou také výpočetní metody pro vícenásobnou regresní analýzu, regresní počet s třemi a více proměnnými. Tyto metody jsou však nepoměrně obtížnější a v praxi je lze zvládnout v přijatelném čase pouze pomocí výpočetní techniky. V oblasti korelace není ale všechno jen rozumné, jak hned uvidíme v dalším oddílu.

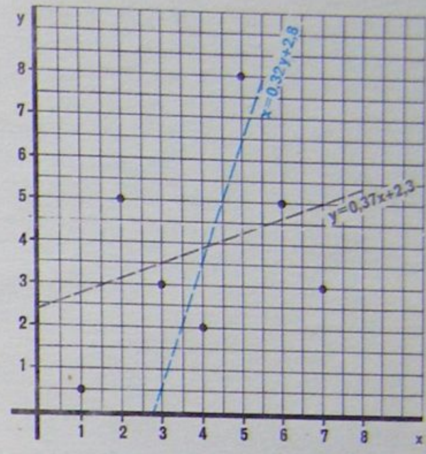
9.4 Zdánlivá korelace — logika a statistika

Jsou-li prokázány závislosti, zbývá většinou ještě otázka o příčině a účinku, o náhodě nebo hlubším významu, o přímé závislosti, společném třetím („podmíněná korelace") nebo klamném zdání. Této zdánlivé korelaci věnujeme na závěr malý oddíl, protože je tak matoucí, a tím nebezpečná, ale také proto, že je někdy i zábavná. V angličtině se o ní ostatně mluví většinou velmi jadrně jako o *„nonsense correlation"* (nesmyslné korelaci), v češtině bývá též nazývána *iluzorní korelací* nebo *pseudokorelací*.

Zdánlivou korelaci, která v posledních letech — zase jednou — prošla celým světovým tiskem, vytváří souvislost mezi burzovními kursy a délkou sukní podle zásady *„čím kratší sukně, tím vyšší kursy"*. Jestliže se nosí mini, na burze stoupají kursy. Někdy se tento žert uvádí se vší vážností: tak prý asi v roce 1967 zjistil jeden finanční poradce ve státě Massachusetts (USA), že tato korelace je dokazatelná za dobu 47 let. Toto tvrzení pravděpodobně není správné (a pokud jde o střední Evropu, byly kursy v roce zakrytých kolen 1962 mnohem vyšší než v roce minisukní 1970), ale i kdyby bylo prokazatelné, šlo by přece jen o *„nonsense correlation"*. Ledaže by snad šlo o teorii, že staří burziáni byli cestou na burzu kratičkými sukýnkami naladěni tak citlivě, že pak jednomyslně a s nadšením hnali kursy nahoru. Jiný příklad udává Helen M. Walkerová, autorka elementární učebnice statistiky ze čtyřicátých let: Zjistí se například, že úhel kladení chodidel na zem při chůzi je u starších žen většinou poněkud větší než u mladších. Stárnou ženy tedy proto, že pokládají chodidla více na vnější stranu? Sotva. Je to tedy snad obráceně — *s přibývajícím věkem se chodidla kladou více k vnější straně*. Vždyť to dokázala statistika! Ve skutečnosti však jde asi pouze o skutečnost, „že starší ženy vyrůstaly v době, kdy se mladé dívky učily klást při chůzi chodidla špičkami od sebe, mladší ženy se však učily držení těla v době, kdy podobná chůze již nebyla v oblibě".

Čistě početní korelace je (nebo snad správněji: byla) tedy zcela dána, jenže nemá žádný význam a klame.

Wallis a Roberts říkají proto zcela právem, že *zdánlivá korelace pochází většinou z nesprávného výkladu skutečně existujících korelací;* při bližším pohledu se ukáže jen jako nesprávná interpretace. Zmiňují se však také o *zdánlivé korelaci, která vzniká pouze při výpočtu*. Když například chceme usuzovat ze tří proměnlivých znaků X, Y a Z na závislosti, lze přijít na myšlenku porovnat poměr $X : Z$ s poměrem $Y : Z$ — a bude



Původní měřené hodnoty z našeho příkladu v textu ukazují krajně malou korelaci. Jestliže se však uvedou do vztahu k třetí měřené hodnotě, může vzniknout dojem vysoké korelace, která však ve skutečnosti neexistuje. Tato třetí hodnota se totiž výpočetním postupem vztahuje sama k sobě a zkresluje tak původní obraz.

možno vypočítat působivou korelaci, která se však uskutečňuje převážně tím, že se ve hře stále znovu uplatňuje hodnota $1/Z$: je-li Z velké, zmenší se X/Z i Y/Z. Je-li Z malé, zvětší se X/Z i Y/Z. Nesprávný početní postup vede v daném případě na scestí a ke zdánlivé korelaci.

Předpokládejme např., že bychom zjistili u tří proměnných — x, y a z — tyto kombinace měřených hodnot:

x	y	z
1	0,5	1
2	5	2
3	3	10
4	2	12
5	8	2
6	5	10
7	3	7

Zaneseme-li dvojice měřených hodnot (x, y) do systému souřadnic, sotva dostaneme věrohodnou regresní přímku; body jsou rozptýleny zcela neuspořádaně. (Viz obr. na str. 305.) Dáme-li však x, případně y do vztahu k z, dostaneme nejdříve tuto tabulku:

x/z	y/z
1	0,5
1	2,5
0,3	0,3
0,33	0,17
2,5	4
0,6	0,5
1	0,43

Z těchto hodnot dostaneme pak velmi působivou regresní přímku a můžeme nepochybně vypočítat i vysoký korelační koeficient. To všechno se však stalo pod vlivem Z, kterému se naše

V domečku pro předvídání počasí je přísná záporná korelace: čím více vystoupí jedna figurka z domečku, tím více ustoupí druhá zpět do domečku. To, co se zde sleduje, může se ve statistice vyskytnout jako nebezpečná myšlenková chyba: jestliže se soubor rozdělí na dva díly, tyto díly nutně vykazují korelaci —1.

zkoumání vlastně ani nemělo věnovat. S takovými metodami lze úplně mimoděk a jen na trochu vyšší úrovni upadnout do hádanek, kdy si někdo nejdříve myslí nějaké číslo, které pak musí několikrát násobit, dělit, něco připočíst nebo odečíst, až se konečně hledané číslo „uhodne".

Hrubá logická myšlenková chyba vzniká konečně i tehdy, když se nějaký soubor rozdělí na dílčí množství a jejich podíl se porovnává v určitém časovém intervalu: nutně musí dojít k výrazné korelaci, neboť jestliže ze dvou procentních podílů, které se navzájem doplňují na 100 %, jeden podíl stoupá, musí druhý klesat: mezi kuřáky a nekuřáky, mezi bílými a barevnými, mezi dny se srážkami a beze srážek bude vždy existovat výrazná záporná korelace tak jako u figurek ve starých *domečcích pro předpovídání počasí*, které jsou od sebe vždy vzdáleny o 180°, ať před domečkem zrovna stojí kterákoliv z nich.

Pravou „nonsense correlation" uvádí Kellerer: „Když se za posledních padesát let zjistí jednak dovoz *pomerančů* (= x), jednak počet úmrtí na *rakovinu* (= y) a každá takto získaná dvojice hodnot x_i, y_i se zanese do pravoúhlého souřadnicového systému, ukáže se, že s výjimkou válečných a bezprostředně poválečných let vysokým (nízkým) hodnotám x odpovídají také vysoké (nízké) hodnoty y. Přes formální souvislost zde není vůbec žádné spojení."

Poté následuje velmi závažná věta: *„Obecně platí, že při použití regresního počtu pro časové řady musí být u každé z obou veličin nejdříve vyznačen trend."* Pak se totiž nemůže stát, že se „prokáže" souvislost mezi cenami pálenky a platy pastorů nebo cenami kyselého zelí ve Vídni a počtem návštěvníků ze Švýcarska v tomto městě. Zbývá

však ještě nesmyslná náhodná korelace.

Je ovšem někdy velmi obtížné určit, do jaké míry je korelace nesmyslná. Jakmile je totiž statistickými metodami testování rozhodně zamítnuta nulová hypotéza „jde o náhodu", zůstane možná ještě „významná" korelace, kterou si nikdo nedovede vysvětlit, která však může mít docela dobře alespoň prognostickou hodnotu.

Vezměme třeba dosti zvláštní až bizarní případ, že na otázku po oblíbené barvě téměř všichni lidé, kteří trpí ledvinovými kameny, odpovědí „zelená", zatímco mezi ostatními pacienty je obliba jednotlivých barev rozdělena náhodně. Jakmile by byla statisticky zjištěna korelace „oblíbená barva zelená — ledvinové kameny", bylo by možno alespoň předpovídat, že mezi tisícem pacientů, jejichž oblíbenou barvou je zelená, bude s největší pravděpodobností také několik nemocných ledvinovými kameny, zatímco mezi těmi, kteří mají v oblibě jinou barvu, nebude asi ani jediný. Lékař by si pak mohl na základě odpovědi „červená" nebo „žlutá" hned říci: „Aha, ledvinové kameny nemá" — ačkoliv mezi oblíbenou barvou a nemocí ledvin není zjistitelná žádná logická nebo medicínská souvislost.

To je ostatně i otázka, která zajímala již Marca Tullia Cicerona a jeho bratra Quinta: zda se korelace (kterým se tehdy ještě tak neříkalo) smějí použít k prognózám, jestliže není vysvětlitelná souvislost. Marcus Tullius to odmítal, Quintus byl pro. My držíme spíše s Quintem. *Tím, že chybí vysvětlení, nestává se ještě prokázaná korelace vratkou.*

Primitivní korelace jsou často zjišťovány nebo potvrzovány v *prastarých selských a povětrnostních úslovích.* Tak zkušenost mnoha desítiletí bezpochyby ukázala, že opožděný začátek zimy znamená často také pozdější jaro, a tak byla jaksi zjištěna vysoká korelace mezi proměnlivými znaky „začátek zimy" a „začátek jara", která se ovšem nevyjadřovala koeficienty, nýbrž úslovím: *„Zelené vánoce — bílé velikonoce; bílé vánoce — zelené velikonoce."*

Při korelaci platí mnohem více než v jiné oblasti statistické práce toto: číselné výsledky, i když byly vypočteny z bezvadných podkladů, ještě nic nedokazují, nýbrž ukazují, upozorňují, vyzývají k vytváření hypotéz a jejich následnému testování.

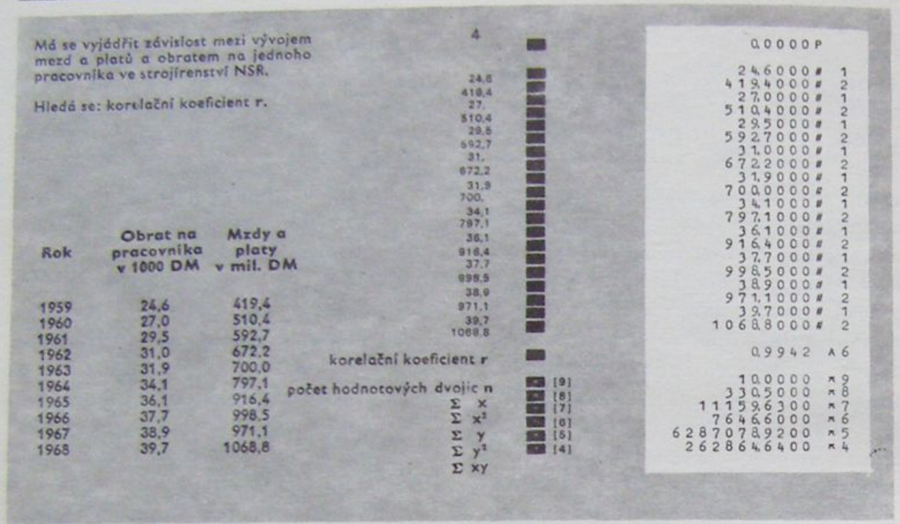
Nikde jinde není tak blízko nebezpečí, že uděláme ukvapené a nesprávné závěry nebo že se necháme nachytat na manipulované interpretace. Několika heslovitými příklady špatně pochopených nebo klamných korelací uzavřeme tento oddíl a necháme na čtenáři, aby sám odhalil nedostatky argumentace.

První příklad: mezi tělesnou váhou a četnými nemocemi existuje kladná korelace. Obézní budou proto zdravější, budou-li užívat odtučňovací tablety.

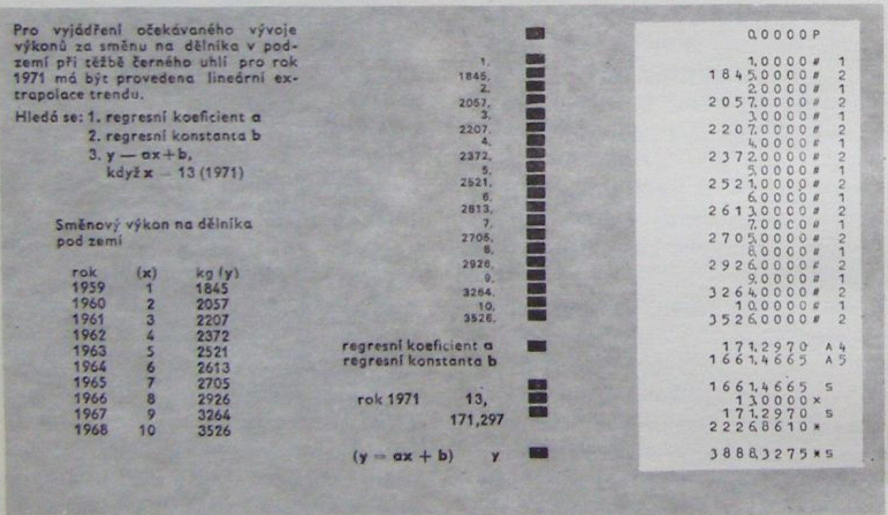
Druhý příklad: ve Švýcarsku je podíl katolíků mezi vinaři velmi vysoký. V NSR jsou zase sedláci v horách skoro bez výjimky katolíci, avšak námořníci jsou většinou evangelíci. Náboženské vyznání má tedy rozhodující vliv na volbu povolání.

Třetí příklad: v letech 1930 až 1950 byla v Anglii jasná korelace mezi počtem udělených rozhlasových koncesí a počtem pacientů v ústavech pro nervově choré. Ohlupující účinek rozhlasu doháněl obyvatelstvo houfně k šílenství.

Čtvrtý příklad: alkoholici vydělávají méně než abstinenti. Pití je činí ne-



Výpočet korelačního koeficientu za pomoci malého elektronického počítače s předprogramovanými statistickými funkcemi. Výpočet všech těch $(x_i - \bar{x})$ atd. provede počítač, který mimo výsledek $r = 0{,}9942$ vypočte také pomocné hodnoty. V tomto příkladu je to např. Σx součet deseti ročních obratů na 1 pracovníka, Σy součet deseti ročních údajů mezd a platů.



Regresní výpočet pomocí počítače. Podle schématu $x = ax + b$ se hledají pro $x = 13$ extrapolované směnové výkony na základě zkušeností z předchozích deseti let. Výsledek je: $x = 13 \,.\, 171{,}3 + 1661 = 3888$. Směnový výkon na dělníka se na základě vývoje z let 1959—1968 odhaduje pro rok 1971 na 3888 kg. (Z návodu k obsluze výpočetního systému „sigmatron".)

schopnými zastávat lépe placená místa. Nebo není tím přece jen spíše dokázáno, že odstrkování v zaměstnání vede k frustraci a ta k alkoholismu?

Pátý příklad: při jednom rozsáhlém šetření byla zjištěna zřetelná korelace mezi špatnou obytnou čtvrtí (staré domy, malé byty, špatné hygienické poměry atd.) a úmrtností na rakovinu. Kdo je chudý, musí tedy dříve umřít na rakovinu.

K tomuto poslednímu příkladu, který před několika lety vyvolal v jednom středoevropském velkoměstě značné vzrušení a který není tak lehce vysvětlitelný jako čtyři předcházející, přece jen dodáme, jak tato klamná korelace vznikla.

Byla to, jak již tak často, chyba v interpretaci. Skutečnou příčinu tvořila společná třetí proměnná: ve starých obytných čtvrtích bydleli nejen převážně chudší, ale *především převážně starší lidé.* Proto zde bylo úmrtí na rakovinu poměrně četnější než ve čtvrtích s novostavbami, kde bydlely v průměru značně mladší rodiny. Dodatečná analýza, provedená specificky z hlediska věku, neukázala významnější rozdíly, z nichž by vyplývala odpovědnost bytových poměrů za úmrtnost na rakovinu. „*Společná třetí*" tvoří jakýsi neznatelný přechod od zdánlivé korelace k důmyslné korelaci. V takových případech se často mluví o „*podmíněné korelaci*". Pfanzagl referuje např. o zjišťování mezi školáky ve věku od 6 do 10 let, které se týkalo korelace mezi tělesnou váhou a manuální zručností. Vyšel korelační koeficient $r = 0{,}45$, tedy dosti zřetelná závislost, která nejdříve překvapí: těžší, robustnější děti mají být šikovnější? Ve skutečnosti to byla korelace obou znaků s věkem: starší děti jsou šikovnější i těžší v poměru k mladším. Proto musí být nejdříve tento faktor početně vyloučen, resp. viditelně vyjádřen. Výsledek po „věkovém vyrovnání" vykázal pak podle očekávání nepatrně zápornou korelaci: mezi stejně starými jsou tlustší poněkud nešikovnější.

Na závěr ještě jedna zdánlivá korelace zvláštního druhu: *nepoměrně více lidí umírá v posteli* než na ulici, na moři nebo ve vykřičené krčmě. Postel je proto prokazatelně nejnebezpečnější místo pobytu.

