MA 0008 – Teorie pravděpodobnosti

PEDAGOGICKÁ FAKULTA MASARYKOVY UNIVERZITY

Břetislav Fajmon, Karel Lepka

Verze textu: leden 2022

Obsah

Úvod

Rádi bychom v tomto textu poskytli materiál, který podepře výuku předmětu MA0008 na Pedagogické fakultě Masarykovy Univerzity v Brně. Text bude psán v průběhu roku 2021 a ověřen ve výuce na jaře 2022.

Představit pravděpodobnost i statistiku v jednom semestru není jednoduché – přesto je kurs někdy zaměřen kromě jistých metod, které mohou pomoci na ZŠ při prezentaci nebo zpracování statistických údajů v Excelu (tomu bude též věnována pozornost), také na výpočty a některá odvození vysokoškolského rázu. Studenti mají částečné povědomí o statistických metodách z kurzů věnovaných tomu, jak psát bakalářskou práci a jak dělat pedagogický výzkum – tento kurs MA 0008 by měl nejen navázat na takové kursy, ale navíc poskytnout jistou matematickou důkladnost v některých věcech, a odvození a vysvětlení základních faktů, je-li to možné zejména s využitím některých znalostí zejména z matematické analýzy (pravděpodobnost je totiž mezioborem matematiky, protože se snaží popisovat náhodnost jak u diskrétních, tak u spojitých veličin, tj. využívá poznatky „diskrétních“ i „spojitých“ matematických disciplín).

Při psaní textu budou využity zkušenosti získané z učebního textu (Fajmon, Hlavičková, Novák 2014) společně s mými bývalými kolegy na FEKT VUT, kde byl hojně využit i text (Loftus, Loftusová 1988). Důležitým textem na téma pravděpodobnosti z Brněnské matematické komunity je (Budíková, Králová, Maroš 2009). Na pedagogické fakultě se ovšem musíme také držet patřičně „při zemi“ a myslet na středoškolskou úroveň, popřípadě úroveň ZŠ – materiálem reprezentujícím tuto oblast je učebnice (Robová, Hála, Calda 2013). A v dokončené verzi textu bude snad něco využito i z anglických učebnic „pro hlupáky“ (for dummies), kterým snad předně jde o srozumitelnost, jako jsou (Rumsey 2006), (Rumsey 2016), (Schmuller 2016).

Obsah předmětu je rozdělen na tří části:

- Cvičení 01-02: Popisná statistika – zpracovává informace na základě naměřených dat pomocí tabulek, grafů, funkcí, číselných hodnot. V podstatě vyučováno na ZŠ. Výuka bude doplněna úkoly zpracovanými v programu Excel.
- Přednáška 01-08, cvičení 03-08: Teorie pravděpodobnosti – zabývá se popisem náhodnosti v experimentech či měřeních veličin, kdy za stejných vstupních podmínek nastávají různé výsledky.
- Přednáška a cvičení 09-12: Úsudková statistika – prezentace některých metod analýzy dat, kdy je informace získaná z náhodného výběru jedinců zobecněna na celou populaci. Její součástí je teorie odhadu, testování statistických hypotéz, statistická predikce (předpověď).

Na textu spolupracujeme společně s kolegou RNDr. Karlem Lepkou, Dr., kterého jsem požádal o korekturu textu a dodání dalších příkladů „ze života“ na oživení přednášky.

Břetislav Fajmon, leden 2022

1 Statistická a axiomatická definice pravděpodobnosti

Ve slově pravděpodobnost vidíme, že se jedná o něco podobného pravdě, nebo blízké pravdě. Při hledání odpovědi na otázku, co je to pravda, bychom asi slyšeli řadu odpovědí. Na tuto otázku se například zeptal i Pilát Ježíše Krista, když ho vyslýchal před jeho popravou. Neměl dost trpělivosti počkat si na odpověď, kterou už dříve Ježíš řekl svým učedníkům (Jan 14,6 – Bible): Já jsem ta cesta, pravda i život, nikdo nepřichází k Otci než skrze mne. Všeobsáhle mluvící apoštol Jan ve svém evangeliu také říká (Jan 1,17): Milost a pravda se stala skrze Ježíše Krista. Tj. pravda, to nejsou jen slova, ale pravdou lze nazvat též skutečnost (věci, které se odehrály – v pojetí Bible ovšem pravda je něco víc než jen to, co se odehrálo, ale ve slově pravda je obsažena i věrnost, že v životě Ježíše Krista byl Pán Bůh věrný vůči nám lidem). Tedy

- Pravda = skutečnost. Ve vědeckém smyslu má skutečnost všeobsáhlý význam, tj. věci dobré i špatné.

- Pravda = věrný či přesný popis skutečnosti. Z matematického hlediska se jedná i o zákonitosti, které stojí za skutečností, např. $2 + 2 = 4$. Z pohledu statistiky se může jednat o záznamy srážek v úhrnu za každý den na daném meteorologickém stanovišti.

V tomto předmětu se budeme zabývat tou částí pravdy, která souvisí se druhým uvedeným bodem: popisem skutečnosti, zejména tedy popisem měření číselných (kvantitativních) veličin, i když někdy lze matematicky zpracovat i veličiny kvalitativní (kvalitu či spokojenost s kvalitou dnes často vyjadřujeme i na číselné stupnici, kde např. 1 = velmi dobrá kvalita, 2 = spíše dobrá kvalita, 3 = průměr, 4 = spíše horší kvalita, 5 = špatná kvalita). Přitom

- Ve statistice (a statistické pravděpodobnosti) většinou jde o popis veličiny **a posteriori**[1], tj. na základě měření či zkušenosti (slovo zkušenost je pokusem o překlad slova empirie – empirické poznání je poznání, které můžeme číselně změřit či zaznamenat, např. informace o tom, jak nám v našem impériu prší).

- V pravděpodobnosti se jedná o popis veličin **a priori**[2], kdy na základě teoretických informací o povaze daného experimentu chceme popsat matematickým aparátem, jak nám v našem impériu budou padat čísla na kostce, líce či ruby na minci, apod. ještě dříve, než k měření dojde.

1.1 Dvě různá pojetí pravděpodobnosti

V Této první přednášce se budeme věnovat dvěma definicím (pojetím) pravděpodobnosti.

Definice 1.1 *Statistickou pravděpodobností náhodného jevu A, který při opakování experimentu či měření za stejných vstupních podmínek nastává jen v některých případech,*

[1] Z latinské předložky post či posteá, což znamená potom, poté.
[2] Z latinského prior = dřívejší, tj. a priori = dříve než, předem.

definujeme jako takové reálné číslo, ke kterému se blíží relativní četnost výskytu jevu A, pokud celý experiment opakujeme nekonečněkrát, tj.

$$P(A) = \lim_{n\to\infty} \frac{n(A)}{n}.$$

Předností této definice je fakt, že pokud jsme schopni dostatečně zajistit, aby měření experimentu bylo vykonáno vždy za stejných vstupních podmínek, máme predikci neboli předpověď pravděpodobnosti výskytu jevu A podloženu reálným měřením. Slabinou definice je to, že přesnou hodnotu limity nikdy nejsme schopni experimentálně zjistit – experiment nelze opakovat nekonečněkrát. Nicméně, při dostatečně velkém n lze tuto pravdě podobnost rozumně odhadnout.

Příklad 1.1 *Jaká je pst náhodného jevu A: při hodu kostkou padne šestka?*

Při statistickém určení psti vycházíme z opakování experimentu, hodíme tedy

- desetkrát, a pro $n = 10$ dostaneme $P(A) \doteq \frac{3}{10} = 0{,}3$;
- stokrát, a pro $n = 100$ dostaneme $P(A) \doteq \frac{20}{100} = 0{,}2$;
- tisíckrát, a pro $n = 1000$ dostaneme $P(A) \doteq \frac{170}{1000} = 0{,}17$; stále se nejedná o přesnou hodnotu, ale hodnotu blízkou hledané limitě.

Zatímco právě popsané statistické pojetí psti nebudeme zanedbávat, ještě častěji budeme užívat pojetí axiomatické, které je představené v následující definici (pomůcka pro zapamatování definice: potřebujete si pamatovat sedm faktů: 1) co je to Ω; 2,3,4) tři axiomy (1), (2), (3) pro náhodné jevy A_i; 5,6,7) tři axiomy (P1), (P2), (P3) pro pst P, která je definována jako zobrazení určitých vlastností).

Definice 1.2 *Pravděpodobnostním prostorem nazveme uspořádanou trojici $(\Omega, \mathcal{A}, \mathcal{P})$, kde*

- *Ω je tzv.* **základní prostor** *neboli množina všech možných elementárních výsledků ω_i daného měření či experimentu;*
- *$\mathcal{A}$ je tzv.* **jevové pole** *neboli taková třída podmnožin A_i (tyto podmnožiny nazýváme* **náhodné jevy***) množiny Ω (nemusí to být nutně všechny podmnožiny množiny Ω, ale takové, ...), že platí axiomy*
  - *(1) $\Omega \in \mathcal{A}$;*
  - *(2) pro $A_1, A_2 \in \mathcal{A}$ také $A_1 - A_2 \in \mathcal{A}$ (třída $\mathcal{A}$ je uzavřená na rozdíl náhodných jevů);*
  - *(3) pro $A_1, A_2, \dots \in \mathcal{A}$ také $\bigcup_{i=1}^{\infty} A_i \in \mathcal{A}$ (třída $\mathcal{A}$ je uzavřená na sjednocení nekonečně mnoha náhodných jevů $A_1, A_2, \dots$).*
- *$P : \mathcal{A} \to \langle \prime; \infty \rangle$ je zobrazení, které nazveme* **pravděpodobností na jevovém poli** *$\mathcal{A}$, když pro ně platí axiomy*

(P1) $P(\Omega) = 1$ *(axiom normovanosti);*

(P2) $P(A) \geq 0 \ \ \forall A \in \mathcal{A}$ *(axiom nezápornosti);*

(P3) *pro navzájem neslučitelné jevy* ($A_i \cap A_j = \emptyset$ *při* $i \neq j$) *platí*

$$P(\bigcup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i)$$

(axiom součtu pravděpodobností nekonečně mnoha neslučitelných náhodných jevů).

Právě uvedená axiomatická definice umožňuje korektně definovat všechny různé pravděpodobnostní modely – budeme se jimi zabývat v následující přednášce. Nejprve ovšem několik poznámek k ní:

Poznámka 1. Axiom (3) u jevového pole $\mathcal{A}$ lze vyslovit i pro konečně mnoho náhodných jevů $A_1, A_2 \ldots, A_n \in \mathcal{A}$. Nemusíme ho ovšem uvádět jako zvláštní axiom, plyne totiž z uvedeného třetího axiomu volbou $\emptyset = A_{n+1} = A_{n+2} = \ldots$: pro $A_1, A_2, \ldots, A_n \in \mathcal{A}$ také $\bigcup_{i=1}^{n} A_i \in \mathcal{A}$ (třída $\mathcal{A}$ je uzavřená i na sjednocení konečně mnoha náhodných jevů $A_1, A_2, \ldots$).

Poznámka 2. Axiom (P3) u pravděpodobnosti P lze vyslovit i pro konečně mnoho navzájem neslučitelných náhodných jevů $A_1, A_2 \ldots, A_n \in \mathcal{A}$. Nemusíme ho ovšem uvádět jako zvláštní axiom, plyne totiž z uvedeného třetího axiomu volbou $\emptyset = A_{n+1} = A_{n+2} = \ldots$:

$$(A_i \cap A_j = \emptyset \ \text{ pro } \ i \neq j) \implies P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i).$$

Poznámka 3. K jednotlivým axiomům psti:

ad P1) Pravděpodobnost jevu Ω je rovna jedné, protože Ω obsahuje všechny možné výsledky měření, které mohou nastat. Jev Ω proto někdy označujeme jako jev jistý, protože nastane vždy.

ad P2) Pravděpodobnost náhodného jevu A bude vždy číslo nezáporné.

ad P3) Jednotlivé náhodné jevy se navzájem vůči sobě vztahují podmínkou aditivity (= podmínkou sečtitelnosti): Pokud výskyt jevu A_1 se vylučuje s výskytem jevu A_2, pak pst jevu $A_1 \cup A_2$ lze vyjádřit jako součet pstí dílčích jevů A_1, A_2. Tento princip je vlastně analogický principu součtu v kombinatorice. A podobně jako kombinatorice počet sjednocení dvou množin konfigurací, které mají některé prvky ve společném průniku, nelze už vyjádřit jako součet prvků dílčích množin, ale musíme brát v úvahu jejich průnik a užít princip inkluze a exkluze, tak i u pravděpodobnosti obecného sjednocení jevů, z nichž dílčí dvojice jevů nemají prázdný průnik, musíme podobně užít složitější vzorec – budeme se mu věnovat ve třetí přednášce.