10 Hledání jistoty

10.1 Hypotézy a kontingence

Moderní statistika sestává z velké části z *testování hypotéz*. Podle výsledků výběrových zkoumání odhaduje *rozdělení četností* nebo *korelací* v základním souboru, a tím zároveň dává doporučení pro chování, pomůcku pro *rozhodování:* tento lék se zdá být významně účinnější, tato nerovnoměrnost výroby nemůže být jen nahodilá, tuto shodu dvou jevů je však možno vysvětlit působením náhody.
Na základě prvních výsledků výběrových zkoumání, a někdy také jen na základě tvůrčí fantazie, se pak vytvářejí hypotézy: „Zde by přece mohla být souvislost". — „To by přece mělo ukázat důsledky." — „Jestliže se tedy předpokládá, že je to pouze náhoda..."
Potom se tyto hypotézy podrobí statistickým testům a skutečnost se vymezuje stále úžeji, pravděpodobnost pro správné, rozumné chování se stále zvětšuje.
O mnohých zásadách testování hypotéz jsme se již dověděli v šesté kapitole, kdy se jednalo o výrobní kontrole, přejímací kontrole, o analýze vzorků všeobecně a o nejmenších vzorcích. Tak jsme mimo jiné poznali princip *nulové hypotézy:* všechno se tak dlouho připisuje náhodě, až se překročí jakýsi práh (ne)pravděpodobnosti.
Pojednali jsme tam také o problému testů hypotéz, při nichž šlo o zjištění, v jakých mezích pravděpodobně platí střední hodnota (nebo četnost) určená ve vzorku i pro základní soubor. K tomu bude třeba dodat již jen málo. Často však jde při ověřování hypotéz o testování *pravděpodobnosti a významu (těsnosti) korelace*. Pacienti spí po léku A o půl hodiny déle než po přípravku B — lze to ještě vysvětlit jako náhodu? Rozvedení muži si berou rozvedené ženy častěji, než by to vlastně odpovídalo podílu rozvedených na počtu lidí uzavírajících sňatek — je to snad jen náhoda? Korelační koeficient je 0,3 — je to v daném případě postačující pro zamítnutí nulové hypotézy? Není snad možno pozorovanou korelaci vztáhnout na dva nebo více faktorů místo pouze na jeden — a jak je asi vysoký podíl jednotlivých faktorů na pozorovaném účinku?
To jsou zhruba charakteristické otázky při testování hypotéz. Odpovědi na ně ovlivňují v každém okamžiku někde na světě rozhodnutí, která mohou podstatně ovlivnit i náš osud. Ať již jde o pevnost stavebních hmot, o vedlejší účinky léků, o jakost profilů pneumatik, o nové způsoby školního vzdělání nebo o neškodnost konzervačního prostředku, vždy jde o získání údajů z experimentů a o jejich ověření statistickými testy z hlediska jejich vypovídací schopnosti.
Nejdříve na příkladu ukážeme, jak v praxi vypadá náhodné rozdělení dvou proměnných. Sestavíme tzv. *kontingenční tabulku*. Slovo *kontingence* se do statistiky dostalo přes angličtinu z latiny — kontingence znamená téměř doslova *setkání, spojení*. V takové tabulce se tedy zaznamenávají výsledky,

které vycházejí ze spojení dvou řad znaků. Z rakouské statistické ročenky vybereme např. tabulku, jež ukazuje rodinný stav nevěsty a ženicha před uzavřením sňatku:

Sňatky v roce 1968 podle rodinného stavu ženicha a nevěsty

Rodinný stav ženicha před sňatkem	Rodinný stav nevěsty před sňatkem				Rodinný stav nevěsty před sňatkem			
	svobodná	ovdovělá	rozvedená	celkem	svobodná	ovdovělá	rozvedená	celkem
	prvotní údaje				ze 100 sňatků			
svobodný	44 833	391	2343	47 567	80,0	0,7	4,2	84,9
ovdovělý	978	581	560	2 119	1,8	1,0	1,0	3,8
rozvedený	3 405	375	2535	6 315	6,1	0,7	4,5	11,3
Celkem	49 216	1347	5438	56 001	87,9	2,4	9,7	100,0

Z tabulky samé nemůže neškolené oko mnoho vyčíst: že u převážného počtu sňatků jsou oba partneři svobodní, není překvapující. I když vypočítáme rozdíly rodinného stavu, nebude ještě nic zvláštního nápadné. Z ženících se mužů bylo 84,9 % svobodných, 3,8 % ovdovělých, 11,3 % rozvedených. Z žen bylo 87,9 % svobodných, 2,4 % ovdovělých a 9,7 % rozvedených.
U mužů mají vdovci a rozvedení poněkud vyšší podíl; to je nyní zcela zřetelné. Nyní, když známe tyto podílové četnosti pro oba partnery, sestavíme kontingenční tabulku nulové hypotézy, tj. předpokládáme, že jde pouze o nahodilost, která podmiňuje rodinný stav manželských partnerů. Protože například 9,7 % nevěst je rozvedených, připadá na 1000 ženichů 97 rozvedených žen. Z 1000 ženichů je však 849 svobodných, takže $\frac{849}{1000}$ z 97 rozvedených žen se provdá za svobodné muže. Protože se nedoporučuje trhat ženy před oddávajícím úředníkem na kousky, řekněme, že $\frac{849 \,.\, 97}{1000}$ se rovná 82. Mělo by se tedy 82 rozvedených žen provdat za svobodné muže, 11 za rozvedené a 4 za vdovce.
Dobře si uvědomme, že je to naše nulová hypotéza čistého nahodilého rozdělení. Je to model házení kostkou: kostkou „rodinný stav nevěsty" se hází nezávisle na kostce „rodinný stav ženicha". Protože tedy domněle jde o nezávislé jevy, platí pravidlo o násobení pravděpodobnosti: pravděpodobnost „svobodný muž" ($p_{sm} = 0{,}849$) násobená pravděpodobností „rozvedená žena" ($p_{rž} = 0{,}097$) dává celkovou pravděpodobnost „svazek mezi svobodným mužem a rozvedenou ženou".
$p_{sm} \cdot p_{rž} = (0{,}849) \,.\, (0{,}097) = 0{,}0825$, tedy, jestliže opět odmítáme roztrhat ženu, 82 nebo 83 takových manželství na 100 sňatků.
Nyní sestavíme *hypotetickou kontingenční tabulku pro všechny myslitelné případy* (v %):

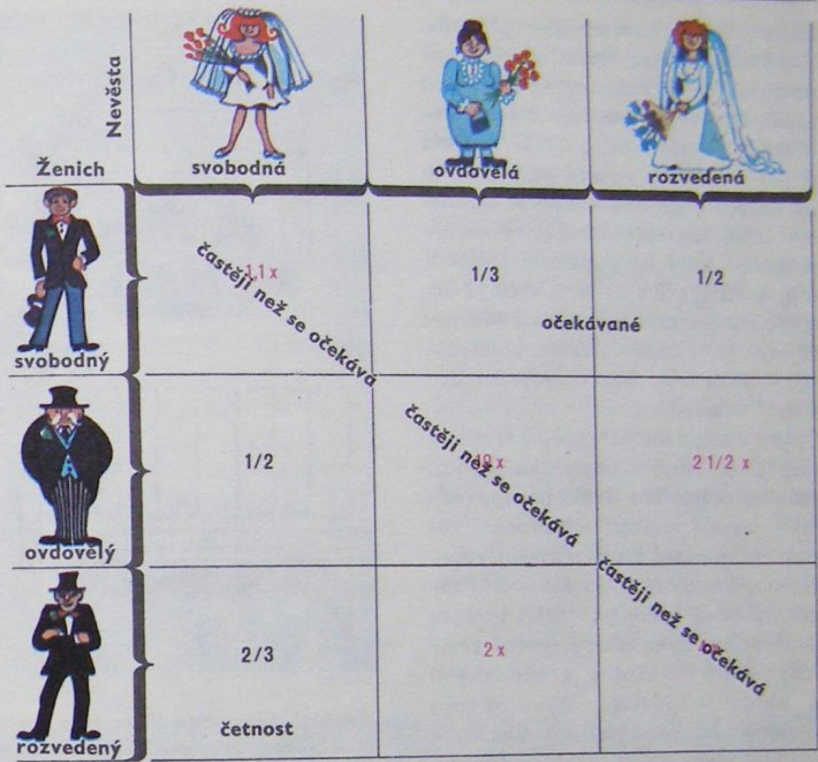
Ženich	Nevěsta			
	svo-bod-ná	ovdo-vělá	roz-vede-ná	celkem
svobodný	74,7	2,0	8,2	84,0
ovdovělý	3,3	0,1	0,4	3,8
rozvedený	9,9	0,3	1,1	11,3
Celkem	87,9	2,4	9,7	100,0

Proti této kontingenční tabulce nulové hypotézy („čistá náhoda") postavíme nyní na procenta přepočítanou původní tabulku skutečných četností:

Ženich	Nevěsta		
	svobod-ná	ovdo-vělá	roz-vedená
svobodný	80,0	0,7	4,2
ovdovělý	1,8	1,0	1,0
rozvedený	6,1	0,7	4,5

Rozdíl mezi hypotetickým a skutečným rozdělením je zřejmý zejména tehdy, jestliže se porovná úhlopříčka — tedy dvojice, kde se druží „rovný s rovným". Např. u oboustranně ovdovělých dvojic je skutečná četnost desetkrát větší než podle nulové hypotézy. (Žádný div: ovdovělé osoby jsou většinou již starší

Pravé náhodné rozdělení sňatků by vzniklo, kdyby se ovdovělí, rozvedení a svobodní k sobě nechali dotápat naslepo. V běžném životě se však ukazuje, že o náhodném rozdělení nemůže být ani řeči: s oblibou se sdružuje rovná s rovným. Zkoumá se, kdo se k sobě více nebo méně hodí.



Proti hypotetickému náhodnému rozdělení vykazují sňatky nápadné rozdíly. Zejména ovdovělí a rozvedení se navzájem evidentně přitahují, zatímco vdovci si berou nápadně zřídka svobodné ženy a svobodní muži zřídkakdy ženy rozvedené.

a častěji volí opět partnera usedlejšího ročníku, který ovšem bývá také často ovdovělý. To však je jen pokus o vysvětlení, který není v žádné přímé souvislosti s číselnými podklady: jak víme, korelace nutně neznamená poměr příčina — účinek. Nejdříve se zjistí neočekávaně velká četnost a potom si musí statistik nebo odborník v příslušné specializované oblasti lámat hlavu nad příčinou.)

Odkaz na neočekávaně vysoké hodnoty v úhlopříčce kontingenční tabulky nám však ještě neposkytuje početní rozbor významnosti nebo bezvýznamnosti. K tomuto účelu musíme použít testovacích metod, s nimiž se seznámíme, neboť porovnání *mezi hypotetickým a skutečným rozdělením* je hlavním úkolem tzv. χ^2 testu (chí kvadrát testu).

Nejdříve si však ještě jednou připomeneme pojem *„stupeň volnosti"*, se kterým jsme se poprvé setkali rovněž při testování hypotéz, a to při *„rozdě-*

lení t". Kdybychom chtěli v příkladu, který zkoumáme, hledat stupeň volnosti, museli bychom např. u mužů říci: především máme volnost určení svobodných — mohlo by to být 90 %. Dále si dovolíme určit rozvedené — mohlo by to být 8 %. A konečně si dovolíme určit libovolně i ovdovělé — ale můžeme to ještě udělat? Máme-li 8 % + 90 % celku a máme určit již jen jednu kategorii, žádná volnost nám nezbývá: je přesně určena poměrem 100 — 98 = 2 %. Máme tedy jen dva stupně volnosti.
Stejně je tomu samozřejmě i u žen: dva stupně volnosti. Pro celou kontingenční tabulku dostáváme dvakrát dva, tedy čtyři stupně volnosti. Nakonec poznámku k *nulové hypotéze*. Není vůbec řečeno, že nulová hypotéza musí vždy mít náhodné rozdělení. Mohli bychom např. nalézt zemi, kde rozvedení téměř vždy uzavírají sňatek s ovdovělými a obráceně. Nulová hypotéza pro zkoumání číselných podkladů Rakouska by pak mohla znít: „Předpokládejme, že by tomu zde bylo také tak — je možno údaje jiného druhu vysvětlit pouhou náhodou?"
Protože tedy máme při volbě testované hypotézy takovou volnost, naskýtá se další velmi důležitá možnost. Výsledek prvního výběrového souboru mohu pokládat za nulovou hypotézu pro výsledek druhého výběrového souboru tak, že řeknu: „Předpokládejme, že by druhý výběrový soubor byl vytvořen ze stejného základního souboru jako první, to však nevím přesně. Má se tento předpoklad zamítnout?"
Jestliže jsem oba výběrové soubory domněle vytvořil ze stejného obyvatelstva, pak významný rozdíl dovolí domněnku, že se ve skutečnosti nejedná o stejný základní soubor. Jestliže jsem

Ještě na nemocničním lůžku, když v zamyšlení hází kostkou a dosáhne překvapivého výsledku, sáhne statistik po tabulce chí kvadrát. Test chí kvadrát ukazuje totiž většinou velmi přesně hranice (meze) náhody.

však — a to je u testování hypotéz častější případ — výběr provedl ze dvou podobných, nikoliv však stejných základních souborů, může prohlášení nulové hypotézy za falešnou poskytnout cenné poznatky o účinnosti rozdílných pracovních postupů, jakosti materiálu, přípravků atd.
Při testování hypotéz pak často otázka zní takto: odpovídá pozorované rozdělení očekávanému? Ve většině případů odpověď na to dává *test* χ^2.

10.2 Test chí kvadrát (χ^2)

Polehávající nemocný statistik hází zamyšleně kostkou po podnosu — třicetkrát, padesátkrát, šedesátkrát. Potom sečte zaznamenané výsledky: sedmkrát padla jednička, devětkrát dvojka, desetkrát trojka, šestkrát čtyřka, patnáctkrát pětka a třináctkrát šestka. „Hmhm," bručí sám pro sebe, „šestkrát čtyřka, ale patnáctkrát pětka? Je to vůbec možné?" A znovu vezme tužku. Mohl by například pomocí směrodatné odchylky binomického rozdělení ($\sigma = \sqrt{npq}$) vypočítat, že pro každý počet ok je směrodatná odchylka $\sqrt{60 \cdot 1/6 \cdot 5/6} = \sqrt{8{,}35} = 2{,}9$, a z toho by mohl poznat, že každý počet ok má četnost ještě uvnitř dvojnásobné směrodatné odchylky — totiž mezi (10 — 5,8 = = 4,2) a (10 + 5,8 = 15,8); zvolí však elegantnější a přesnější metodu. Poznamená si:

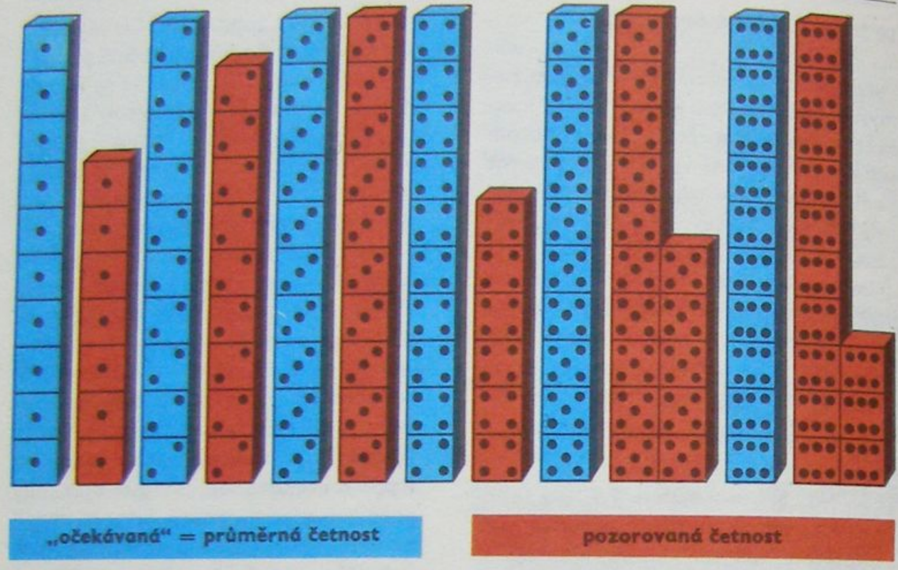
síme si postup zopakovat. Co se stalo? *Byly vypočítány rozdíly*, a to zcela podobným způsobem, s jakým jsme se seznámili na počátku druhé kapitoly při výpočtu rozptylu. Měřená hodnota nebyla ovšem porovnána s aritmetickým průměrem všech pozorovaných hodnot, nýbrž s hypotetickou očekávanou hodnotou, očekávanou četností. Ta byla ve všech případech vždy 10, neboť — protože kostka má 6 ploch a u správné kostky je pravděpodobnost pro každý jednotlivý počet ok 1/6 — očekávaná hodnota (která je podle definice průměrem) činí při celkem 60 hodech vždy 10 pro každý počet ok. (Obr. str. 316.) S touto očekávanou četností byla tedy porovnána pozorovaná četnost, rozdíl byl umocněn známým způsobem, ale pak dělen očekávanou četností. Přitom se umocněním zvlášť zřetelně ukázaly extrémní odchylky: rozdíl 1 dal nakonec pouze 0,1, rozdíl 5 však 2,5.

Počet ok	Četnost očekávaná	Četnost pozorovaná	Rozdíl	Mocnina rozdílu	Mocnina rozdílu dělená očekávanou četností
1	10	7	— 3	9	0,9
2	10	9	— 1	1	0,1
3	10	10	0	0	0
4	10	6	— 4	16	1,6
5	10	15	5	25	2,5
6	10	13	3	9	0,9
					6,0

„Přesně šest," mumlá si statistik a sáhne po tabulce chí kvadrát, kterou má samozřejmě u postele. „Šest a pět stupňů volnosti." Nahlédnutí do tabulky a již se spokojeně usmívá: „Vše v pořádku, kostka je prima."
Zjistil totiž, že předpoklad, že kostka je v pořádku, nebyl testem chí kvadrát zamítnut. Dříve však než se i my budeme spokojeně usmívat, mu-

Nakonec byly mocniny odchylek sečteny a dostali jsme číslo 6 jako kritickou veličinu pro tabulku chí kvadrát. (Obr. str. 317.) Náš výsledek jsme dostali podle tohoto vzorce:

$$\chi^2 = \sum \frac{(B - E)^2}{E}$$



Porovnání očekávané a pozorované četnosti je základem pro výpočet chí kvadrátu. Při 60 hodech „správné“ kostky je očekávaná hodnota každé oko 60/6 = 10. Pozorované četnosti v našem příkladu se od sebe více nebo méně odchylují.

Slovy: *chí kvadrát je součet všech umocněných rozdílů mezi pozorovanými (B) a očekávanými (E) hodnotami, dělený očekávanou (E) četností.* Dostali jsme tedy $\chi^2 = 6$. A co teď s tímto číslem?

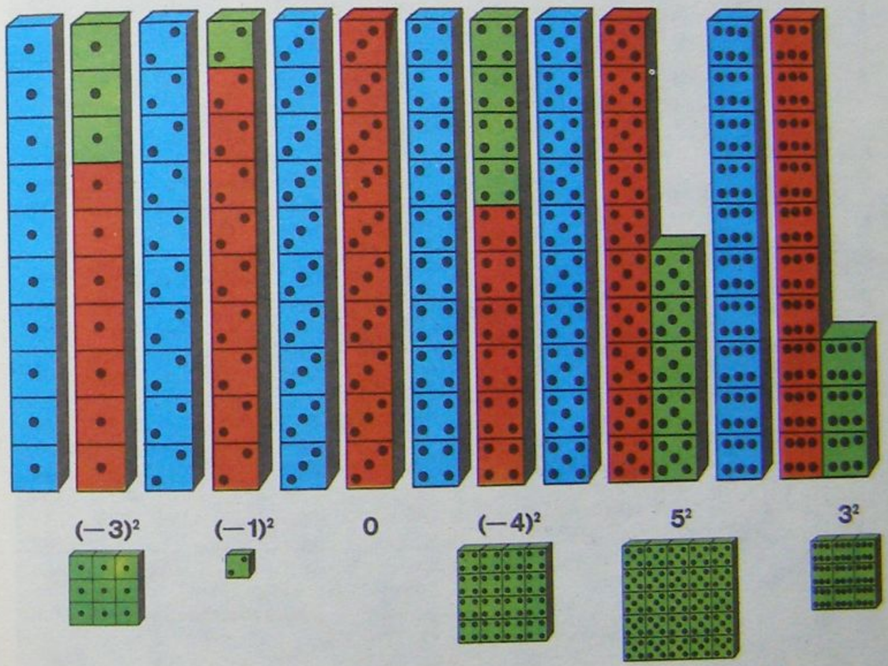
Stejně jako v případě „rozdělení t“ pouze toto číslo samo nestačí ještě ke zjištění jeho významnosti z tabulky. Mimo toto číslo χ^2 *musí být dále znám počet stupňů volnosti*. V našem případě dostaneme $6 - 1 = 5$ stupňů volnosti, protože při celkem šesti možnostech je poslední určována ostatními pěti (jak jsme podrobněji vysvětlili v oddílu 6. 41). V tabulce rozdělení chí kvadrát vyhledáme řádek pro „pět stupňů volnosti“ a sloupec s hledanou významností. Chceme-li například zkoumat, zda je náš výsledek v 95 % všech náhodných výsledků správné kostky, vyhledáme ve sloupci 95 % řádek 5 a najdeme tam porovnávací hodnotu 11,1 — a to je mnohem více než 6,0, které jsme dostali při našem pokusu. Nulová hypotéza („kostka je v pořádku, všechno byla pouze náhoda“) nebyla v úrovni 95 % zamítnuta. (Viz obr. na str. 318.) Zkusíme-li znovu své štěstí při hladině významnosti 97,5 %, dostaneme samozřejmě ještě vyšší hodnotu, protože jsme tím rozšířili prostor náhodnosti nulové hypotézy. Pro zúžení prostoru bychom museli hledat jiným směrem, v tomto případě se však ukazuje, že např. i 80 % pravděpodobnosti je ještě u 7,3, takže náš výsledek nejen není nevýznamný, ale není ani zcela neobvyklý.

Tabulky chí kvadrát mají však ještě

zvláštní vlastnost: *kromě krajně vysoké a 50% pravděpodobnosti obsahují i pravděpodobnosti krajně nízké* — a to má svůj účel. (Viz tabulku na str. 319.) To, co jsme nazývali nulovou hypotézou, je u testu chí kvadrát nulovou hypotézou zvláštního druhu — je to hypotéza, že „se hodí“ hypotetické rozdělení. Čekali jsme, že se kostka bude chovat jako pravá kostka, že se tedy bude řídit zákony binomického rozdělení. Porovnávali jsme pozorované hodnoty s „očekávanými“ hodnotami binomického rozdělení a naše domněnka, že snad takové rozdělení existuje, nebyla zamítnuta.

Taková domněnka však může být zamítnuta nebo alespoň zpochybněna také jiným způsobem — totiž tehdy, jestliže pozorovaný výsledek „se hodí“ až příliš dobře, když náhoda jí vyhrazený prostor takříkajíc vůbec nevyža-

Rozdíl mezi očekávaným a pozorovaným



$$\frac{\text{mocniny rozdílu:}}{\text{očekávané:}}\quad \frac{9}{10}+\frac{1}{10}+\frac{0}{10}+\frac{16}{10}+\frac{25}{10}+\frac{9}{10}=\frac{60}{10}=6=\chi^2$$

Ve druhém kroku výpočtu chí kvadrátu zjistíme „chybějící“, případně „přebytečné“ hody podle jednotlivých počtů ok a potom je umocníme. Součet těchto umocněných odchylek dává hodnotu chí kvadrátu.

Rozdělení χ^2

n \ α	1 %	5 %	50 %	95 %	99 %
1	0,0	0,0	0,5	3,8	6,6
2	0,0	0,1	1,4	6,0	9,2
3	0,1	0,4	2,4	7,8	11,3
4	0,3	0,7	3,4	9,5	13,3
5	0,6	1,1	4,4	11,1	15,1
6	0,9	1,6	5,4	12,6	16,8
7	1,2	2,2	6,4	14,1	18,5
8	1,6	2,7	7,3	15,5	20,1
9	2,1	3,3	8,3	16,9	21,7
10	2,6	3,9	9,3	18,3	23,2
11	3,0	4,6	10,3	19,7	24,7
12	3,6	5,2	11,3	21,0	26,2
13	4,1	5,9	12,3	22,4	27,7
14	4,7	6,6	13,3	23,7	29,1
15	5,2	7,3	14,3	25,0	30,6
16	5,8	8,0	15,3	26,3	32,0
17	6,4	8,7	16,3	27,6	33,4
18	7,0	9,4	17,3	28,9	34,8
19	7,6	10,1	18,3	30,1	36,2
20	8,3	10,9	19,3	31,4	37,6
21	8,9	11,6	20,3	32,7	38,9
22	9,5	12,3	21,3	33,9	40,3
23	10,2	13,1	22,3	35,2	41,6
24	10,9	13,8	23,3	36,4	43,0
25	11,5	14,6	24,3	37,7	44,3
26	12,2	15,4	25,3	38,9	45,6
27	12,9	16,2	26,3	40,1	47,0
28	13,6	16,9	27,3	41,3	48,3
29	14,3	17,7	28,3	42,6	49,6
30	15,0	18,5	29,3	43,8	50,9
34	17,8	21,7	33,3	48,6	56,1
40	22,2	26,5	39,3	55,8	63,7
50	29,7	34,8	49,3	67,5	76,2
70	45,4	51,7	69,3	90,5	100,4
100	70,1	77,9	99,3	124,3	135,8

Rozdělení chí kvadrát

Stupně volnosti	Pravděpodobnost[1] 20 %	10 %	5 %	2 %	1 %	0,1 %
1	1,64	2,70	3,84	5,41	6,63	10,82
2	3,21	4,60	5,99	7,82	9,21	13,81
3	4,64	6,25	7,81	9,83	11,34	16,26
4	5,98	7,77	9,48	11,66	13,27	18,46
5	7,28	9,23	11,07	13,38	15,08	20,51
6	8,55	10,64	12,59	15,03	16,81	22,45
7	9,80	12,01	14,06	16,62	18,47	24,32
8	11,03	13,36	15,50	18,16	20,09	26,12
9	12,24	14,68	16,91	19,67	21,66	27,87
10	13,44	15,98	18,30	21,16	23,20	29,58
11	14,63	17,27	19,67	22,61	24,72	31,26
12	15,81	18,54	21,02	24,05	26,21	32,90
13	16,98	19,81	22,36	25,47	27,68	34,52
14	18,15	21,06	23,68	26,87	29,14	36,12
15	19,31	22,30	24,99	28,25	30,57	37,69
16	20,46	23,54	26,29	29,63	32,00	39,25
17	21,61	24,76	27,58	30,99	33,40	40,79
18	22,76	25,98	28,86	32,34	34,80	42,31
19	23,90	27,20	30,14	33,68	36,19	43,82
20	25,03	28,41	31,41	35,02	37,56	45,31
21	26,17	29,61	32,67	36,34	38,93	46,79
22	27,30	30,81	33,92	37,65	40,28	48,26
23	28,42	32,00	35,17	38,96	41,63	49,72
24	29,55	33,19	36,41	40,27	42,98	51,17
25	30,67	34,38	37,56	41,56	44,31	52,62
26	31,79	35,56	38,88	42,85	45,64	54,05
27	32,91	36,74	40,11	44,14	46,96	55,47
28	34,02	37,91	41,33	45,41	48,27	56,89
29	35,13	39,08	42,55	46,69	49,58	58,30
30	36,25	40,25	43,77	47,96	50,89	59,70

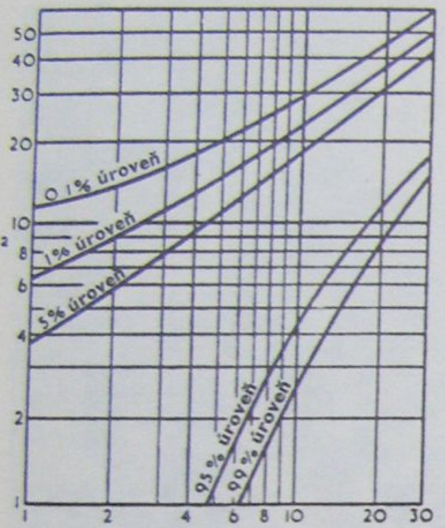
[1] Procentní pravděpodobnost dostat určitou v tabulce uvedenou nebo vyšší hodnotu chí kvadrátu.

Tři různé tabulky rozdělení chí kvadrátu podle Kreyszigа, Haseloffa-Hoffmanna a Pfanzagla. Při správném použití dávají samozřejmě vždy správné hodnoty.

Rozdělení chí kvadrát

$F(x)$	Počet stupňů volnosti 1	2	3	4	5	6	7	8	9	10
0,001	0,00	0,00	0,02	0,09	0,21	0,38	0,60	0,86	1,15	1,48
0,005	0,00	0,01	0,07	0,21	0,41	0,68	0,99	1,34	1,73	2,16
0,01	0,00	0,02	0,11	0,30	0,55	0,87	1,24	1,65	2,09	2,56
0,025	0,00	0,05	0,22	0,48	0,83	1,24	1,69	2,18	2,70	3,25
0,05	0,00	0,10	0,35	0,71	1,15	1,64	2,17	2,73	3,33	3,94
0,1	0,02	0,21	0,58	1,06	1,61	2,20	2,83	3,49	4,17	4,87
0,25	0,10	0,58	1,21	1,92	2,67	3,45	4,25	5,07	5,90	6,74
0,5	0,45	1,39	2,37	3,36	4,35	5,35	6,35	7,34	8,34	9,34
0,75	1,32	2,77	4,11	5,39	6,63	7,84	9,04	10,22	11,39	12,55
0,9	2,71	4,61	6,25	7,78	9,24	10,64	12,02	13,36	14,68	15,99
0,95	3,84	5,99	7,81	9,49	11,07	12,59	14,07	15,51	16,92	18,31
0,975	5,02	7,38	9,35	11,14	12,83	14,45	16,01	17,53	19,02	20,48
0,99	6,63	9,21	11,34	13,28	15,09	16,81	18,48	20,09	21,67	23,21
0,995	7,88	10,60	12,84	14,86	16,75	18,55	20,28	21,96	23,59	25,19
0,999	10,83	13,82	16,27	18,47	20,52	22,46	24,32	26,13	27,88	29,59

$F(x)$	Počet stupňů volnosti 11	12	13	14	15	16	17	18	19	20
0,001	1,83	2,21	2,62	3,04	3,48	3,94	4,42	4,90	5,41	5,92
0,005	2,60	3,07	3,57	4,07	4,60	5,14	5,70	6,26	6,84	7,43
0,01	3,05	3,57	4,11	4,66	5,23	5,81	6,41	7,01	7,63	8,26
0,025	3,82	4,40	5,01	5,63	6,26	6,91	7,56	8,23	8,91	9,59
0,05	4,57	5,23	5,89	6,57	7,26	7,96	8,67	9,39	10,12	10,85
0,1	5,58	6,30	7,04	7,79	8,55	9,31	10,09	10,86	11,65	12,44
0,25	7,58	8,44	9,30	10,17	11,04	11,91	12,79	13,68	14,56	15,45
0,5	10,34	11,34	12,34	13,34	14,34	15,34	16,34	17,34	18,34	19,34
0,75	13,70	14,85	15,98	17,12	18,25	19,37	20,49	21,60	22,72	23,83
0,9	17,28	18,55	19,81	21,06	22,31	23,54	24,77	25,99	27,20	28,41
0,95	19,68	21,03	22,36	23,68	25,00	26,30	27,59	28,87	30,14	31,41
0,975	21,92	23,34	24,74	26,12	27,49	28,85	30,19	31,53	32,85	34,17
0,99	24,73	26,22	27,69	29,14	30,58	32,00	33,41	34,81	36,19	37,57
0,995	26,76	28,30	29,82	31,32	32,80	34,27	35,72	37,16	38,58	40,00
0,999	31,26	32,91	34,53	36,12	37,70	39,25	40,79	42,31	43,82	45,32



Tabulka chí kvadrát v grafické podobě podle Fisherových a Yatesových „Statistických tabulek". Deset stupňů volnosti dává na úrovni 5 % $\chi^2 = 19$. Tabulka udává přesnou hodnotu 18,3. (Co zde znamená úroveň 5 %, uvádí se ve většině tabulek jako 95% úroveň atd.)

duje. Kdybychom totiž v průběhu 60 hodů skutečně hodili každé ze šesti čísel přesně desetkrát, byl by to opravdu neobvyklý případ náhody. Výpočet by samozřejmě dal $\chi^2 = 0$ a pohled do tabulky by ukázal, že (5 stupňů volnosti, 1 %) se vyžaduje hodnota $\chi^2 = 0{,}6$. Jinými slovy: je nasnadě podezření, že se neuplatnilo binomické rozdělení, protože se manipulovalo. V případě naší kostky to snad byla přece jen vzácná náhoda, při jiných pokusech však může překvapivě dobře „se hodící" výsledek být testem chí kvadrát *odhalen jako „významně podezřelý"*. (Pro test chí kvadrát platí ostatně jako pro „rozdělení t": opatrně při používání tabulek, zvláště ze starších prací — procenta jsou někdy „zrcadlově převrácena", místo stupňů volnosti jsou uvedeny třídy znaků.)

Také rozdělení chí kvadrát lze nakreslit jako křivku, má však — podle počtu stupňů volnosti — různý tvar. Mezi ním, rozdělením t, rozdělením F (se kterým se ještě krátce seznámíme) a normálním rozdělením jsou zajímavé, ale složité souvislosti, o jejichž popis se zde nebudeme pokoušet. Rozdělení χ^2 poprvé objevil v roce 1876 německý matematik a geodet Friedrich Robert Helmert, upadlo však úplně v zapomenutí, až je v roce 1900 nezávisle na Helmertovi znovu objevil a pro použití v matematické statistice připravil Karl Pearson, kterého již známe jako tvůrce korelačního koeficientu. Teprve když tento Pearsonův objev vyvolal pozornost a posunul statistiku o veliký krok kupředu, zjistil Bortkievicz, že Helmert již dříve formuloval podobné myšlenky.

Rozdělení χ^2 má v praxi testování hypotéz nepřecenitelnou úlohu, protože je schopno určit, *zda ověřovaný názor na rozdělení bude nebo nebude zamítnut šetřením*. Přitom je možno pracovat s malými, ne však nejmenšími výběry, což je pochopitelné, protože k poznání struktury rozdělení je třeba poměrně mnoho jednotlivých údajů.

Vzhledem k významu tohoto testu uvedeme ještě jiný příklad, pro který však použijeme údajů o sňatcích z předcházejícího oddílu. Předpokládejme, že bychom údaje tam uvedené pro základní soubor zjistili v odpovídající četnosti ve výběrovém souboru o rozsahu $n = 1000$, a nyní chceme zkoumat, zda tím může být zamítnuta nulová hypotéza „čistě náhodného rozdělení".