Kromě daných tří axiomů psti platí i další vlastnosti (viz následující dvě věty). Ty už nepokládáme za axiomy, protože jejich zdůvodnění plyne z axiomů už uvedených.

Věta 1.1
$$\forall A, B \in \mathcal{A}: \quad \mathcal{A} \subseteq \mathcal{B} \implies \mathcal{P}(\mathcal{A}) \leq \mathcal{P}(\mathcal{B}).$$

Důkaz: Množinu B lze rozdělit na dvě disjunktní části A, $B - A$, pro které na základě axiomu (P3) platí: $P(B) = P(A) + P(B - A)$. A protože $P(A) \geq 0$, $P(B - A) \geq 0$ na základě (P2), vidíme, že $P(B)$ získáme jako součet dvou nezáporných čísel, z nichž jedním je číslo $P(A)$ – proto platí tvrzeníčko věty. □

Věta 1.2 *Pro každý náhodný jev A a jev k němu opačný*[3] $\overline{A}$ *platí* $P(A) + P(\overline{A}) = 1$.

Důkaz: Celou množinu Ω lze rozdělit (rozložit) na dvě disjunktní množiny A, $\overline{A}$. Tedy na základě (P1) a (P3) platí $P(\Omega) = 1 = P(A) + P(\overline{A})$. □

Právě uvedená větička bude užitečná, namísto výpočtu $P(A)$, když bude pro nás jednodušší vypočítat pst $P(\overline{A})$ jevu opačného, a pak vyjádříme $P(A)$ jako $1 - P(A)$.

Než kapitolu ukončíme, tak ještě série dvou příkladů, která ilustruje rozdíl mezi pstí statistickou (empirickou) a pstí axiomatickou (teoretickou):

Příklad 1.2 *Byla získána data tím způsobem, že každá z dvaceti osob hodila čtyřikrát korunou (měříme tedy hodnotu veličiny X = počet líců ve čtyřech hodech korunou). V tabulce 1.1 jsou zaznamenány počty líců ve čtyřech hodech u každé z osob. Určete empirické rozdělení pravděpodobnosti veličiny X.*

Tabulka 1.1: K př. 1.2: Naměřené hodnoty veličiny X.

osoba	1	2	3	4	5	6	7	8	9	10
X-hodnota	3	1	1	3	1	2	0	2	4	4
osoba	11	12	13	14	15	16	17	18	19	20
X-hodnota	1	2	2	1	2	1	2	3	3	3

Řešení: Nejprve si všimněme, že naše veličina X nabývá pouze pěti hodnot, a to $0, 1, 2, 3$ *nebo* 4. Zpracování této úlohy je založeno na pojmu **četnost**, který udává počet výskytů dané hodnoty v našem souboru. Například ze všech dvaceti měření je jen jedna hodnota 0, tj. veličina X nabývá hodnoty 0 s četností 1 (budeme značit $c(0) = 1$). Hodnota 1 se vyskytuje s četností 6, atd. Všechny četnosti jsou zaznamenány v tabulce 1.2:

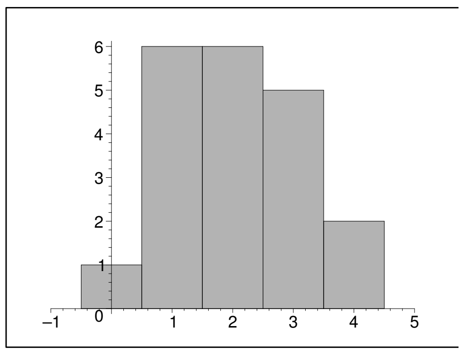
Musí platit jednoduchá kontrola, že součet všech četností ve druhém řádku tabulky je roven počtu hodnot (v našem případě 20).

[3] V teorii psti mají všechny množinové pojmy svůj terminologický ekvivalent. Množinově je $\overline{A}$ doplňkem množiny A, ale v teorii psti mluvíme o $\overline{A}$ jako o jevu opačném.

Tabulka 1.2: K př. 1.2: Tabulka empirických četností hodnot veličiny X.

X-hodnota	0	1	2	3	4
četnost	1	6	6	5	2

Uvedené četnosti lze také znázornit v tzv. **histogramu četností** - viz obr. 1.1, kde výšky jednotlivých obdélníčků jsou rovny konkrétním četnostem a délka základny každého z obdélníčků je rovna 1.



Obrázek 1.1: K příkladu 1.2: Histogram četností veličiny X.

K určení statistického (či empirického = naměřeného) rozdělení pravděpodobnosti nám zbývá poslední krok – vydělit četnosti délkou souboru (= počtem hodnot), v našem případě číslem 20. *Tak dostaneme tabulku 1.3 relativních četností vzhledem k počtu měření. Tyto relativní četnosti můžeme prohlásit za naměřené (statistické) psti různých hodnot veličiny X.*

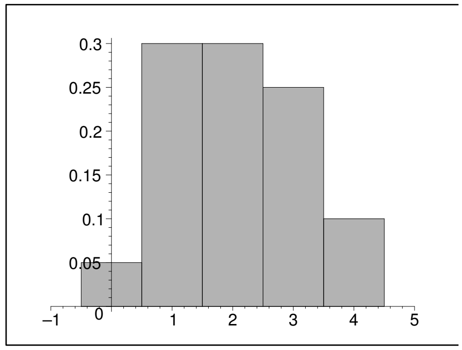
Tabulka 1.3: K př. 1.2: Funkce $p(x)$ empirického rozdělení pravděpodobnosti veličiny X.

X-hodnota	0	1	2	3	4
p(x)	0,05	0,3	0,3	0,25	0,1

Součet těchto relativních četností je roven jedné, jak bychom asi čekali.

Jediný rozdíl mezi obrázky 1.1 a 1.2 je v tom, že v prvním případě se na osu y nanáší hodnoty četnosti a ve druhém případě pravděpodobnosti. Na pravděpodobnostním histogramu je zajímavé to, že součet obsahů všech obdélníků na obrázku je roven jedné.