Nulovou hypotézu čistě náhodného rozdělení jsme prve (v ‰) načrtli takto:

Ženich	Nevěsta			
	svobodná	ovdovělá	rozvedená	Celkem
svobodný	747	20	82	849
ovdovělý	33	1	4	38
rozvedený	99	3	11	113
Celkem	879	24	97	1000

Zjištěný výsledek, sestavený ve stejném uspořádání, však vypadal takto:

Ženich	Nevěsta		
	svobodná	ovdovělá	rozvedená
svobodný	800	7	42
ovdovělý	18	10	10
rozvedený	61	7	45

Rozdíly ($B - E$), tedy „pozorované četnosti minus očekávané četnosti", vypadají takto, přičemž hned také vypočítáme mocniny těchto rozdílů:

Ženich	Nevěsta					
	svobodná		ovdovělá		rozvedená	
	$(B-E)$	$(B-E)^2$	$(B-E)$	$(B-E)^2$	$(B-E)$	$(B-E)^2$
svob.	53	2809	—13	169	—40	1600
ovd.	—25	625	9	81	6	36
rozv.	—38	1444	4	16	34	1156

Mocniny rozdílů $(B - E)^2$ se nyní dělí očekávanou četností a dostáváme tuto tabulku:

2809/747 =	3,8	169/20 =	8,5
625/ 33 =	18,9	81/ 1 =	81,0
1444/ 99 =	14,5	16/ 3 =	5,3
	37,2	+	94,8
1600/82 =	19,5		
36/ 4 =	9,0		
1156/11 =	105,1		
+	133,6 = 265,6 celkem		

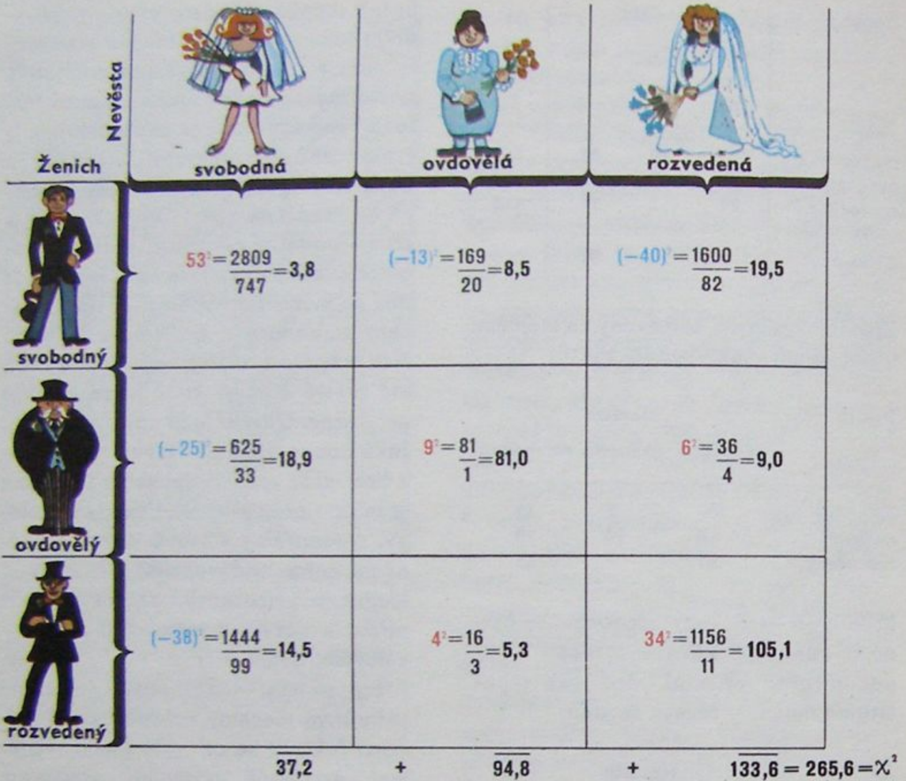
Počet stupňů volnosti jsme již dříve určili jako hodnotu 4. Ať však hledáme kdekoli ve sloupci „4 stupně volnosti", prakticky na každé úrovni je tato nulová hypotéza („vše je pouze náhoda") zcela rozhodně zamítnuta: i když jdeme až do 99,9 % a ponecháme tedy náhodnému rozdělení volný prostor až přes tři směrodatné odchylky, vždy je v tabulce uvedeno nejdříve 18,5.

Bez nejmenší pochybnosti je tím hypotéza zamítnuta a můžeme s určitostí říci: *uzavírání sňatků, jak jsme zjistili, má takové důvody, které nelze vysvětlit ani nejopovážlivější hrou náhody.*

Jaké jsou to důvody? To nelze vyčíst ani z čísel, ani z tabulky chí kvadrát. O tom je možno přemýšlet, radit se se sociology, s odborníky v manželských poradnách nebo oddávajícími úředníky a hledat v příručkách, potom vytvořit novou hypotézu a testovat ji znovu — s novými údaji!

Přece jen nám však již pouhý pohled na jednotlivé mocniny rozdílů dělené četnosti řekl, že se od náhodného rozdělení extrémně odchylují především sňatky mezi dvěma ovdovělými, resp. mezi dvěma rozvedenými. Snad se především zde budou hledat možné důvody.

K samé technice provedení musíme dodat toto upozornění na malou závadu: *když se počítá chí kvadrátem, mělo by být v každém řádku nejméně 5 měření.* To se právě stalo při našich pozorováních, nikoliv však v kontingenční tabulce náhodného rozdělení, kde je v jedné řádce 1 měření, v další řádce 3 a v jedné řádce 4. Nám jde o princip výpočtu a mimoto víme, že bychom mohli plným právem uvést každé číslo desetkrát větší, protože podklady pocházejí z mnohem většího výběrového souboru.



Ještě zřetelněji než porovnání absolutních čísel ukazuje test chí kvadrát nápadné odchylky na obě strany. Umocněné rozdíly mezi očekávanými a pozorovanými hodnotami se dělí očekávanými hodnotami. Jako nejnápadnější se teď ukazuje četnost sňatků mezi rozvedenými: $\chi^2 = 105{,}1$, nejvyšší číslo v tabulce.

Co se však stane, máme-li desetkrát více pozorování (se stejným podílovým rozdělením)? I hypotetické hodnoty pak musí být dosazeny v desetinásobné velikosti — a mocniny rozdílů stokrát tak veliké! U obou svobodných by to znamenalo: očekáváno 7470, pozorováno 8000, rozdíl 530, čtverec rozdílu 280 900, děleno 7470 dává nyní 38 místo dosavadních 3,8! Podobným způsobem se zdesateronásobí všechny dílčí zkušební hodnoty chí kvadrátu, a tak dostaneme zkušební hodnotu 2656 místo dřívějších 265,6.

V našem příkladu bylo ovšem toto menší číslo již více než dostatečně veliké pro zamítnutí nulové hypotézy. V jiných případech může být zamítnutí možné teprve po přiměřeném zvětšení rozsahu výběrového souboru. Kdyby

např. kostkou házející statistik v 600 hodech (místo v původních 60) hodil sedmdesátkrát jedničku, devadesátkrát dvojku atd., vyšly by z toho pro jednotlivé počty ok tyto rozdíly, mocniny rozdílů a hodnoty chí kvadrátu:

1	— 30	900	9
2	— 10	100	1
3	0	0	0
4	— 40	1600	16
5	50	2500	25
6	30	900	9
			60

Nyní je to zcela jasné: *ve sloupci pro 5 stupňů volnosti je těch 60 všude příliš mnoho.* Už to není náhoda, nyní je dokázáno, že s kostkou není něco v pořádku.

Rozsah výběrového souboru je již zahrnut v obecném výpočetním postupu a nemusí se k němu zvlášť přihlížet — *„očekávaná" čísla se zvyšují s rostoucím rozsahem výběrového souboru.* Menší výběrové soubory dávají menší hodnoty pro χ^2, takže při zmenšujícím se rozsahu výběrového souboru je zamítnutí nulové hypotézy stále obtížnější.

Test chí kvadrát se proto často používá k zamítnutí korelace, kterou nejdříve bezděčně předpokládáme, protože je „viditelná pouhým okem", a to tak, že se testem připisuje náhodě a tím nabádá k opatrnosti. Předpokládejme např., že by bylo v nějakém podniku zaznamenáno během jednoho roku v ranní směně pět těžkých pracovních úrazů, v odpolední třináct a v noční směně dvanáct. Není nic snadnějšího než říci, že rozdíl je nápadný — dělníci první směny jsou lépe odpočinutí, nebo je možno uvést jako důvod cokoliv jiného.

Test chí kvadrát však odkáže rozdělení těchto úrazů do možné oblasti náhody. Nulová hypotéza („nehody jsou v každé směně stejně pravděpodobné, při celkovém počtu třiceti by mělo na každou směnu připadnout v průměru deset") se totiž nezamítá. Dostaneme:

$$\chi^2 = \frac{(5-10)^2}{10} + \frac{(13-10)^2}{10} + $$

$$+ \frac{(12-10)^2}{10} = 3{,}8.$$

V řádce pro dva stupně volnosti (3 minus 1) najdeme však u 90 % ještě 4,6, což nestačí k zamítnutí nulové hypotézy. Teprve na příliš nízké úrovni stěží 85 % se zdá tento výsledek nápadný. Jinak by již tomu bylo při pouze třech úrazech v ranní směně a patnácti, případně dvanácti odpoledne a večer. Pak bychom dostali:

$\chi^2 = 4{,}9 + 2{,}5 + 0{,}4 = 7{,}8.$

7,8 je však přesně tabulková hodnota pro 95 %. Zde již sotva budeme mluvit o náhodném výsledku.

Pro test chí kvadrát je tedy charakteristické jeho použití k průzkumu hypotéz z hlediska rozdělení. Odpovídá na otázku: „Jak se hodí zjištěné hodnoty k naší představě o domnělém rozdělení?" Nebo, vyjádřeno ve směru nulové hypotézy: „Je ze zjištěných hodnot prokazatelné, že předpoklad náhodného rozdělení — nebo jiného zcela určitého rozdělení — lze pokládat za vyvrácený?"

Testuje se tedy, *zda a jak dalece „se hodí" hypotetické rozdělení na skutečnost,* a proto anglická literatura označuje chí kvadrát jako *„goodness-of-fit-test"*, postup ke zjištění, jak dobře se hodí hypotéza, pokud jde o rozdělení. A protože se test týká této hypotézy, těchto „očekávaných hodnot", dělí se mocniny rozdílů očekávanými, a nikoliv zjištěnými četnostmi. Bylo by opravdu nesmyslné chtít testovat, zda něco skutečně zjištěného může nastat nebo ne. Skutečnost nemůže být zkoumána z hlediska své pravděpodobnosti, protože je skutečností, a tím jistotou. Nejisté je pouze to, zda ji lze či nelze uvést do souladu s určitým předpokladem.

10.3 Analýza rozptylu

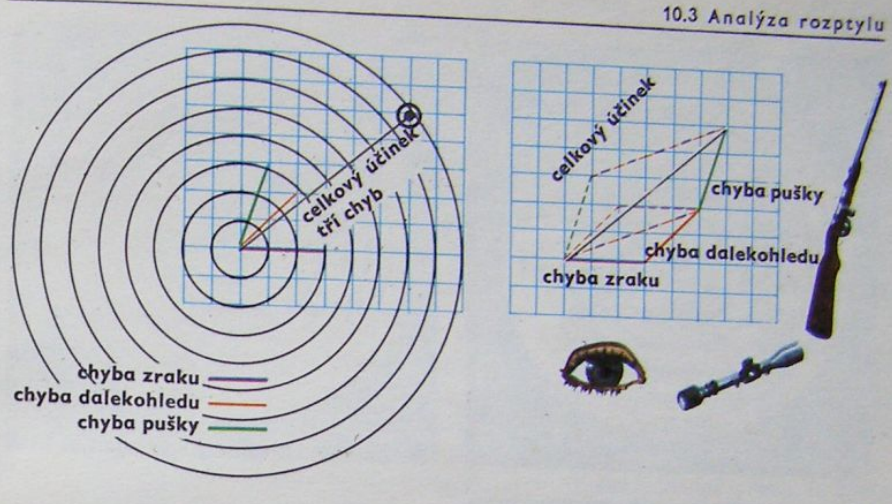
Střelec pálí desetkrát na terč. Později se ukáže, že všechny střely šly daleko mimo střed a skončily v pravé horní části terče. Kdyby se byly rozptýlily přibližně stejnoměrně kolem středu terče, nebylo by to nijak překvapující. Takto však usuzuje pečlivý pozorovatel, že tu není něco v pořádku. Snad není zaměřovací dalekohled zcela bez závad, snad je poněkud nepřesně postavena hlaveň pušky, snad má střelec vadný zrak.

Zaměřovací dalekohled přezkouší optik a zjistí: skutečně, malá nepřesnost znemožňuje přesné zacílení. Puškař pečlivě přezkouší zbraň a zjišťuje: opravdu, hlaveň pušky má maličkou nepřesnost, kterou lze vysvětlit odchýlení střel vpravo. Nakonec jde sám střelec k očnímu lékaři a ten zjišťuje: celkem je zrak v pořádku, trochou nejistoty při odhadu vzdáleností by však bylo případně možno vysvětlit špatné výkony při střelbě.

Konečný výsledek: *máme tři faktory, které asi spolupůsobily, a tak vedly k neočekávanému výsledku na střelnici.* V tomto případě by bylo nepochopitelné, kdyby střelec, jakmile se dověděl výsledek první série, byl systematicky mířil nikoliv do černého, nýbrž jeden bod vlevo dolů od středu. Nám však jde o zásadu následujícího problému: dosáhli jsme takového výsledku pokusu, který se významně odchyluje od střední hodnoty základního souboru, a domníváme se, že jeho příčinou jsou dva nebo více faktorů. Jak je máme od sebe oddělit?

Chceme-li především ušetřit výlohy na optika, puškaře a očního lékaře, doporučuje se např. toto uspořádání pokusu: střelec nejdříve dostane nový zaměřovací dalekohled na starou pušku. Pak se mu dá jiná puška se starým dalekohledem. Za třetí se mu dá nová puška s novým dalekohledem. To máme celkem čtyři výběry. Je však možno jít ještě dál a dát různé kombinace pušek a dalekohledů dvěma jiným střelcům, takže nakonec dostaneme dvanáct výběrů nebo střeleckých souborů.

Když střelec opět netrefí do černého, může být pro to více důvodů. Zrak? Puška? Zaměřovací dalekohled? Rozlišením více vlivů na výsledek testu se zabývá analýza rozptylu.



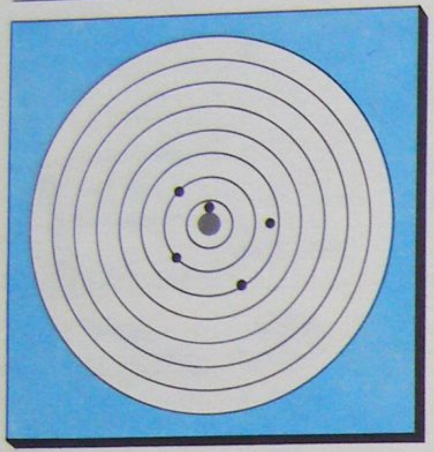
Zásah na terči je příliš vysoko vpravo nahoře. V našem příkladu předpokládáme, že to zavinily tři různé faktory, které se ve svém účinku doplňují (někdy se však také mohou vylučovat): chyba zraku, chyba dalekohledu, chyba pušky. Samozřejmě lze připojit další zdroj chyb — např. nedostatečná praxe ve střelbě.

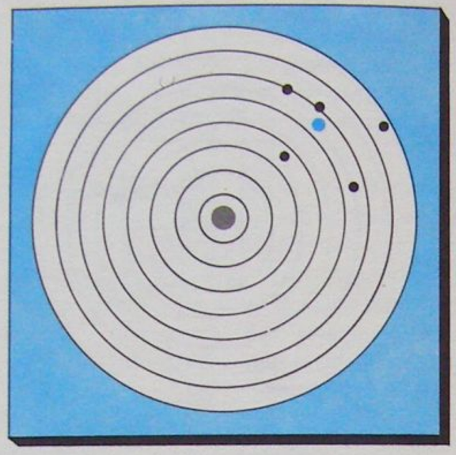
Z porovnání hodnot docílených při všech těchto pokusech (kruhy terče, poloha zásahů) je pak snad možno postupně vytřídit jednotlivé faktory: tak přispívá k odchylce puška, tak zaměřovací dalekohled, tak střelec. Zpočátku nepřehledně promíšený celkový rozptyl se rozkládá, rozptyl se analyzuje — dostali jsme se k *analýze rozptylu (variance)*, k *„rozkladu rozptylu"*.

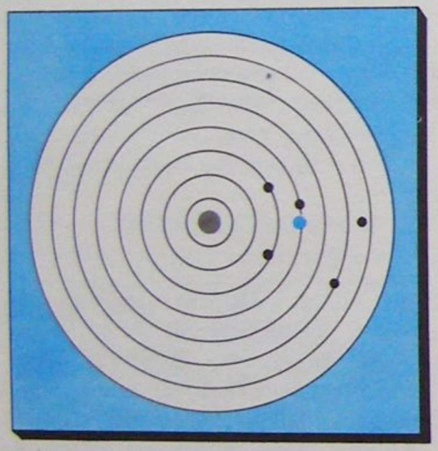
Statistikovi vzniká často tento problém v *obrácené podobě:* pokus byl několikrát obměňován a nyní se má odpovědět na otázku, zda navzájem odlišné výsledky jsou skutečně důsledkem změn v uspořádání pokusů nebo zda jde o náhodu. Možná že výsledky se odlišují pouze proto (což se také stává), že někdo vytáhne z balíčku karet pět zelených a jiný z téhož balíčku pět červených karet.

Z oddílu 5.4 známe již dvě „normalizované chyby", které mohou dát vysvětlení, zda střední hodnota (nebo četnost) nějakého výběru může být uvedena do souladu se srovnatelnými hodnotami jiného výběru, takže lze předpokládat, že oba výběry pocházejí z téhož základního souboru. Není to však pouze střední hodnota, která charakterizuje nějaké rozdělení. I rozptyl může být důležitý pro posouzení situace.

Nejsou-li vzorky příliš velké, může se







Princip rozložení rozptylu zobrazený na terči. Jen chyby střelce vytvářejí při pěti střelách charakteristický vzorec kolem středu. Další chybou (např. chybou pušky) přesune se tento vzorec doprava a ještě další chybou (např. chybou dalekohledu) se přesune nahoru. Výsledek: zásahy se rozptylují o bod doprava nahoru od středu.

střední hodnota z nich vypočítaná velmi podstatně odchylovat od skutečného průměru základního souboru, ale ani rozptyl vzorku nebude stejný jako u základního souboru.

Již v šesté kapitole při výkladu „Studentova rozdělení t" jsme poukázali na to, že rozptyl vypočítaný z malého výběrového souboru je možno použít jako „odhad" pro rozptyl základního souboru jen tehdy, když součet kvadratických odchylek se nedělil počtem měřených hodnot (n), nýbrž pouze počtem stupňů volnosti ($n - 1$). Nebo se součet dělený n násobí korekčním faktorem, tzv. „Besselovým faktorem" (podle astronoma Friedricha Wilhelma Bessela, 19. století): $n/(n - 1)$.

Proč nám najednou tolik záleží na přesné znalosti (nebo pokud možno přesném odhadu) toho, snad jen smyšleného základního souboru? Především

proto, že nechceme porovnávat výsledky dvou nebo více výběrů pouze navzájem, nýbrž pokoušíme se vyjasnit, zda tyto výběry mohou vycházet ze společného základního souboru. Lze dosažené výsledky vysvětlit jako náhodné z jediného souboru, nebo je nutno předpokládat, že pocházejí z různorodých souborů?