Obrázek 1.2 tedy představuje první krok k určení pro určení psti v tom prvním představeném, statistickém pojetí. Pro naprostou přesnost bychom opět museli provést



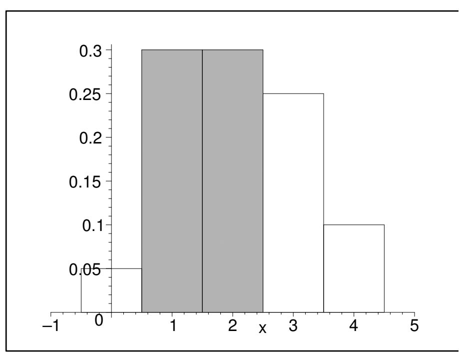
Obrázek 1.2: K př. 1.2: Histogram pravděpodobností veličiny X.

limity relativních četností pro celkový počet hodů jdoucí k nekonečnu. V této chvíli se ovšem spokojíme jen s tímto přibližným, ne naprosto přesným statistickým určením psti.

Pokud chceme s využitím histogramu pravděpodobnosti v našem diskrétním případě vyčíslit třeba pravděpodobnost,že při 4 *hodech mincí padl líc jednou nebo dvakrát, dostáváme*

$$P(X \in < 1, 2 >) = P(X = 1) + P(X = 2) = 0{,}3 + 0{,}3 = 0{,}6,$$

což je rovno součtu obsahů obdélníků histogramu nad hodnotami 1 *a* 2 *(viz obrázek):*



Nyní se věnujme popisu celého jevu teoreticky, aniž bychom jakkoli házeli korunou a počítali počty líců, tj. bez měření. Při popisu nám pomůže pst axiomatická (teoretická):

Příklad 1.3 *Nalezněte teoretické rozdělení veličiny X, která udává počet líců při čtyřech hodech mincí.*

Řešení: Podrobíme naši situaci teoretickým úvahám za předpokladu, že mince je vyvážená a vyrobená ze stejnorodého materiálu. V tabulce 1.4 jsou uvedeny všechny možné výsledky čtyř hodů mincí (druhý sloupec udává vždy počet líců v dané variantě):

Bystrému pozorovateli asi neušlo, že všech možných výsledků je 16*. A protože líc padá s pravděpodobností $\frac{1}{2}$, každý z těchto* 16 *výsledků je stejně pravděpodobný. A proto můžeme z tabulky určit četnosti počtu líců (viz tabulka 1.5)*

Tabulka 1.4: K př. 1.3: přehled všech možných výsledků při čtyřech hodech mincí.

výsledek	počet líců	výsledek	počet líců
LLLL	4	LRRL	2
LLLR	3	RLRL	2
LLRL	3	RRLL	2
LRLL	3	LRRR	1
RLLL	3	RLRR	1
LLRR	2	RRLR	1
LRLR	2	RRRL	1
RLLR	2	RRRR	0

Tabulka 1.5: K př. 1.3: Tabulka teoretických četností hodnot veličiny X.

X-hodnota	0	1	2	3	4
četnost	1	4	6	4	1

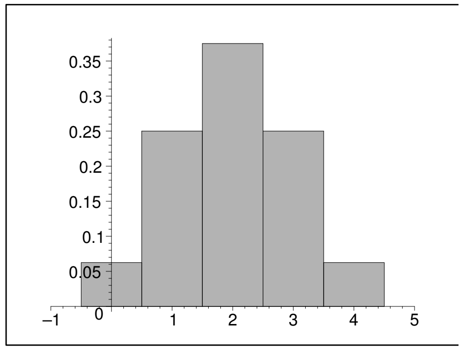
Tabulka 1.6: K př. 1.3: Funkce $p(x)$ teoretického rozdělení pravděpodobnosti veličiny X.

X-hodnota	0	1	2	3	4
p(x)	0,0625	0,25	0,375	0,25	0,0625

a vydělením hodnotou 16 *pak i relativní četnosti, které už jsou hodnotami hledané teoretické pravděpodobnostní funkce* $p(x)$ *(viz tabulka 1.6).*

Příslušný histogram pravděpodobnosti je znázorněn na obrázku 1.3.

K teoretickému rozdělení pravděpodobnosti v příkladu 1.3 lze jednoduše sestrojit teoretické rozdělení četnosti, a dokonce si můžeme vybrat, kolikrát se má experiment „prakticky“ provádět. Například pro 128 opakování experimentu čtyř hodů mincí má teoretické rozdělení četnosti stejný tvar jako pravděpodobnostní histogram 1.3, jen na osu y vynášíme hodnoty reprezentující četnost $c(i)$ (obrázek zde už není uveden, od 1.3 se liší jen měřítkem svislé osy):



Obrázek 1.3: K př. 1.3: Histogram pravděpodobnosti teoretického rozdělení veličiny X.

$$\begin{aligned}
c(0) &= p(0)\cdot 128 = 0{,}0625\cdot 128 = 8\\
c(1) &= p(1)\cdot 128 = 0{,}25\cdot 128 = 32\\
c(2) &= p(2)\cdot 128 = 0{,}375\cdot 128 = 48\\
c(3) &= p(3)\cdot 128 = 0{,}25\cdot 128 = 32\\
c(4) &= p(4)\cdot 128 = 0{,}0625\cdot 128 = 8
\end{aligned}$$

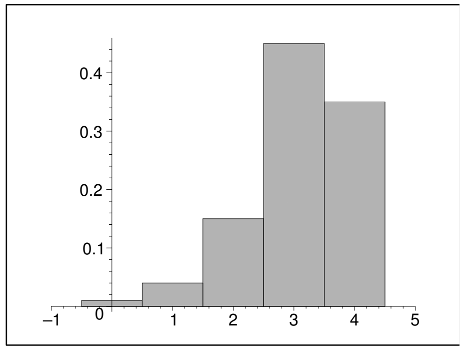
Čili kdybychom učinili 128 pokusů, z nichž jeden sestává ze čtyř hodů mincí, náš nejlepší teoretický odhad je ten, že v 8 pokusech by nepadl žádný líc, ve 32 pokusech jeden líc, atd.

Teoretické rozdělení pravděpodobnosti je jakési očekávané rozdělení, které nastane za jistých předpokladů. Například při pokusu 4 hodů mincí těmito předpoklady jsou:

- Mince je vyrobena tak, že rub a líc padá se stejnou pravděpodobností.
- Mincí je házeno „normálně“, ne nějakým divným stylem, který by zvýhodňoval buď rub, nebo líc.
- Každý účastník pokusu pravdivě nahlásí své výsledky.

Rozdělení získané empiricky v příkladu 1.2 „zhruba“ odpovídá teoretickému rozdělení z příkladu 1.3. Zdá se tedy rozumné uzavřít, že se světem je všechno v pořádku: mince je pravděpodobně dobře vyvážená, lidé jí házou dobrým způsobem a nahlašují výsledky poctivě.

Pokud by data z příkladu 1.2 vedla na empirické rozdělení pravděpodobnosti uvedené na následujícím obrázku, bylo by patrné, že tři nebo čtyři líce padaly ve čtyřech hodech mnohem častěji, než jsme očekávali, na úkor výsledků 0 líců, 1 líc, 2 líce. To by zpochybnilo některý z našich předpokladů. Uzavřeli bychom, že buď je mince nějak divně vyvážená, nebo lidé jí házejí divným stylem:



1.2 Shrnutí

Při popisu náhodnosti jevů, které se mohou odehrát, vycházíme ze dvou možných popisů: statistická pst vlastně popisuje budoucí náhodnost na základě minulého měření – i když není stoprocentně přesná, je založena na pozorování, měření dané veličiny, a tedy se snaží předpovědět chování veličiny podle toho, jak se (za analogických vnějších podmínek experimentu) chovala dosud. Příklad: pst toho, že na kostce padne šestka, stanovíme na základě toho, že danou kostkou hodíme dostatečněkrát a sledujeme výsledky hodu.

Naproti tomu, axiomatická definice psti se snaží podpořit popis náhodnosti který je teoretický – nevychází z měření, ale z teoretických předpokladů o tom, jak se bude veličina, kterou měříme, při zopakování analogických vnějších podmínek experimentu, chovat. Tyto teoretické předpoklady jsou sestrojeny na základě toho, že při měření veličiny zajistíme takové podmínky, ze kterých dané teoreticky spočtené psti vychází.

Porovnání obou pojetí je vidět na příkladech 1.2 a 1.3. Na obrázku 1.2 vidíme histogram pstí získaných měřením = empiricky = statisticky, na obrázku 1.3 histogram pstí získaných pouze teoretickými úvahami. Obě pojetí se při popisu náhodnosti používají: někdy máme k dispozici měření konkrétních hodnot nebo je můžeme snadno provést, jindy měření nemáme možnost zjistit a musíme se omezit na odhad dílčích pstí na základě předpokládaného charakteru a podmínek dané situace.

1.3 Otázky k opakování

U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.

Otázka 1.1 *Pravděpodobnost se zabývá otázkou: pokud vycházíme z jistého stavu světa, jaké důsledky budou následovat?*

Otázka 1.2 *Empirické rozdělení pravděpodobnosti je rozdělení, které získáme z naměřených dat.*

Otázka 1.3 *Empirické pravděpodobnosti jsou vlastně relativní četnosti.*

Otázka 1.4 *Třetí axiom psti je řečen jen pro sjednocení nekonečně mnoha jevů, nelze ho formulovat pro konečně mnoho disjunktních jevů.*

Otázka 1.5 *Pst opačného jevu* $\overline{A}$ *k jevu* A *vypočteme, když* $P(A)$ *vynásobíme jednou polovinou.*

Otázka 1.6 *Prvky jevového pole* $\mathcal{A}$ *jsou všechny možné podmnožiny* A *množiny* Ω*, které existují.*

Otázka 1.7 *Pro každé dva náhodné jevy platí:* $P(A \cup B) = P(A) + P(B)$.

Odpovědi na otázky viz 13.1.

2 Čtyři různé modely popisu pravděpodobnosti

Následující čtyři modely pravděpodobnosti se liší svou použitelností při popisu náhodnosti. Vždy záleží na vlastnostech množiny Ω všech možných elementárních výsledků měření, které mohou nastat.

2.1 Klasická pravděpodobnost

Klasickou pst můžeme užít v následujících podmínkách, a užíváme ji takto:

- Ω má konečně mnoho možných elementárních výsledků,
- všechny elementární výsledky mají stejnou možnost nastat.
- Pak lze počítat pst náhodného jevu $A \in \mathcal{A}$ podle vzorce $P(A) = \frac{|A|}{|\Omega|}$ (podíl počtu prvků obou množin)

Příklad 2.1 *Dvakrát hodíme hrací kostkou – jaká je pst, že součet hodnot obou hodů je roven 5?*

Řešení: Množina všech elementárních výsledků na dvou hodech (záleží na pořadí, ve kterém hodu co padlo) je

$$\Omega = \{[1;1],[1;2],[1;3],\dots,[6;5],[6;6]\},$$

náhodný jev A, na který se zaměřujeme, je množina

$$A = \{[1;4],[4;1],[2;3],[3;2]\}, \quad \text{a tedy} \quad P(A) = \frac{4}{36} = \frac{1}{9} \doteq 0{,}1111.$$

Příklad 2.2 *Z karet na mariáš (32 karet) vybereme náhodně po zamíchání čtyři karty. Jaká je pst, že aspoň jedna z nich bude eso?*

Řešení: V množině Ω budou elementárními výsledky všechny možné čtveřice karet (ve kterých nezáleží na pořadí, ale jen na tom, jaké karty jsou v té skupině čtyř vytažených). Náhodný jev A, na který se chceme zaměřit, obsahuje všechny čtveřice, ve kterých je právě jedno eso, právě dvě esa, právě tři esa, nebo čtyři esa. Najít celkový počet prvků množiny A znamená najít počty prvků ve všech čtyřech právě popsaných navzájem se vylučujících případech a sečíst je. Vnímavý student, který absolvoval předmět Diskrétní matematika, už nyní ví a píše výsledek, protože klasická pst převede otázku psti na otázku počtu prvků:

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\binom{4}{1}\cdot\binom{28}{3} + \binom{4}{2}\cdot\binom{28}{2} + \binom{4}{3}\cdot\binom{28}{1} + 1}{\binom{32}{4}} = 0{,}4305895 \doteq 0{,}4306.$$