Tento hypotetický základní soubor můžeme ovšem zkoumat jen pomocí rozboru výsledků našich výběrů. Situace je tedy tato: máme dva nebo více výběrů a známe jejich střední hodnoty a rozptyly. Jako nulovou hypotézu předpokládáme nejdříve, že různé složení výběrů by bylo možno vysvětlit tím, že při náhodném výběru z hypotetického základního souboru se vyskytly obvyklé, náhodou podmíněné výkyvy.

Tato náhoda má takříkajíc dvě tváře. V každém výběru je náhodný rozptyl kolem příslušné střední hodnoty souboru. *Střední hodnoty* výběrových souborů, které se od sebe liší, dávají možnost dvou různých výkladů: mohly vzniknout náhodným výběrem z jednoho a téhož hypotetického základního souboru (z téhož pytle ořechů vytáhnu jednou osm dobrých ořechů z desíti, pak jen šest z desíti), nebo má rozdíl hlubší význam (výběr osm z desíti pochází z jiného pytle než výběr šest z desíti).

Výchozí situace, která je obvykle podkladem analýzy rozptylu, je tedy taková, že jsem zpravidla načal dvě nebo více pokusných sérií a chci zjistit, zda provedené pokusné obměny skutečně výsledky významně ovlivnily. Proto se někdy označuje rozptyl „*uvnitř*" jednotlivých výběrů jako „*náhodný*", rozptyl „*mezi*" výběry jako „*vysvětlený*" — vysvětlený totiž tím, že uspořádání pokusu se úmyslně obměňuje, je tedy zcela „*vysvětlitelné*", jestliže se výběry navzájem odlišují.

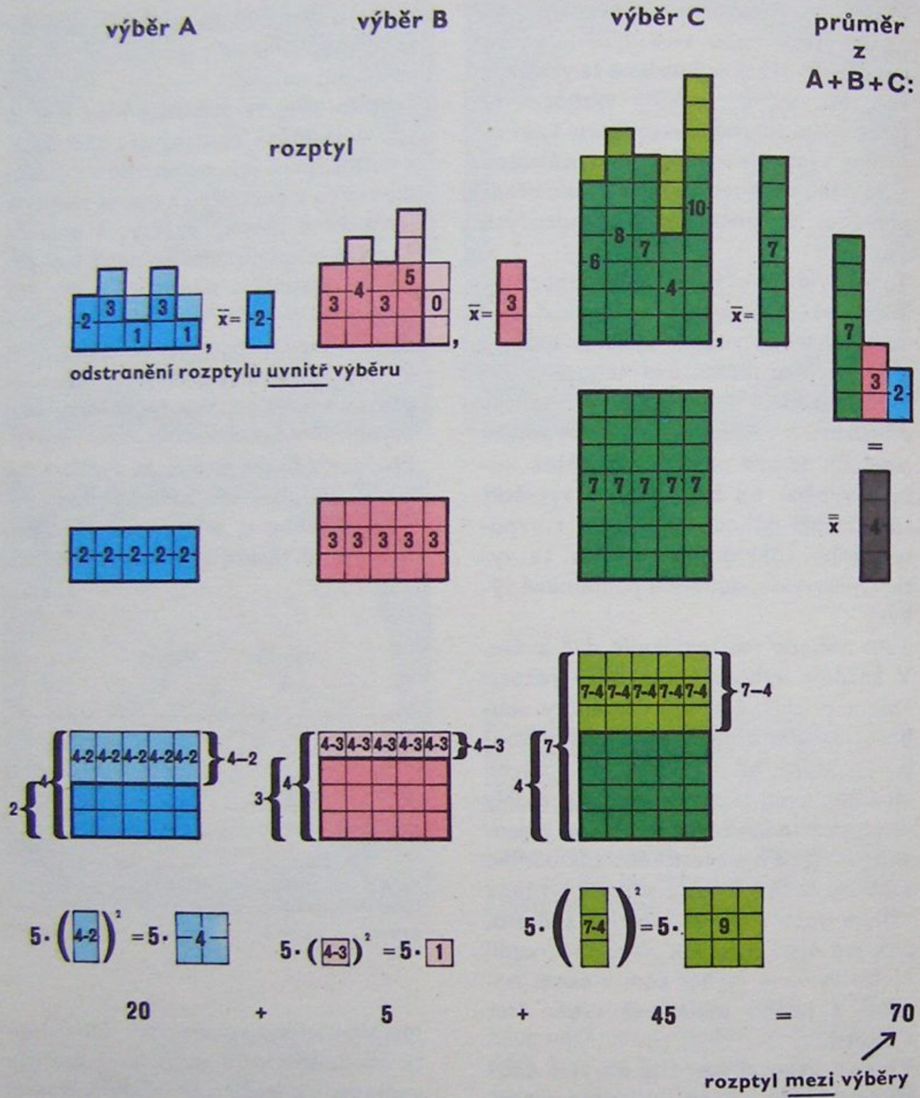
Protože však ve statistice není možno být dostatečně opatrný při zacházení s pojmy, jako je „vysvětlený", zůstaňme raději u neutrálních pojmů rozptylu „uvnitř" a „mezi" výběry. A protože analýza rozptylu mimoto patří k obtížnějším metodám statistiky a my zde nechceme uvádět nekonečné matematické vývody, musíme se omezit na krajně jednoduché příklady, které však přesto vysvětlují zásady analýzy rozptylu.

Předpokládejme proto, že bychom vytvořili tři výběrové soubory vždy stejného rozsahu $n_1 = n_2 = n_3 = 5$ a přitom dostali těchto 15 měřených hodnot:

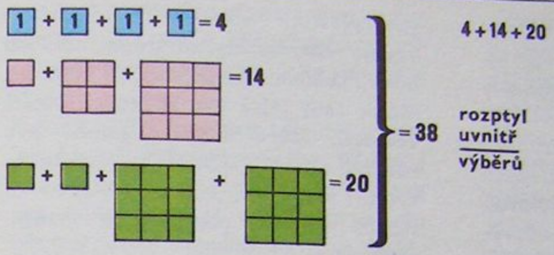
	Vzorek A	Vzorek B	Vzorek C
	2	3	6
	3	4	8
	1	3	7
	3	5	4
	1	0	10
	10	15	35
Proto:	$\bar{x}_1 = 2$	$\bar{x}_2 = 3$	$\bar{x}_3 = 7$

Nejdříve se podíváme na střední hodnotu onoho výběru, o kterém — protože nemáme nic lepšího — se domníváme, že je složen z uvedených tří vzorků. Všech patnáct měřených hodnot dohromady dává střední hodnotu tohoto pseudosouboru, průměr vyplývá z průměrů vzorků $\bar{\bar{x}} = \frac{60}{15} = \frac{2 + 3 + 7}{3} = 4$.

Jak se nyní mají jednotlivé měřené hodnoty k celkovému průměru?



Zkoumání tří výběrů z hlediska jejich rozptylu. Při prvním kroku (nahoře) se zjišťuje, jak se odchylují jednotlivé měřené hodnoty každého výběru od svého vlastního průměru vzorku. Pak (uprostřed) se všechny tři výběry jaksi „vnitřně očistí“: každá měřená hodnota odpovídá nyní průměru svého výběru. Zatím byl vypočítán také (docela vpravo) průměr průměrů výběrů, tedy průměr všech patnácti měřených hodnot (zde: $\bar{\bar{x}} = 4$). Nyní se — třetí krok — vypočítá, o kolik se odchylují vnitřně očištěné měrné hodnoty výběrů od celkového průměru. Nakonec se tyto odchylky umocní; součet těchto odchylek dává rozptyl „mezi“ výběry.



Rozptyl uvnitř výběrů lze téměř pouhým okem vyčíst z první řady předchozího obrázku, jen se nesmí zapomenout umocnit. Výběr A vykazuje čtyřikrát odchylku + nebo —1, takže při umocňování se jen dvakrát mění znaménko. V B a C byly však větší odchylky, které se nyní projevují přiměřeně silněji. Součet těchto mocnin dává rozptyl „uvnitř“ výběrů.

Výběr A	B	C
— 2	— 1	+ 2
— 1	0	+ 4
— 3	— 1	+ 3
— 1	+ 1	0
— 3	— 4	+ 6

K tomu mocniny těchto rozdílů:

A	B	C
4	1	4
1	0	16
9	1	9
1	1	0
9	16	46
24 +	19 +	65 = 108

V této druhé tabulce jsme vlastně neudělali nic jiného, než že jsme vypočítali rozptyl patnácti měření oproti jejich průměru, stejně jako již ve druhé kapitole této knihy. Vypočtený součet povýšený na druhou však nyní nedělíme počtem měření, abychom dostali rozptyl, nýbrž uvažujeme, z čeho se tento celkový rozptyl vlastně skládá.

Především se v něm zrcadlí rozptyl „uvnitř“ vzorků, dále však také rozptyl „mezi“ jednotlivými vzorky (s jejich specifickými středními hodnotami) a průměrem celého souboru. Jestliže chceme oba od sebe pečlivě oddělit, nabízí se k tomu tento postup: rozptyl „mezi“ zůstane nepochybně zachován, pokud zůstanou naše tři vzorky. Nyní budeme však postupovat tak, jako by každý výběr dal pětkrát střední hodnotu daného výběru: řekněme $\bar{x}_1$ se rovná stále 2, ale vzniklo z pěti měření, která všechna dala 2. Dosadíme rovněž měřené hodnoty z výběru *B* rovné 3, 3, 3, 3 a 3 — s $\bar{x}_2$ i nadále 3, a $\bar{x}_3$ rovné 7 rovněž vzniklé z pětkrát 7. (Viz obr. na str. 328.)

Nyní opakujeme dřívější postup: tyto „vnitřně vyrovnané“ fiktivní hodnoty se porovnají s nezměněným celkovým průměrem $\bar{\bar{x}} = 4$. To dává v případě vzorku *A* pětkrát umocněnou odchylku $(—2)^2$ $(= 20)$, pro *B* pětkrát $(—1)^2$ $(= 5)$ a pro *C* pětkrát $(+3)^2$ $(= 45)$. Součet dává $20 + 5 + 45 = 70$. To je tedy onen podíl rozptylu, který připadá na rozdíly „mezi“ vzorky a nepřihlíží k rozdílům uvnitř vzorků.

Jestliže však byl celkový rozptyl 108 a z toho 70 se vysvětlilo rozdíly „mezi“ vzorky, zbývá ještě 38 na podíl rozptylu „uvnitř“. Toto tvrzení lze snadno dokázat. Je nutno jen vypočítat rozptyly „uvnitř“ jednotlivých vzorků — tedy rozptyl pěti měření z *A* proti $\bar{x}_1 = 2$, z *B* proti $\bar{x}_2 = 3$ a z *C* proti $\bar{x}_3 = 7$. To dává u *A*: $1^2 + (—1)^2 + 1^2 + (—1)^2 =$

$= 4$; u B: $1^2 + 2^2 + (-3)^2 = 14$; a konečně u C: $(-1)^2 + 1^2 + (-3)^2 + 3^2 = 20$. A to všechno dohromady dává jako součet odchylek „uvnitř“: $4 + 14 + 20 = 38$. (Viz obr. na str. 329.)
Všechny tyto, při těžších příkladech dost často obtížné výpočty nevyvolávají zvláštní zájem na takovém pseudovýběru a miniaturním výběru o 15 měřeních v našem příkladu, který je navíc slepen ze tří výběrů. Jsou spíše prostředkem k jinému účelu — ke zjištění, *dá-li se vůbec rozumně předpokládat, že existuje jednotná populace, ze které by naše tři (obecně: k) výběry mohly pocházet.*
To je totiž naše nulová hypotéza. Tajně předpokládáme, že jsme našimi pokusnými obměnami vystopovali podstatné rozdíly. Musíme si však být vědomi, že všechno po nulové hypotéze je vlastně jen náhoda. Jako správní vědci zkoumáme tedy nejdříve nulovou hypotézu. Tuto nulovou hypotézu bude pak možno snadno zamítnout, jestliže je rozptyl „mezi“ výběry velký nebo rozptyl „uvnitř“ výběrů malý. Zkrátka: bude třeba zkoumat relativní poměr velikosti obou částí rozptylu.

10.31 Rozdělení F

Tato zkouška se opět provádí s porovnávací tabulkou, a to se „*Snedecorovým rozdělením F*“ (F k poctě sira Ronalda A. Fishera). Dříve než začneme tuto zkoušku provádět, musíme samozřejmě vypočítat zkušební veličinu. Proto musíme znovu mluvit o stupních volnosti.
Pokud jde o soubor našich měření, máme (při $n = 15$ měření) $n - 1 = 14$ stupňů volnosti. Každý jednotlivý výběr se však skládal jen z pěti měření, takže „uvnitř“ každého z těchto výběrů jsou jen čtyři stupně volnosti. Tři (obecně: k) výběry vždy se čtyřmi stupni volnosti dávají celkem dvanáct stupňů volnosti. Máme tedy ještě dva zbývající stupně volnosti, které můžeme použít pro rozptyl „mezi“. (Obecně: jestliže se vytvoří k výběrů, je k dispozici $k - 1$ stupňů volnosti pro rozptyl „mezi“ výběry.)
Jako celkový rozptyl jsme již zjistili 108, z toho připadlo 38 na rozptyl „uvnitř“ a 70 na rozptyl „mezi“. Nyní vypočítáme rozptyl „uvnitř“ tím, že dělíme součet rozptylů počtem na něj připadajících stupňů volnosti:

$$\text{Rozptyl „uvnitř“: } \frac{38}{12} = 3{,}17.$$

Stejně postupujeme s rozptylem „mezi“:

$$\text{Rozptyl „mezi“: } \frac{70}{2} = 35.$$

Již jsme uvažovali, že zkušební veličina by musela zjistit poměr mezi oběma typy rozptylu, a skutečně tomu tak je. Testovací veličinu pro test F vypočítáme podle tohoto schématu:

$$\textit{zkušební veličina testu } F = \frac{\text{rozptyl „mezi“}}{\text{rozptyl „uvnitř“}}.$$

V našem příkladu to dává zkušební veličinu $35/3{,}17 = 11{,}0$, a protože jsme měli velkou varianci „mezi“ a malou „uvnitř“, doufáme, že pomocí relativně velké zkušební veličiny budeme moci zamítnout nulovou hypotézu. Je ovšem třeba uvážit, že rozsah našeho výběru byl velmi malý. Ale podívejme se, co říká tabulka F.
Tabulka rozdělení F se zdá být ještě zmatenější než tabulka rozdělení chí kvadrát — Studentovo t je proti tomu malá násobilka. Rozdělení F má totiž pro každou hladinu významnosti matici řádek a sloupců: v řádcích jsou stupně volnosti menšího rozptylu, ve sloupcích většího.
Vyčerpávající soustava tabulek rozdělení F tvoří tedy celou knihu. Většinou je však možno se spokojit s tabulkami pro několik zvlášť důležitých hladin významnosti, zejména 95 % a 99 %. Platí zde, stejně jako u chí kvadrátu a Studentova t, že se nevyhledává přímo zjištěná testovací veličina, která slouží jen pro účely porovnání s tabulkovou hodnotou. Zkoumáme tedy nejdříve situaci na podkladě tabulky „rozdělení F“ na úrovni 95 %: mohly by naše tři výběry pocházet z téhož základního souboru, jestliže náhodě vyhradíme 95% prostor (= dvě směrodatné odchylky)?
Menší rozptyl (3,17) měl dvanáct stupňů volnosti. Hledáme vlevo řádek 12 a přejdeme dále ke sloupci 2 (neboť náš větší rozptyl měl dva stupně volnosti.) V tomto řádku najdeme hodnotu 3,89. Naše testovací veličina 11,0 je mnohem vyšší, odchylky jsou tedy určitě významné. Nyní zkontrolujeme ještě rozdělení F v tabulce 99 % a nacházíme tam v řádku 12 sloupec 2 hodnotu 6,93. Tedy i na

Dvě krátké tabulky rozdělení F na úrovni 95 a 99 % (druhá na násl. str.). Protože se musí přihlížet vždy ke dvěma stupňům volnosti, jsou řádky i sloupce rozděleny podle stupňů volnosti.

Rozdělení F; 95% úroveň

	1	2	3	4	5	6	8	10	12	20	50	100	∞
1	161	200	216	225	230	234	239	242	244	248	252	253	254
2	18,51	19,00	19,16	19,25	19,30	19,33	19,37	19,39	19,41	19,44	19,47	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,84	8,78	8,74	8,66	8,58	8,56	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,04	5,96	5,91	5,80	5,70	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,82	4,74	4,68	4,56	4,44	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,15	4,06	4,00	3,87	3,75	3,71	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,73	3,63	3,57	3,44	3,32	3,28	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,44	3,34	3,28	3,15	3,03	2,98	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,23	3,13	3,07	2,93	2,80	2,76	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,07	2,97	2,91	2,77	2,64	2,59	2,54
12	4,75	3,89	3,49	3,26	3,11	3,00	2,85	2,76	2,69	2,54	2,40	2,35	2,30
14	4,60	3,74	3,34	3,11	2,96	2,85	2,70	2,60	2,53	2,39	2,24	2,19	2,13
16	4,49	3,63	3,24	3,01	2,85	2,74	2,59	2,49	2,42	2,28	2,13	2,07	2,01
18	4,41	3,55	3,16	2,93	2,77	2,66	2,51	2,41	2,34	2,19	2,04	1,98	1,92
20	4,35	3,49	3,10	2,87	2,71	2,60	2,45	2,35	2,28	2,12	1,96	1,90	1,84
22	4,30	3,44	3,05	2,82	2,66	2,55	2,40	2,30	2,23	2,07	1,91	1,84	1,78
24	4,26	3,40	3,01	2,78	2,62	2,51	2,36	2,26	2,18	2,02	1,86	1,80	1,73
26	4,22	3,37	2,98	2,74	2,59	2,47	2,32	2,22	2,15	1,99	1,82	1,76	1,69
28	4,20	3,34	2,95	2,71	2,56	2,44	2,29	2,19	2,12	1,96	1,78	1,72	1,65
30	4,17	3,32	2,92	2,69	2,53	2,42	2,27	2,16	2,09	1,93	1,76	1,69	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,18	2,07	2,00	1,84	1,66	1,59	1,51
50	4,03	3,18	2,79	2,56	2,40	2,29	2,13	2,02	1,95	1,78	1,60	1,52	1,44
100	3,94	3,09	2,70	2,46	2,30	2,19	2,03	1,92	1,85	1,68	1,48	1,33	1,28
∞	3,84	2,99	2,60	2,37	2,21	2,09	1,94	1,83	1,75	1,57	1,35	1,24	1,00

této hladině významnosti je naše testovací veličina 11,0 rozhodně větší. Můžeme tedy nulovou hypotézu prakticky zamítnout a konstatovat: *naše tři výběry nepocházejí s největší pravděpodobností z téhož základního souboru.*

Ukázali jsme jen základní rysy nejjednodušší analýzy rozptylu na krajně zjednodušeném příkladu, ale s trochou fantazie je možno vytvořit si představu, jaké jsou zde skryty možnosti. Vhodným vytvářením časových řad mohou být zkoumány, tříděny a analyzovány domnělé, žádoucí, obávané účinky i vedlejší důsledky všeho druhu z hlediska jejich pravděpodobnosti.