Při výpočtu psti jevu se někdy ukazuje výhodné zapřemýšlet nad tím, zda existuje rychlejší cesta k cíli – zdá se, že nyní by existovala cesta vyčíslení opačného jevu a odečtení jeho psti od hodnoty 1, viz věta 1.2. Opačný jev $\overline{A}$ znamená, že ze čtyř náhodně vytažených karet nebude žádné eso. Pak

$$P(A) = 1 - P(\overline{A}) = 1 - \frac{\binom{28}{4}}{\binom{32}{4}} \doteq 0{,}4306.$$

2.2 Geometrická pravděpodobnost

Geometrickou pst můžeme užít v následujících situacích, a užíváme ji takto:

- Ω má nekonečně mnoho možných elementárních výsledků, dokonce tak velkou, že je oblastí kladné míry (úsečkou kladné délky, plochou kladného obsahu, tělesem kladného objemu).
- všechny elementární výsledky mají stejnou možnost nastat.
- Pak lze počítat pst náhodného jevu $A \in \mathcal{A}$ podle vzorce $P(A) = \frac{\mu(A)}{\mu(\Omega)}$ (podíl měr obou množin, tj. podle charakteru množin podíl délek, obsahů nebo objemů).

Příklad 2.3 *Tramvaj jezdí v pracovní době každých 7 minut. Student, který se nedívá na hodinky a přichází na zastávku tramvaje náhodně, bude určitou dobu čekat na příjezd další tramvaje. Jaká je pst, že bude čekat 4 a více minut?*

Řešení: Nelze počítat podle klasického modelu, protože množiny Ω, A jsou obě nekonečné; ale všimněme si, že při náhodném příchodu studenta na zastávku jsou všechny doby čekání stejně pravděpodobné, tj. splňují předpoklady geometrického modelu psti.

Množina všech možných výsledků je totiž $\Omega = \langle 0; 7\rangle$, náhodný jev A na který se chceme zaměřit, je $A = \langle 4; 7\rangle$ ... doba čekání na nejbližší tramvaj, která je dlouhá 4 a více minut. Díky splněným oběma předpokladům geometrického modelu psti můžeme využít délky obou množin:

$$P(A) = \frac{d(A)}{d(\Omega)} = \frac{3}{7} = 0{,}4285714 \doteq 0{,}4286.$$

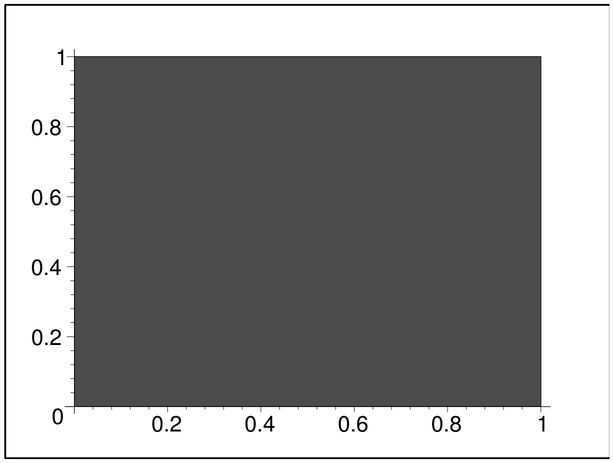
Příklad 2.4 *Honza a Marek se domluvili, že se setkají na jistém místě mezi osmou a devátou hodinou, kam každý z nich v tu dobu náhodně přijde. Ale řekli si, že ten, kdo přijde první, bude na toho druhého čekat jen 15 minut, a pak odejde. Jaká je pravděpodobnost, že se setkají?*

Řešení: Označme
$8 + x \ldots$ čas příchodu Honzy (v hodinách);
$8 + y \ldots$ čas příchodu Marka.

Víme, že oba přijdou určitě do devíti hodin, tedy $0 \leq x \leq 1$, $0 \leq y \leq 1$. Každý výsledek jejich příchodu lze vyjádřit jako uspořádanou dvojici (x, y), což lze znázornit – a uvidíme, že to bude pomocí – jako bod v rovině, jehož obě souřadnice leží v intervalu $< 0, 1 >$. Všechny tyto body modelující možný výsledek příchodů vytvářejí tedy čtverec v rovině. Tento čtverec $\Omega = \{(x, y) : 0 \leq x \leq 1,\ 0 \leq y \leq 1\}$ je množinou všech možných výsledků dané situace (viz obrázek 2.4).

Počet všech možných případů je sice nekonečný, ale jsme schopni spočítat obsah čtverce: $S(\Omega) = 1 \cdot 1 = 1$.

Označme dále



Obrázek 2.4: K př. 2.4: Množina všech možných výsledků.

$A \ldots$ Honza a Marek se setkají

Příznivým případům jevu A odpovídají ty příchody (x, y) obou studentů, ve kterých se x od y liší nanejvýš o 15 minut, což je asi $\frac{1}{4}$ hodiny. Pro tyto „příznivé" body čtverce Ω tedy musí platit nerovnost

$$|y - x| \leq \frac{1}{4}.$$

Vyřešme tuto nerovnost. Při odstraňování absolutní hodnoty musíme rozlišit dvě situace:

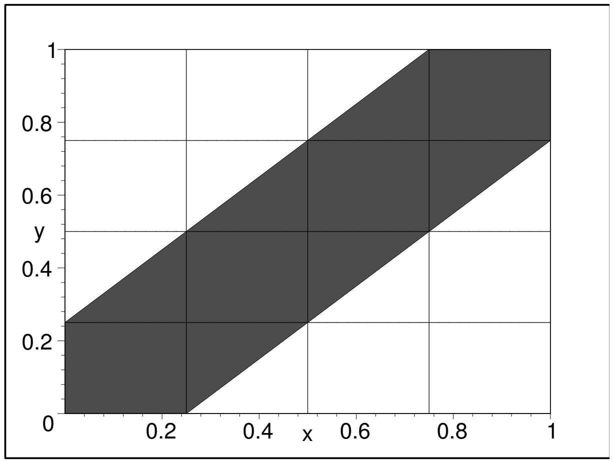
- *Pro $y - x \geq 0$ se znaménka nemění, tj $y - x \leq \frac{1}{4}$, odtud $y \leq x + \frac{1}{4}$.*
- *Pro $y - x < 0$ musíme při odstraňování absolutní hodnoty na levé straně nerovnosti změnit znaménka: $-y + x \leq \frac{1}{4}$, odtud $y \geq x - \frac{1}{4}$.*

Body splňující obě z uvedených nerovností získáme jako průnik polorovin každou z nerovností určených, tj. 2.5:

Jev A lze tedy vyjádřit jako množinu bodů v rovině:

$$A = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1, y \leq x + \frac{1}{4}, y \geq x - \frac{1}{4}\}.$$

Příznivých případů je také nekonečně mnoho, ale jsme schopni vypočítat míru této nekonečnosti, konkrétně řečeno obsah množiny A: nejjednodušeji $S(A)$ vypočteme z grafického znázornění na obrázku 2.5, když budeme brát v úvahu rozdělení čtverce Ω na šestnáct menších čtverečků o straně délky $\frac{1}{4}$. Je vidět, že množina A zabírá plochu sedmi z těchto čtverečků, a protože $S(\Omega) = 1$, máme $S(A) = \frac{7}{16} \cdot S(\Omega) = \frac{7}{16}$.