Na analýze rozptylu je totiž podstatné to, že má umožnit poznání, *jaká část rozdílu mezi výběry se má přisoudit náhodilostem a jaká část se asi musí vykládat jinak,* a poukazuje na více nebo méně významnou jinou druhovost dvou (nebo více) výběrů.

Tato analýza nám teď také poskytuje — pozdě, ale přece — *matematické zdůvodnění „Besselovy korekce“*, případně toho, že se dnes v matematické statistice vypočítává rozptyl výběrového souboru s^2 zpravidla nikoliv dělením součtu kvadratických odchylek číslem n, nýbrž pouze $n - 1$. S tímto problémem jsme se neustále potýkali od prvního popisu odchylky a směrodatné odchylky přes rozdělení t až do tohoto oddílu. Až dosud jsme také nemohli zcela jasně zdůvodnit, že u výběru malého rozsahu (např. $n < 100$) se musí použít korekční faktor, resp. že se o tuto korekci postará již definice s^2 a s.

Rozdělení F; 99% úroveň

	1	2	3	4	5	6	8	10	12	20	50	100	∞
1	4052	4999	5403	5625	5764	5859	5981	6056	6106	6208	6302	6334	6366
2	98,49	99,00	99,17	99,25	99,30	99,33	99,36	99,40	99,42	99,45	99,48	99,49	99,50
3	34,12	30,82	29,46	28,71	28,24	27,91	27,49	27,23	27,05	26,69	26,35	26,23	26,12
4	21,20	18,00	16,69	15,98	15,52	15,21	14,80	14,54	14,37	14,02	13,69	13,57	13,46
5	16,26	13,27	12,06	11,39	10,97	10,67	10,27	10,05	9,89	9,55	9,24	9,13	9,02
6	13,74	10,92	9,78	9,15	8,75	8,47	8,10	7,87	7,72	7,39	7,09	6,99	6,88
7	12,25	9,55	8,45	7,85	7,46	7,19	6,84	6,62	6,47	6,15	5,85	5,75	5,65
8	11,26	8,65	7,59	7,01	6,63	6,37	6,03	5,82	5,67	5,36	5,06	4,96	4,86
9	10,56	8,02	6,99	6,42	6,06	5,80	5,47	5,26	5,11	4,80	4,51	4,41	4,31
10	10,04	7,56	6,55	5,99	5,64	5,39	5,06	4,85	4,71	4,41	4,12	4,01	3,91
12	9,33	6,93	5,95	5,41	5,06	4,82	4,50	4,30	4,16	3,86	3,56	3,46	3,36
14	8,86	6,51	5,56	5,03	4,69	4,46	4,14	3,94	3,80	3,51	3,21	3,11	3,00
16	8,53	6,23	5,29	4,77	4,44	4,20	3,89	3,69	3,55	3,25	2,96	2,86	2,75
18	8,28	6,01	5,09	4,58	4,25	4,01	3,71	3,51	3,37	3,07	2,78	2,68	2,57
20	8,10	5,85	4,94	4,43	4,10	3,87	3,56	3,37	3,23	2,94	2,63	2,53	2,42
22	7,94	5,72	4,82	4,31	3,99	3,76	3,45	3,26	3,12	2,83	2,53	2,42	2,31
24	7,82	5,61	4,72	4,22	3,90	3,67	3,36	3,17	3,03	2,74	2,44	2,33	2,21
26	7,72	5,53	4,64	4,14	3,82	3,59	3,29	3,09	2,96	2,66	2,36	2,25	2,13
28	7,64	5,45	4,57	4,07	3,76	3,53	3,23	3,03	2,90	2,60	2,30	2,18	2,06
30	7,56	5,39	4,51	4,02	3,70	3,47	3,17	2,98	2,84	2,55	2,24	2,13	2,01
40	7,31	5,18	4,31	3,83	3,51	3,29	2,99	2,80	2,66	2,37	2,05	1,94	1,81
50	7,17	5,06	4,20	3,72	3,41	3,18	2,88	2,70	2,56	2,26	1,94	1,82	1,68
100	6,90	4,82	3,98	3,51	3,20	2,99	2,69	2,51	2,36	2,06	1,73	1,59	1,43
∞	6,64	4,60	3,78	3,32	3,02	2,80	2,51	2,32	2,18	1,87	1,52	1,36	1,00

Skutečný (neznámý) rozptyl základního souboru, variance σ^2, sestává zřejmě ze dvou faktorů: z rozptylu měření výběrového souboru kolem průměru výběrového souboru $\bar{x}$ a z rozptylu četných možných $\bar{x}$ kolem skutečného průměru μ. Tímto druhým hlediskem jsme se zabývali již v oddílu 5. 4, kde jsme také popsali náhodnou chybu střední hodnoty výběru: činila $\sigma/\sqrt{n}$.

Protože celkový rozptyl výběrového souboru tvoří souhrn vnitřního výběrového rozptylu a náhodné chyby průměru výběru, musíme od „naivně“ vypočítaného rozptylu — jak by to bylo správné u jednoho souboru a jak jsme jej již ve druhé kapitole poznali a pak často používali — odečíst náhodnou chybu střední hodnoty a zbytek (právě rozptyl „uvnitř“ výběru) použít jako odhad pro rozptyl souboru. Provedeme-li však tuto operaci, odečteme-li tedy od souhrnu kvadratické odchylky dělené n náhodnou chybu střední hodnoty $(\sigma/\sqrt{n})$, dostáváme po několika úpravách *korigovanou hodnotu rozptylu*, rozptyl výběrového souboru očištěný od náhodné chyby střední hodnoty.

Toto „korigované“ s^2, případně s, se zpravidla obecně používá v učebnicích matematické statistiky, protože výběrový soubor se vždy chápe jako směrový ukazatel pro odhad celku, není tedy uzavřeným základním souborem.

Kdyby však skutečný průměr základního souboru (μ) byl znám — což se v praxi stěží stane — pak samozřejmě nedojde k náhodné chybě střední hodnoty a korekce odpadá.

Rozdělení F také umožňuje poznat, zda se rozptyl dvou výběrových souborů od sebe liší, a tím tvoří důležitý doplněk testu t, který má měřit odchylku od středních hodnot. Vzpomeňte si na příklad, který jsme uvedli v oddílu 6. 41: z pěti měření jsme dostali hodnoty 82, 87, 104, 92 a 90 a průměr výběrového souboru $\bar{x} = 91$. Tam nám šlo o zkoumání otázky, zda má být nulová hypotéza $\mu = 100$ zamítnuta či nikoliv.

Nyní předpokládejme, že bychom byli vytvořili druhý výběrový soubor, z kterého by vyšla tato měření: 98, 88, 95, 89 a 100. Průměr výběrového souboru je nyní $\bar{x}_2 = 94$, ale náš zájem se teď ani tak nezaměřuje na tento průměr jako na otázku, zda rozptyl zůstane konstantní, tj. zda se změnil jen uvnitř prostoru volného pro náhodu.

Pro první výběrový soubor vyšlo $s_1^2 = 67$. Druhý výběrový soubor dává:

$$s_2^2 = \frac{16 + 36 + 1 + 25 + 36}{4} = 28,5.$$

Větší rozptyl dělený menším dává $67/28,5 = 2,34$. Vyhledáme porovnávací hodnotu v tabulce F vždy u čtyř stupňů volnosti a najdeme (úroveň 95 %) 6,39. Hypotéza o nezměněném rozptylu není zamítnuta.

Nakonec můžeme ještě zkusit, zda se průměr výběrového souboru změnil pouhou náhodou nebo zda se za tím něco skrývá. To provedeme se Studentovým t a dostaneme:

$$t = \frac{94 - 91 \sqrt{5}}{\sqrt{67 + 28,5}} = \frac{6,7}{9,77} = 0,685.$$

Tabulka „rozdělení t“ udává pro 8 stupňů volnosti (totiž po čtyřech ze dvou výběrů) na úrovni 95 % ještě 1,86, na úrovni 80 % pak ještě 0,89. *Odchylka středních hodnot je zřejmě náhodná.*

Analýza rozptylu se používá často v lékařském výzkumu. Zde ovšem platí,

jako bohužel u všech statistických metod, že všechny vypočítané zákonitosti se zakládají na předpokladu, že se provede náhodná volba ze stejného druhu. Je ovšem dobře možné vypěstovat myší kmeny, u kterých jsou jednotlivá individua téměř totožná co do organické konstituce: vstříkne-li se deseti z nich přípravek A a deseti jiným přípravek B, lze předpokládat, že by výsledek byl stěží jiný, kdyby se obě skupiny zaměnily. (K tomu ještě někdy přistupuje možnost oboustranného pokusu na jednom zvířeti.)
Při léčbě lidí však ještě nejsme v „Krásném Novém Světě" Aldouse Huxleyho, aby bylo k dispozici 64 jednovaječných množenců Bokanovského jako pokusný materiál stejného druhu. V daném případě také nemůže být nemoc vyvolána stejnoměrnými dávkami bakterií nebo škodlivých látek, ale musí se brát, jak přijde — v počátečním i pozdějším stadiu, u staršího i mladšího člověka, u odolného stejně jako u citlivého pacienta. Jde-li o rozšířenou nemoc, je možno nanejvýš vytvořit skupiny a získat pomocí analýzy rozptylu cenné poznatky. Jde-li však o poměrně vzácnou nemoc, bude obtížné vytvořit stejné podmínky nutné pro teoretické úvahy o pravděpodobnosti.
M. J. Moroney, na jehož znamenitou knihu „*Facts from Figures*" (Skutečnosti z čísel) zde chceme zvlášť upozornit, k tomu jednou řekl: „Žádná oblast není pro neopatrného tak nebezpečná a plná »vlčích jam« jako závěry z lékařské statistiky. Je spíše výjimkou než pravidlem, že jsou totožné všechny jiné porovnávané skupiny v každém významném ohledu — mimo ten, který má být ověřován. Každý statistik, který se kdy zabýval lékařskou statistikou, ví, jak často se musel vzdát, protože skupiny nabízené pro porovnání nejsou prostě dostatečně nebo vůbec porovnatelné."
K tomu přistupuje ještě jiný problém, o kterém jsme se již zmínili. Na kom se má ověřovat nový, avšak prozatím jen při pokusech na zvířatech vyzkoušený lék? Je možné, že nebezpečí bude větší, možná však i léčebný efekt lepší, jestliže se použije pro „beznadějné" případy, u kterých každé jiné léčení zklamalo, nebo když se použije pro lehčí případy, které v nejhorším snáze překonají neočekávaný vedlejší účinek? Výsledky pokusu budou mít rozdílné výsledky v závislosti na tomto rozhodnutí. Pro statistika by byl případ jasný: vybrat pokud možno pacienty stejného druhu se stejným průběhem nemoci a pak dělat pokusy, např. na základě schématu latinských čtverců.
Latinské čtverce tvoří schéma pokusného zařízení, jehož se používá zejména při zemědělských experimentech a které umožňuje zcela zřetelně vyřadit náhodné výkyvy. Při pokusech s novými druhy rostlin, novými hnojivy atd. je vždy nebezpečí, že výsledky sklizní nedovolí spolehlivé závěry, protože i nepatrné rozdíly v jakosti půdy ovlivnily výsledky více než jakost nového druhu nebo nového hnojiva. (Viz obr. str. 335.)
Proto se používá takového uspořádání pokusu, které podobné nahodilosti vyřadí tím, že při pokusech např. se třemi druhy se veliké pole vhodně rozdělí na devět čtverců a dané tři druhy se vysadí tak, že v každé řadě (řádku) a v každém „sloupci" je každý ze tří druhů. Jestliže druhy označíme písmeny A, B a C, vypadá schéma latinského čtverce takto:

A	B	C
B	C	A
C	A	B

Příklad latinského čtverce: např. použijí se hnojiva ve čtyřech koncentracích nebo čtyřech druzích (A, B, C, D); tímto uspořádáním se do značné míry vyrovnají případné rozdíly v jakosti půdy atd.

Pro čtyři druhy by muselo být k dispozici $4^2 = 16$ čtverců, pro 6 druhů $6^2 = 36$ čtverců. Pak se porovnají rozdíly „mezi" řádky, sloupci a druhy, a tím se může dalekosáhle rozlišit vliv půdních podmínek od vlivu druhů. Mimoto dává rozdíl z celkového rozptylu a rozptylů „mezi" ještě jako zbytek rozdíly „uvnitř", které vyjadřují náhodné faktory, k nimž se nepřihlíželo nebo které nebyly poznány.
Schéma latinského čtverce není ovšem omezeno jen na zemědělství. Podobným způsobem lze mimo jiné provádět analýzu odchylek v továrnách, například tak, že se nový výrobní postup porovnává s dosavadním na třech různých strojích, které jsou střídavě obsluhovány třemi více nebo méně schopnými dělníky. Na základě tohoto pokusného schématu probíhají často také lékařské, biologické a psychologické experimenty a pomocí analýzy odchylek jsou pak vyhodnoceny.

10.4 Testy pořadí

Jestliže střetnutí boxerů neskončí k. o., vyhlásí ringový rozhodčí jednoho z nich vítězem na body, někdy i za projevů nesouhlasu diváků, kteří pokládají poraženého za lepšího. Když se pořádá soutěž krásy, nevolí porota jen jednu

Schéma uspořádání pokusu podle latinského čtverce: 4 léky A, B, C, D se zkoušejí u čtyř věkových skupin. V důsledku nevyhnutelných rozdílů v tělesném stavu a průběhu nemoci nepostačí ovšem 16 pacientů, spíše je nutno říci podle obrázkových statistik: „každá figurka znamená 10 (nebo ještě lépe více) osob“.

„královnu krásy“, ale zařadí do pořadí také další kandidátky, aniž se nejdříve provedla analýza odchylek v rozměrech boků nebo třetí odmocnina běloskvoucího chrupu. Když jsou vyřazováni mladí důstojníci amerických vojenských akademií, dostane každý absolvent pořadové číslo, a když později zemře jako generál, konstatuje se v nekrologu, že byl kdysi vyřazen jako 87. ze 453 svého ročníku. Zeptáme-li se známého, ve kterém z deseti měst by nejraději žil, udělá — snad po určitém váhání — preferenční seznam: od města svých snů až po to, ve kterém by nechtěl být ani pohřben.

Když... a tento výpočet by mohl pokračovat donekonečna. Jsou zkrátka nesčetné situace, události, předměty, *které sice lze sestavit do určitého pořadí, jejich rozdíly je však možno jen obtížně zjistit s matematickou přesností*, i když vědecké výzkumy vždy směřují k nalezení mnoha měřitelných, zvažitelných a vypočítatelných hodnot. (Dokonce i superkrálovnu krásy lze nechat zvolit počítačem pomocí zaprogramovaných ideálních měr.)

Kromě toho existuje však mnoho případů, ve kterých by sice bylo zásadně možné provést přesně změřitelné rozlišení, kdy však je zařazení do pořadí jednodušší, méně nejisté a pro sledovaný účel plně postačuje.

Při každé sportovní soutěži je jeden vítěz, jeden druhý a jeden třetí, případně mohou být rovným dílem obsazena dvě první nebo dvě druhá místa, ale nikoho nenapadne umocňovat rozdíly mezi časy nebo dálkami anebo je dokonce složitým způsobem uvádět do vztahu k jiným výkonům či rekordům. Na stupni vítězů stojí první, vedle druhý a třetí. Má-li čtvrtý smůlu, přišel jen o vlásek o stupeň a o medaili. Pro všechny je však rozhodující jen jedno: *pořadí.*

V deváté kapitole při pojednání o korelaci jsme se zmínili o příkladu záporné korelace mezi odeslanými dopisy a telegramy a nyní seřadíme tam uvedených 13 zemí podle jejich „pořadí“ v obou

druzích zasílání zpráv a pokusíme se zjistit, zda pomocí takto stanoveného pořadí můžeme vypočítat míru korelace.

Kdybychom chtěli zjistit přesnou korelaci, musely by být nejdříve zjištěny střední hodnoty odeslaných telegramů a dopisů a pro každou jednotlivou dvojici měřených hodnot vypočítána odchylka a její mocnina atd.

Korelace pořadí však nejdříve porovná, na kterém místě (v kterém pořadí) jsou jednotliví nositelé znaků v obou případech. V našem případě to vypadá takto:

	Pořadí podle telegramů	Pořadí podle dopisů	Rozdílnost pořadí (d)	d^2
NDR	1	11	—10	100
Řecko	2	13	—11	121
Španělsko	3	10	— 7	49
Itálie	4	9	— 5	25
Jugoslávie	5	12	— 7	49
Švédsko	6	6	0	0
Rakousko	7	7	0	0
Belgie	8	2	6	36
Francie	9	5	4	16
NSR	10	8	2	4
Švýcarsko	11	1	10	100
Velká Británie	12	4	8	64
Nizozemí	13	3	10	100

$\Sigma\, d^2 = 664$

Upustili jsme od výpočtu přesné korelace podle Pearsonova korelačního koeficientu a jen jsme zdůraznili, že korelace bude bezpochyby velmi vysoká a samozřejmě záporná. Nyní použijeme jiný korelační koeficient, tzv. *Spearmanův koeficient korelace pořadí*, abychom dodatečně dokázali toto tvrzení. Nazýváme jej také ϱ (ró) a uvádíme zde vzorec (aniž bychom se pouštěli do jeho odvození):

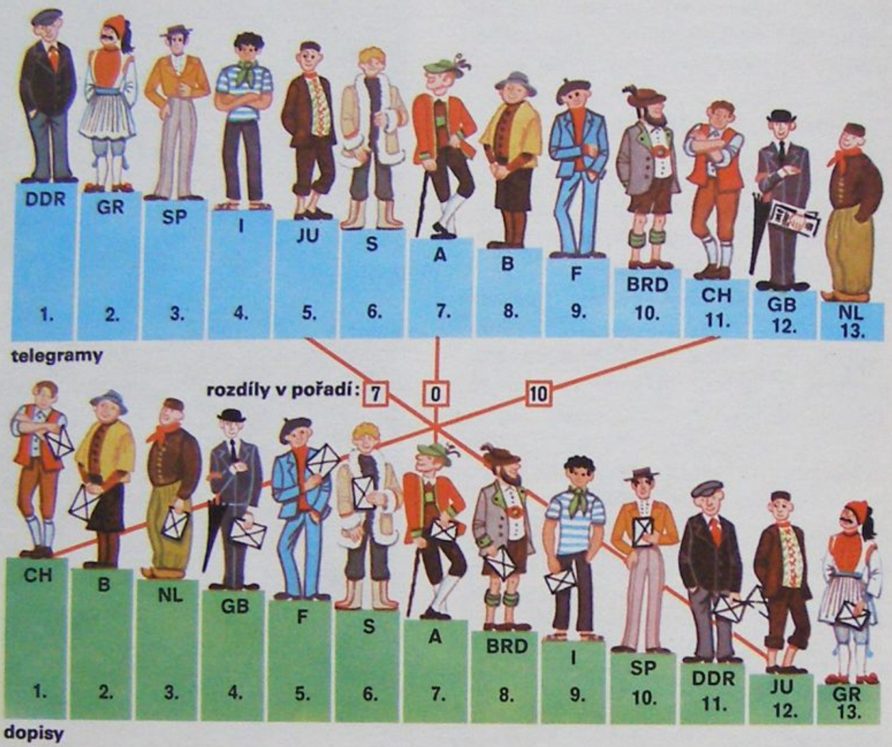
Pořadí krásy: královna krásy 1969 a její největší konkurentky.

$$\boxed{\varrho = 1 - \frac{6\,\Sigma d^2}{n(n^2-1)}}$$

Počet pozorovaných nositelů znaků n se rovná v našem příkladu 13, Σd^2 (součet umocněných rozdílů pořadí) = = 664, a tak dostáváme:

$$\varrho = 1 - \frac{6 \,.\, 664}{13 \,.\, 168} = 0{,}82.$$

To je, jak jsme předpokládali, velmi vysoká záporná korelace, neboť i u ko-



1. NDR 2. Řecko 3. Španělsko 4. Itálie 5. Jugoslávie 6. Švédsko 7. Rakousko 8. Belgie 9. Francie 10. NSR 11. Švýcarsko 12. V. Británie 13. Nizozemí

1. Švýcarsko 2. Belgie 3. Nizozemí 4. V. Británie 5. Francie 6. Švédsko 7. Rakousko 8. NSR 9. Itálie 10. Španělsko 11. NDR 12. Jugoslávie 13. Řecko

Rozdíly v pořadí mezi postavením 13 států v odesílání telegramů, případně dopisů. Státy s největším počtem odesílaných telegramů jsou na konci pořadí sestaveného podle odeslaných dopisů a obráceně: silná záporná korelace.

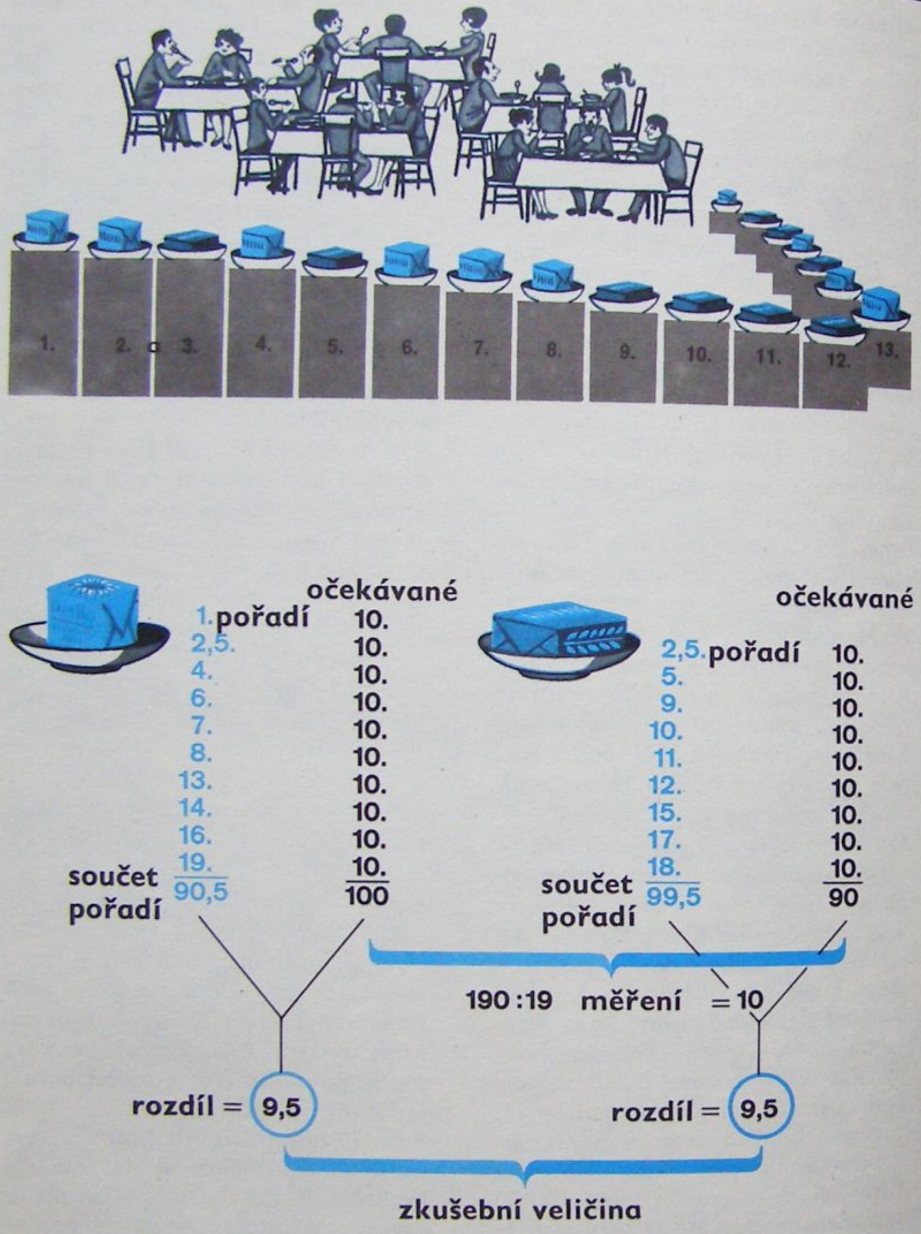
relačního koeficientu pořadí platí, že může být pouze mezi —1 a +1.
Je přirozené, že ani testy pořadí nejsou rychle a snadno proveditelné. Jsou mezi nimi i velmi obtížné a namáhavé postupy, které se ovšem použijí jen tehdy, jsou-li přece jen koneckonců poněkud méně pracné a obtížné než jiné metody (a přesto postačující a spolehlivé), nebo jestliže žádné jiné metody nejsou.
Testy pořadí jsou totiž — a v tom je jejich vlastní význam — zásadně *„nezávislé na tvaru rozdělení“* (*„neparametrické“*).
Máme-li důvod k domněnce, že měření nelze zařadit do normálního, binomického, Poissonova nebo jiného poměrně jednoduchého rozdělení, nebo jsou-li pochybnosti o tom, které rozdělení by bylo nejvhodnější, použije se testu pořadí.
Z „širokého sortimentu nabídky“ testů pořadí předvedeme pouze jediný: *Wilcoxonův test*. Nebo přesněji: jeden z Wilcoxonových testů, tzv. *Wilcoxonův test dvou výběrů*.

10.41 Wilcoxonův test dvou výběrů

Ve velké kuchyni mají být vyzkoušeny dva druhy jedlého tuku, přičemž kritériem jsou jejich chuťové vlastnosti. Polovina porcí jídel se tedy připraví s jedním, druhá polovina s druhým tukem, hosté dostanou dotazník a mají odpovědět, zda jídla pokládají za chutná nebo zda jim vytýkají nějakou závadu. Pokus probíhá po dva týdny a pak se odpovědi vyhodnocují, přičemž pro každý den a každý z obou druhů jídel se vypočítá kvocient spokojenosti (např. poměr hlasů „dobré“ a „špatné“).
Jen na okraj budiž poznamenáno, že takový experiment v sobě skrývá nesčetné zdroje chyb. Je samozřejmé, že v zájmu náhodného rozdělení se nepodávají jídla připravená s tukem A jen v pravé polovině jídelny a v levé polovině s tukem B, nýbrž že se uspořádání každý den mění, protože jinak by stále stejní strávníci dostávali jídla se stejným tukem. Takový postup vyžaduje úzkostlivě přesnou organizaci dotazníků, protože každá záměna (strávníci A dostanou dotazník B) test znehodnotí. Může se také stát, že některý host odmítne jídlo z důvodů, které s použitím tuku nemají nic společného — např. proto, že dostal nepropečenou porci masa atd. Lze však doufat, že podobné rušivé příhody se rozloží přibližně stejnoměrně na oba druhy tuků, a je-li testovaný počet dostatečně velký, nezkreslí výsledek.
Předpokládejme, že jsme zjistili následující „koeficienty spokojenosti“. Pro tuk A za 10 dnů pokusu: 3,5; 4,2; 4,3; 5,0; 2,3; 3,3; 3,6; 5,1; 4,4; 4,9. Pro B za 9 dní pokusu: 3,7; 4,0; 4,1; 4,8; 2,4; 3,0; 3,4; 5,0; 3,8. (V tomto případě by bylo účelné škrtnout test A toho dne, kdy B nebylo použito. V četných jiných případech rozdílný rozsah výběru nepůsobí rušivě.)
Nyní se zjištěné hodnoty z obou vzorků společně uspořádají, přičemž písmeny a nebo b vyznačíme původ zjištěných hodnot. Dostaneme těchto 19 uspořádaných hodnot:

1) 5,1 a	10) 4,0 b
2) 5,0 a	11) 3,8 b
3) 5,0 b	12) 3,7 b
4) 4,9 a	13) 3,6 a
5) 4,8 b	14) 3,5 a
6) 4,4 a	15) 3,4 b
7) 4,3 a	16) 3,3 a
8) 4,2 a	17) 3,0 b
9) 4,1 b	18) 2,4 b
	19) 2,3 a

Nyní se spočítají *„součty pořadí“* obou pokusných řad, přičemž je účelné začít s menším výběrem. Pro výběr B dostaneme celkem 99,5. Protože celkový součet pořadí je 190, je součet pořadí pro A (R_A) = 90,5.
A teď je zase třeba vypočítat zkušební veličinu. Vzpomeňme si na test chí kvadrát, ve kterém se očekávané a pozorované četnosti navzájem porovnávaly, a uvažujme takto: celkový souhrn pořadí z celkem 19 měření je 190; „očekávaná hodnota“ libovolného



Pořadí hodnocení jídel připravovaných s použitím dvou různých druhů tuků a vyhodnocení těchto pořadí pomocí Wilcoxonova testu výběrů.

pořadí by podle toho byla 190/19 = 10. Podívejme se nyní na jeden z obou souhrnů pořadí, např. na souhrn pořadí pro B: $R_B = 99{,}5$ sestává z 9 měření, očekávaná hodnota E_B by tedy byla jen $9 \,.\, 10 = 90$. Nyní odečteme pozorovanou hodnotu od očekávané hodnoty a dostáváme zkušební veličinu: $99{,}5 - 90 = 9{,}5$, přičemž znaménko je bezvýznamné. (Viz obr. na str. 340.)
Stejně jako pro Studentovo t, pro chí kvadrát a pro test F existuje také pro tento Wilcoxonův test srovnávací tabulka, i tentokrát s řádky a sloupci, které udávají rozsah výběrů.
Hledáme v řádce 9 (rozsah B), sloupci 10 (rozsah A) a v tabulce najdeme pro 99% hodnotu 32, pro 95% stále ještě 25.
O zamítnutí nulové hypotézy („náhodný výsledek") nemůže být na této úrovni vůbec řeč. Pokud nedostaneme nový materiál, musíme vycházet z předpokladu, že oba druhy tuků přijímají strávníci přibližně stejně.
Testy pořadí jsou vždy trochu ne-

Wilcoxonův test 99%

4	5	6	7	8	9	10	11	12	13	14	n_2 / n_1
—	—	—	—	—	13,5	15,0	16,5	17,0	18,5	12,0	3
—	—	12,0	14,0	15,0	17,0	18,0	20,0	21,0	22,0	24,0	4
—	12,5	14,0	15,5	18,0	19,5	21,0	22,5	24,0	25,5	28,0	5
15	61,5	16,0	18,0	20,0	22,0	24,0	26,0	27,0	29,0	31,0	6
14	59,0	62,0	20,5	22,0	24,5	26,0	28,5	30,0	32,5	34,0	7
13	55,5	58,0	61,5	25,0	27,0	29,0	31,0	33,0	35,0	38,0	8
12	53,0	55,0	58,0	61,0	29,5	32,0	33,5	36,0	38,5	41,0	9
11	49,5	52,0	54,5	57,0	59,5	34,0	36,0	39,0	41,0	44,0	10
10	46,0	49,0	51,0	53,0	56,0	58,0	39,5	42,0	44,5	47,0	11
9	42,5	45,0	47,5	50,0	52,5	54,0	56,5	44,0	47,0	50,0	12
8	40,0	42,0	44,0	46,0	48,0	50,0	52,0	54,0	50,5	53,0	13
7	36,5	38,0	40,5	42,0	44,5	46,0	48,5	50,0	51,5	56,0	14
6	33,0	35,0	36,0	38,0	40,0	42,0	44,0	45,0	47,0	49,0	51,0
5	29,5	31,0	32,5	34,0	35,5	37,0	38,5	41,0	42,5	44,0	45,5
4	25,0	27,0	28,0	30,0	31,0	32,0	34,0	35,0	37,0	38,0	40,0
3	20,5	22,0	23,5	25,0	25,5	27,0	28,5	29,0	30,5	32,0	32,5
2	—	—	—	—	19,0	20,0	21,0	22,0	23,0	24,0	25,0
n_1 / n_2	15	16	17	18	19	20	21	22	23	24	25

Zkušební tabulka pro Wilcoxonův test dvou výběrů. V políčku, které tvoří řádek a sloupec obou rozsahů výběrů, lze vyčíst porovnávací hodnotu. Pro každou úroveň pravděpodobnosti jistoty je třeba zvláštní tabulky.

přesnější než metody analýzy rozptylu, ale většinou tyto poněkud jednodušší postupy postačí pro rychlé zjištění podstatného. V našem příkladu by se ani analýzou rozptylu nedosáhlo jiného výsledku než poznání, že nulová hypotéza nemůže být zamítnuta; k tomu bychom ale potřebovali mnohem více času a práce. Především však: *testy pořadí jsou nezávislé na rozdělení, a proto vždy použitelné.*

Tím jsme dospěli téměř ke konci. Snad by bylo možné pojednat ještě o mnohém, avšak pak bychom museli ponenáhlu vystoupit do větších výšek matematické statistiky, což již nemůže být naším úkolem. Kdo se chce do těchto výšin odvážit, najde v dodatku stručné odkazy na další učebnice statistiky.

Pouze o jedné věci ještě pojednáme, protože tím se jaksi uzavírá kruh a ještě jednou jasně ukazuje úzké spojení mezi počtem pravděpodobnosti a statistikou, které jsme již na začátku knihy tolik zdůrazňovali. Tentokrát pojednáme o pravděpodobnosti zvláštního druhu — ne o „probability" (pravděpodobnosti), nýbrž o „likelihood" (věrohodnosti).

10.5 Maximální věrohodnost

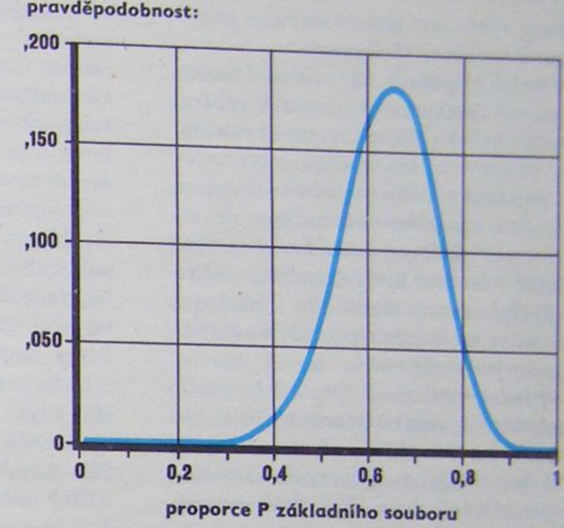
„Ta spousta krásných slov, která Fisher a jeho stoupenci uvádějí k odůvodnění teorie »věrohodnosti«, je mi nepochopitelná." Jestliže Richard von Mises, jeden z vedoucích teoretiků pravděpodobnosti našeho století, takto odmítá teorii věrohodnosti, resp. pokládá ji za zbytečnou komplikaci obvyklého pojmu pravděpodobnosti („od okamžiku, kdy se vymyslelo pojmenování, nakládá Fisher s »věrohodností« jako s úplným symbolem toho, co se v obyčejné řeči nazývá pravděpodobnost"), nemuseli bychom se vlastně vůbec zabývat důvtipnými úvahami o případné existujících *rozdílech mezi pravděpodobností a věrohodností*. Když však přesto poslední oddíl této knihy věnujeme právě „maximu věrohodnosti", má to svůj zvláštní důvod: v „maximu věrohodnosti" se jaksi soustřeďují mnohé myšlenky, s nimiž jsme se opětovně seznamovali v průběhu našeho pojednání.