Obrázek 2.5: K př. 2.4: Množina všech příznivých výsledků.

Pravděpodobnost jevu A teď určíme jako podíl míry množiny příznivých případů a míry množiny všech možných případů:

$$P(A) = \frac{S(A)}{S(\Omega)} = \frac{\frac{7}{16}}{1} = \frac{7}{16} = 0{,}4375.$$

2.3 Diskrétní pravděpodobnost

Diskrétní pst můžeme užít v následujících situacích, a užíváme ji takto:

- Ω má konečně mnoho možných elementárních výsledků, nebo je jich stejně najko přirozených čísel (říkáme, že Ω je množina nejvýše spočetně nekonečná).
- všechny elementární výsledky mají obecně RŮZNOU možnost nastat (elementární výsledek ω_i nastane s pstí $p(\omega_i)$).
- Pak lze počítat pst náhodného jevu $A \in \mathcal{A}$ podle vzorce

$$P(A) = \sum_{\omega_i \in A} p(\omega_i)$$

V případě diskrétní psti tři axiomy z předchozí kapitoly přecházejí v následující axiomy pro pstní funkci p:

1. $\sum_{w_i \in \Omega} p(\omega_i) = 1$ (axiom normovanosti),
2. $p(\omega_i) \geq 0$ pro každý elementární výsledek $\omega_i \in \Omega$ (axiom nezápornosti),
3. $P(A) = \sum_{\omega_i \in A} p(\omega_i)$ (axiom součtu pstí neslučitelných jevů, protože dílčí výsledky ω_i jsou navzájem neslučitelné).

Příklad 2.5 *Házeme hrací kostkou tak dlouho, až poprvé padne šestka; pak skončíme s házením.* ω_i *je jakákoli přípustná sekvence hodů za těchto podmínek. Určete pst, že na první šestku budeme potřebovat více než tři hody.*

Řešení: Nejnižší možný počet hodů je jeden, kdy právě tímto pokusem hned padne šestka:

$$P(6) = \frac{1}{6} = 0{,}1667.$$

Dále může nastat sekvence $N6$, tj. prvním hodem šestka nepadne a druhým ano – a to s pstí

$$P(N6) = \frac{5}{6} \cdot \frac{1}{6} = \frac{5}{36} \doteq 0{,}1389.$$

Pak se také může stát, že nastane sekvence $NN6$, a sice s pravděpodobností

$$P(NN6) = \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{1}{6} \doteq 0{,}1157.$$

Teoreticky je možné, že nastane pro přirozené číslo k nastane sekvence $NN \ldots N6$ s pstí

$$P(\underbrace{NN \ldots N}_{\text{k-krát}}6) = \underbrace{\frac{5}{6} \cdot \frac{5}{6} \cdot \cdots \cdot \frac{5}{6}}_{\text{k-krát}} \cdot \frac{1}{6} = \left(\frac{5}{6}\right)^k \cdot \frac{1}{6}.$$

Je jasně vidět, že jednotlivé dílčí sekvence nastávají s různými pstmi. Takových sekvencí je vlastně nekonečně mnoho a víme, že musí splňovat vztah

$$\sum_{k=0}^{\infty} P(\underbrace{NN \ldots N}_{\text{k-krát}}6) = 1$$

(axiom normovanosti). Funkce, jejíž hodnoty jsme právě určili, se nazývá **pravděpodobnostní funkce**. V kapitole 8 si o tomto příkladu řekneme ještě něco více. Nyní se vraťme k zadání příkladu a spočteme pst, že pro první hozenou šestku potřebujeme více než tři hody:

$$P(A) = p(NNN6) + p(NNNN6) + p(NNNNN6) + \cdots .$$

Zdá se, že součet nekonečné řady by mohl dát práci, ovšem využijeme opět věty 1.2 a odečteme hledanou pst od opačného jevu:

$$P(A) = 1 - P(\overline{A}) = 1 - (p(6) + p(N6) + p(NN6)) = 1 - (\frac{1}{6} + \frac{5}{36} + \frac{25}{216}) \doteq 0{,}5787.$$

2.4 Spojitá pravděpodobnost

Spojitou pst můžeme užít v následujících situacích, a užíváme ji takto:

- Ω má nespočetně nekonečně mnoho možných elementárních výsledků, a tedy se jedná o interval reálných čísel nebo o R^+, nebo R.

- Tyto elementární výsledky mají obecně RŮZNOU možnost nastat.
- Pak pro pst jevu $(a;b) \subseteq \Omega$ nebo $\langle a;b) \subseteq \Omega$ nebo $(a;b\rangle \subseteq \Omega$ nebo $\langle a;b\rangle \subseteq \Omega$ platí

$$P(a;b) = P(\langle a;b) = P((a;b\rangle) = P\langle a;b\rangle) = \int_a^b f(x)dx,$$

kde $f(x)$ je nezáporná po částech spojitá funkce.

V případě spojité psti tři axiomy z předchozí kapitoly přecházejí v následující axiomy pro hustotu psti $f(x)$:

1. $\int_{-\infty}^{+\infty} f(x)dx = 1$ (axiom normovanosti),
2. $\forall a, b \in R \ : \ \int_a^b f(x)dx \geq 0$ (axiom nezápornosti),
3. $P(a;b) = P(\langle a;b) = P((a;b\rangle) = P\langle a;b\rangle) = \int_a^b f(x)dx$ (axiom součtu neslučitelných jevů má formu integrálu).