To, pro co R. A. Fisher, tento dominující zjev statistiky první poloviny našeho století, poprvé v roce 1912 použil pojmu „*maximum likelihood*", se někdy překládá do češtiny jako „*maximální věrohodnost*", častěji se však ponechává anglické označení nebo se zkracuje na *metodu m. l.* Matematická teorie, která je podkladem metody m. l., přesahuje zaměření naší knihy (a dokonce i mnohých odborných statistických knih), ale základní myšlenku lze snadno popsat:

Často jsme viděli, že cílem analýzy výběrových souborů je vyhodnotit poznatky získané z výběrů se zřetelem na neznámý základní soubor. Jak jsme poznali, počet pravděpodobnosti udává především to, jak pravděpodobný je specifický výsledek výběrového souboru ze známého základního souboru. Naproti tomu statistická analýza se snaží usuzovat ze známého výběrového souboru na neznámý základní soubor. Výběrový soubor nám poskytuje přibližně odhadované hodnoty pro základní soubor.

Tuto situaci jsme v oddílu 5.3 zvlášť zdůrazňovali a vylíčili obrácený proces; například z normální karetní hry se určí pravděpodobnost „2 piky, 3 jiné" při pěti libovolně tažených kartách, v jiném případě, např. z výběrového souboru „2 piky, 3 jiné", se usuzuje na rozdělení

Maximální věrohodnost, ať maximálně sebenižší, je nejvyšší vždy tam, kde proporce výběru odpovídá proporci základního souboru. V našem příkladu Wallise a Robertse, v němž mezi dvaceti koulemi výběru bylo 13 červených, je tedy maximální pravděpodobnost nad předpokladem 0,65 — např. 650 červených mezi 1000 koulí.



piků a „jiných" v karetní hře neznámého složení. Na styku obou těchto operací stojí tedy Fisherovo „maximum likelihood" a je matematickou základnou pro všechny „odhadové funkce". Otázka, kterou Fisher nadhazuje, je totiž koneckonců tato: *Jak pravděpodobné je, že vyjde právě tento výsledek výběrového souboru?* Jak „likely", jak „*hodnověrné*" je, že dostaneme právě tento výběr, když lze vycházet z tolika četných hypotetických základních souborů.

Chceme-li tento myšlenkový postup ozřejmit pomocí nějakého příkladu, maximální věrohodnost vzniká z této úvahy:

Z balíčku dvaceti karet, jejichž složení neznám, jsem vytáhl čtyři karty, a to tři červené a jednu černou. Je složení základního souboru (dvaceti karet) nyní takové, že z něho pochází *s relativně největší pravděpodobností, s maximální věrohodností*, tento výběr?

Jisté je pouze jedno: mezi dvaceti kartami jsou nejméně tři červené a jedna černá. Má-li být z jakýchkoli důvodů vyloučena možnost, že jsou tam mimo červené a černé ještě také karty jiných barev, pak je 17 „možných" základních souborů: počínaje „1 černou, 19 červenými" až po „17 černých, 3 červené". Jak je tedy pravděpodobné, že například z prvního hypotetického balíčku („1 černé, 19 červených") dostaneme výsledek výběru „1 černá, 3 červené"? Zřejmě menší než např. z balíčku „4 černé, 16 červených", ale nepochybně větší než z balíčku „17 černých, 3 červené". Bylo by možno provést všechny výpočty pro všechny myslitelné základní soubory, ale nakonec bychom dostali zcela všední výsledek: maximální věrohodnost má ten balíček, který přesně odpovídá „předběžnému výpočtu" výběru: mezi čtyřmi taženými kartami jsme měli jednu černou a tři červené, přičemž základní soubor má dvacet karet. Balíček, který s „maximální pravděpodobností" poskytne takový výběr, má složení stejného poměru: 5 černých, 15 červených.

V jiných případech, kdy základní soubor netvoří celek násobku rozsahu výběru, není snadné výsledek vysypat z rukávu, a nejde-li o binomickou alternativu z poměrně malého základního souboru, nýbrž o komplikované rozdělení ve velkém základním souboru, změní se obyčejné násobení brzy v opravdu složité výpočty — tak složité, že si většinou musíme vypomáhat jednoduššími „odhadovými funkcemi".

Výsledek výběru: „1 černá, 3 červené" je ostatně nejpravděpodobnějším výsledkem základního souboru „4 černé, 16 červených" a „6 černých, 14 červených". Kdybychom však chtěli provést výpočet, ukázalo by se, že pravděpodobnost pro „5 černých, 15 červených" je ještě o něco vyšší než pro obě sousední možnosti složení. U velmi velkých základních souborů může být tato maximální pravděpodobnost zcela malá, ale musí představovat relativně největší pravděpodobnost. Je značně nepravděpodobné, že by se při stu hodech mincí dosáhlo přesně padesátkrát orla, ale „maximum likelihood" by přesto muselo upozornit na tento výsledek, protože je minimálně pravděpodobnější než každá jiná kombinace z 99 možných.

Tento myšlenkový postup nevyvinul teprve R. A. Fischer. V podobné formě jej zpracoval v roce 1846 již Quételet, aby doložil svou teorii o „homme moyen". Na přelomu století shrnul Lexis úvahy o Quételetovi takto: „Jestliže na vývoj tělesné výšky dospělého člověka působí např. vždy 1000 elementárních poruch, dosáhlo by se typické normální výšky jen tehdy, kdyby se setkalo 500 kladných a 500 záporných elementárních poruch. Tento případ je relativně nejpravděpodobnější, ale absolutně je jeho pravděpodobnost pro nesmírně veliký počet všech vůbec možných kombinací přece jen velmi malá."

A již Quételet sestavil tabulku pravděpodobnosti výsledků výběru 999 tahů bílých a černých kuliček z téměř nekonečně velikého osudí, kde budou obě barvy stejně zastoupeny: podle této tabulky je např. pravděpodobnost tahu 499 bílých a 500 černých (nebo také obráceně) 0,0252, pro 497 bílých a 502 černých (nebo obráceně) ještě 0,0249 atd.

A opět se už úplně na závěr setkáváme se vzájemným sepětím *normálního rozdělení a teorie výběru, s maximální věrohodností a zákonem velkých čísel, popisné a analytické statistiky*. Tím se také uzavírá kruh, který vedl od teoretických pravděpodobnostních základů statistiky k odhadovým postupům statistické analýzy a který nás zase přivedl pod krycím jménem „maximální věrohodnost" zpět k teorii pravděpodobnosti.

Závěrem můžeme tedy ještě jednou připomenout to, co jsme se snažili neustále zdůrazňovat: *Statistika pomáhá zmenšovat nejistotu a zacházet realisticky s odhady — může odlišit nepravděpodobné od pravděpodobného a kvantifikovat přesnost rozlišení*. Avšak nemůže vydupat ze země jistoty. Je cennou, dokonce nepostradatelnou *pomůckou* při rozhodování, ale neposkytuje hotové recepty.

Doslov k českému vydání

Pro čtenáře, který celou Swobodovu knihu dočetl až sem, není třeba mnoho dodávat. Redakce českého vydání tedy připojuje závěrem jen několik poznámek.

Když se Nakladatelství Svoboda rozhodovalo vydat tuto knihu, vycházelo především ze skutečnosti, jak velký význam získává v současné době statistika jako zdroj informací a předpoklad rozhodování. V době vědeckotechnické revoluce prudce roste masa nejrůznějších informací, kterou musí každý nějakým způsobem absorbovat a „strávit". Má proto pro něj nesmírný význam kvantifikace těchto informací, jejich správné zpracování a vyhodnocení. A právě to umožňují statistické metody.

Veliký význam má statistika především pro politické a veřejné pracovníky, pro řídící hospodářské kádry, zejména v naší socialistické společnosti, která nese právem označení společnost „vědecky řízená".

Většina veřejných a hospodářských pracovníků, a to i mnoho z těch, kdo absolvovali vysokou školu už před lety, neměla často možnost hlouběji se s metodami a způsobem práce statistiky seznámit. Tato kniha, jak se už každý čtenář přesvědčil, je vynikající zejména tím, jak zpřístupňuje i tomu, kdo nemá velké matematické znalosti, poměrně složité otázky matematické statistiky. Jde o vynikající popularizační práci, jakých není ve světové literatuře mnoho. Tento rys Swobodovy knihy rozhodl o jejím vydání v češtině.

Popularizace u Swobody neznamená zlehčování. Naopak, autor na mnoha místech přímo nebo nepřímo před diletantským přístupem ke statistice varuje. Ať již je to příklad ze „Zbohatlíkova", k němuž se několikrát vrací, nebo sedmá kapitola, výstižně nazvaná „Proč statistiky lžou", kde je čtenář nejen seznamován s úskalím statistických metod, ale i názorně varován před zneužíváním statistiky a jejích metod a před povrchními interpretacemi. Autor zároveň vlastně mimoděk kritizuje tyto četné praktiky zneužívání statistiky v kapitalistických zemích. Je to dobré pro zasmání, ale i pro zamyšlení...

Místy, kde kniha kriticky hodnotí různé nesprávné statistické postupy nebo jejich interpretace, se vytvořila určitá návaznost na leninskou kritiku buržoazních statistik, jak se s ní čtenář může seznámit v práci B. M. Volina „Statistika a politika" (Orbis, Praha 1951) nebo přímo v řadě Leninových prací o rozvoji kapitalismu a imperialismu. Tím nechceme tvrdit, že autor této knihy o statistice Lenina studoval, naopak, alespoň u jedné otázky statistických metod by bylo možné na základě Leninových prací tuto knížku doplnit, a to při výkladu o třídění.

Oddíl o třídění (2.6) je vyložen podle našeho názoru poněkud zjednodušeně proto, že opomíjí sociálně ekonomická kritéria, jimž se zejména statistika obyvatelstva a hospodářská statistika nemohou vyhnout. Třídění opírající se o náhodně zvolené rovnoměrné intervaly rozdělení četností je použitelné při statistickém rozboru přírodních nebo technických jevů. Statistika společenských jevů musí při třídění vycházet jako z metodologického základu i z teorie společnosti. Jen tak můžeme hlouběji poznat i sociálně ekonomické složení obyvatel Zbohatlíkova a vyhnout se nic neříkajícím průměrům.

Kniha „Moderní statistika" není zaměřena na ekonomickou statistiku, i když se jí na některých místech dotýká. Je pravda, že zatím je ekonomická statistika povýtce to, co autor nazývá „popisnou statistikou", i když i v této oblasti se stále více využívá výběrových šetření, mikrocensů, regresní a korelační analýzy a dalších metod, s nimiž jsme se u Swobody seznámili. Takové oblasti ekonomiky, jako např. pojišťovnictví, svého času podstatně přispěly ke zpracování různých oblastí počtu pravděpodobnosti a matematické statistiky a daly vzniknout tzv. aktuárské vědě. V poslední době se tyto problémy dále zpracovávají v matematické teorii rizik a jejich optimálního rozložení a v dalších metodách tzv. operačního výzkumu, což v budoucnu postaví nové úkoly i před moderní statistiku. Nelze ale požadovat, aby v jedné knize bylo všechno.

Kniha, jejíž poslední stránky jste právě dočetli, se opírá především o německou a anglosaskou literaturu. Čtenář jistě nebude mít dojem, že v jiných zemích v oblasti statistiky nic nového nevzniklo nebo nevzniká. Je všeobecně uznáván velký přínos ruských a sovětských matematiků a statistiků pro vytvoření moderní statistiky. I našim čtenářům jsou známa jména Čebyšev (zakladatel petrohradské školy počtu pravděpodobnosti), Markov (Markovovy řetězce), Kolmogorov, Bojarskij i jména dalších ruských učenců, jméno Poláka Oskara Langeho, Maďara Renyiho aj. Z českých autorů jsme do textu zařadili jméno prof. Jaroslava Janko.

Autor knihy vyslovil souhlas s tím, aby v českém vydání byly některé příklady doplněny či pozměněny tak, aby byly bližší našemu čtenáři. V několika případech jsme tohoto svolení využili. Zejména však byl nově zpracován bibliografický dodatek a tento doslov. Větším zásahům do textu jsme se vyhnuli: vždyť jde o překlad a ne o novou knihu. V řadě případů si český čtenář jistě poradí sám.

Těm, koho kniha zaujala, přejeme hodně úspěchů při využívání získaných poznatků a při dalším studiu.

Redakce

Pro další informaci

Na závěr své knihy radí autor čtenářům, kteří se zajímají o další poznatky ze statistiky a nechtějí se zrovna dát zapsat jako mimořádní posluchači na vysokou školu, co by si podle svého zájmu a úrovně znalostí matematiky měli vybrat ke studiu. V českém překladu pochopitelně nebudeme čtenáře orientovat na německé knihy, ale snažíme se analogicky podat několik informací o knihách, které vyšly v češtině a jsou tedy alespoň v knihovnách dostupné českému čtenáři.

Takováto rada není nejlehčí nejen proto, že v českém originále je přirozeně menší výběr, ale i proto, že náš knižní trh je zásobován především učebnicemi statistiky, které se používají k výuce na odborných a vysokých školách.

Rada je obtížná i proto, že metody moderní statistiky jsou obsaženy především v náročných vysokoškolských učebnicích a vědeckých monografiích, jejichž studium je pochopitelně matematicky daleko náročnější, než byla tato kniha.

V dalším studiu je účelné věnovat pozornost především počtu pravděpodobnosti. První informaci je možné získat ze středoškolské učebnice *E. Kraemera - J. Hájka* ad. *„Matematika pro III. ročník SVVŠ"* — část V (SPN, Praha 1967).

Stručný výklad počtu pravděpodobnosti bez použití vyšší matematiky je obsažen v knize *V. Dupače - J. Hájka „Pravděpodobnost ve vědě a technice"* (NČSAV, Praha 1962).

Podrobnější výklad, který se však v některých kapitolách opírá o vyšší matematiku, podává český překlad knihy *V. I. Glivenka „Theorie pravděpodobnosti"* (Přírodovědecké nakladatelství, Praha 1950).

Nejnovější vysokoškolskou učebnici, schválenou pro vysoké školy ekonomické, je kniha *J. Hátla - J. Likeše „Základy počtu pravděpodobnosti a matematické statistiky"* (SNTL - Alfa, Praha - Bratislava 1972), která zároveň tvoří přechod ke studiu matematické statistiky.

Základní informaci o metodách matematické statistiky bez použití vyšší matematiky napsal *R. Reisenauer*. Jsou to *„Metody matematické statistiky a jejich aplikace v technice"* (Polytechnická knižnice, Práce - SNTL, II. vyd., Praha 1970).

Vysvětlení základních pojmů matematické i ekonomické statistiky obsahuje též *„Stručný statistický slovník pro hospodářské pracovníky"*, zpracovaný kolektivem autorů za redakce V. Roubíčka (Svoboda, Praha 1967).

Vysokoškolskou učebnici statistických metod napsali *L. Cyhelský - I. Novák: „Statistika I"* (SNTL, Praha 1967) a *V. Čermák: „Statistika II"* (SNTL - Alfa, Praha-Bratislava 1968). Druhý díl je věnován teorii výběrových šetření.

Některé z uvedených publikací obsahují rovněž výběr z nejdůležitějších statistických tabulek (např. Reisenauerova). Uceleně byly statistické tabulky zpracovány v dnes již obtížně přístupné publikaci *Jaroslava Janko „Sta-*

tistické tabulky" (NČSAV, Praha 1958). Použitím statistiky v různých odvětvích se zabývá řada speciálních knih.

Ekonomickou statistiku jako soustavu ekonomických ukazatelů a agregátů zpracovali nejpodrobněji, a to z národohospodářského hlediska, autoři *J. Kašpar, J. Jílek, V. Matějka* a *V. Roubíček* v knize „*Ekonomická statistika*" (SNTL - Alfa, Praha-Bratislava 1969).

Statistice průmyslu je věnována kniha *F. Egermayera - L. Šauera* „*Základy průmyslové a stavební statistiky*" (SNTL, Praha 1961).

Jiné publikace se zabývají především některými statistickými metodami. Tak grafickým metodám je věnován překlad sovětské knihy *L. A. Byzova* „*Grafické metody ve statistice, plánování a početnictví*" (Orbis, Praha 1955).

Regresní a korelační analýza je předmětem knihy *F. Egermayera - J. Nováka* „*Regresní a korelační analýza pro ekonomy*" (SNTL, Praha 1964).

Teorií výběru se kromě již zmíněné učebnice V. Čermáka zabývá *J. Hájek* v knize „*Teorie pravděpodobnostního výběru s aplikacemi na výběrová šetření*" (NČSAV, Praha 1960).

Hodnocení experimentů je věnována práce *E. Mitteneckera* „*Plánování a statistické hodnocení experimentů. Úvod pro psychology, biology a lékaře*" (SPN, Praha 1968) a kniha *Z. Rotha, M. Jozífko, V. Malého* a *V. Trčky* „*Statistické metody v experimentální medicíně*" (SZN, Praha 1962).

Pro toho, kdo se chce naučit statistické metody prakticky používat, je neocenitelnou pomůckou práce *L. Cyhelského* „*Statistika v příkladech*" (SNTL, Praha 1967).

Podrobnou bibliografii statistické literatury zpracoval *J. Podzimek* ve svém „*Průvodci statistickou literaturou*" (vyd. Výzkumný ústav statistiky a účetnictví, Praha 1970).

Údaje československé statistiky jsou publikovány především v každoročně vydávané *Statistické ročence ČSSR* (vydává SNTL - Alfa) a v publikaci *Čísla pro každého* (vydává též SNTL).

Rejstřík

moderní statistika HELMUT SWOBODA

Z německého originálu Knaurs Buch der modernen Statistik, vydaného nakladatelstvím Droemer Knaur, München-Zürich 1971, přeložil doc. Ing. J. Císař, CSc. Obálku a vazbu navrhl Jan Jánský. Vydání I. Praha 1977. Vydalo Nakladatelství Svoboda jako svou 3971. publikaci. Odpovědná redaktorka Marie Veszeláková. Technický redaktor Petr Havlík.
Vytisklo Rudé právo, tiskařské závody, Praha. Náklad 15 000. AA 28,87, VA 33,16. Tematická skupina 01/1. Cena brož. výt. 43,30 Kčs, váz. výt. 50,— Kčs. 73/505-21-8.5

ČLENSKÁ KNIŽNICE 25 — 004 — 77 Kčs 50,—