Příklad 2.6 *mobily jisté značky a typu mají životnost patnáct let. Jaká je pst, že náhodně koupený mobil tohoto typu vydrží více než deset let?*

Řešení: Jak zdůvodníme později, lze odvodit, že za hustotu psti, která dobře modeluje životnost, lze vzít funkci, která je pro záporná x rovna nule, a pro kladná x je $f(x)$ klesající exponenciální funkcí, ve které vystupuje číslo 15 z důvodu průměrné životnosti. Tedy za funkci $f(x)$ lze vzít funkci

$$f(x) = \begin{cases} 0 \ \ldots \ \text{pro } x < 0; \\ \frac{1}{15} \cdot e^{\frac{-x}{15}} \ \ldots \ \text{pro } x \geq 0 \end{cases}$$

A pokud máme počítat pravděpodobnost, že životnost je větší než deset let (časová jednotka je tedy jeden rok), integrujeme na intervalu od 10 do ∞:

$$P((10;\infty)) = \int_{10}^{\infty} \frac{1}{15} \cdot e^{\frac{-x}{15}} dx = e^{\frac{-10}{15}} \doteq 0{,}5134.$$

2.5 Shrnutí

Pokud výsledky jistého pokusu, hry nebo experimentu mohou nastat se stejnou pravděpodobností, používáme k jeho popisu klasickou (2.1) nebo geometrickou (2.2) pravděpodobnost. Ovšem pokud některé z elementárních výsledků nastávají častěji než jiné, situaci znázorníme pomocí diskrétní (2.3) nebo spojité (2.4) pravděpodobnosti. Z modelů je vidět, že (2.1) je speciálním případem modelu (2.3) a že model (2.2) je speciálním případem modelu (2.4). Tedy dvěma hlavními modely psti je distrétní model (2.3) a spojitý model (2.4).

Když studujeme jistou veličinu, jako první věc bychom si měli uvědomit, zda se jedná o veličinu diskrétní (ta nabývá hodnot z konečné (např. $\{1, 2, 3, 4, 5, 6\}$) nebo spočetné (např. **N**, **Z**) množiny Ω) nebo veličinu spojitou (ta nabývá hodnot z reálného intervalu $\Omega =< a, b >$ nebo z celé množiny reálných čísel). Popis těchto dvou typů veličin se totiž v některých věcech liší. A používané vzorce nebo způsob popisu se neustále odvíjí od jednoho z těchto dvou typů. V následujících kapitolách (a i v úlohách praxe) se potřebuje občas určit pravděpodobnost, že náhodná veličina nabývá hodnot z jistého intervalu $(a, b\rangle$. S ohledem na typ veličiny budeme užívat vzorec

$$P(X \in (a, b\rangle)) = P(a < X \leq b) = \begin{cases} \sum_{a<k\leq b} p(k) & \text{pro diskrétní veličinu } X, \\ \int_a^b f(x)dx & \text{pro spojitou veličinu } X. \end{cases}$$

V diskrétním případě se funkce $p(k)$ nazývá pravděpodobnostní funkce, ve spojitém případě funkci $f(x)$ říkáme hustota psti.

2.6 Otázky k opakování

U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.

Otázka 2.1 *Všech možných výsledků experimentu, který lze popsat pravděpodobnostním modelem, může být nejvýše spočetně mnoho.*

Otázka 2.2 *Geometrická pravděpodobnost je speciálním případem spojité pravděpodobnosti.*

Otázka 2.3 *Diskrétní náhodná veličina nemůže nabývat všech hodnot se stejnou pravděpodobností.*

Otázka 2.4 *U spojité náhodné veličiny X je pravděpodobnost, že X nabude konkrétní hodnoty, vždy rovna nule.*

Otázka 2.5 *Hustota spojité náhodné veličiny nemůže nikdy mít bod nespojitosti.*

Otázka 2.6 *Každou nezápornou funkci $f(x)$, pro kterou $\int_{-\infty}^{\infty} f(x)dx = 1$, lze označit za hustotu jisté náhodné veličiny.*

Otázka 2.7 *Žádná hodnota pstní funkce $p(k)$ nemůže být větší než* 1.

Otázka 2.8 *Žádná funkční hodnota hustoty psti $f(x)$ nemůže být větší než* 1.

Odpovědi na otázky viz 13.2.

3 Věta o součtu pravděpodobností, věta o součinu pravděpodobností, podmíněná pst

Od axiomu (P3) nyní pokročíme ke slibované psti sjednocení obecných náhodných jevů, a pak se budeme zabývat samozřejmě i pstí průniku jevů, protože psti průniku se vlastně objevují právě i ve větě o psti sjednocení jevů.

Začneme pstí dvou obecných jevů:

Věta 3.1 *(věta o součtu pstí dvou náhodných jevů) Pro libovolné náhodné jevy $A_1, A_2 \in \mathcal{A}$ platí:*

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2).$$

Důkaz by byl velmi jednoduchý: protože na základě axiomu (P3) pro disjunktní jevy $A_1 - A_2$ a $A_1 \cap A_2$ platí $P(A_1) = P(A_1 - A_2) + P(A_1 \cap A_2)$ a podobně platí $P(A_2) = P(A_2 - A_1) + P(A_1 \cap A_2)$, dosazením těchto vztahů do pravé strany rovnosti dostaneme $P(A_1 - A_2) + P(A_2 - A_1) + P(A_1 \cap A_2)$, což je součet pstí tří disjunktních množin, a tak podle axiomu (P3) přesně roven jejich sjednocení $A_1 \cup A_2$.

Příklad 3.1 *Uvažujme situaci*[4]*, kdy zásilkový prodejce vyřizuje objednávky zboží telefonicky, emailem, nebo formulářem při osobním odběru. Objednávky dělíme podle typu na malé, střední, velké a prioritní. Lze označit následující náhodné jevy, které mohou nastat:*

T … *náhodně vybraná-přišlá objednávka je telefonická;*

E … *náhodně vybraná-přišlá objednávka je emailem;*

F … *náhodně vybraná-přišlá objednávka je formulářem při osobním odběru;*

M … *náhodně vybraná-přišlá objednávka je malá;*

S … *náhodně vybraná-přišlá objednávka je střední;*

V … *náhodně vybraná-přišlá objednávka je velká;*

H … *náhodně vybraná-přišlá objednávka je prioritní (high priority).*

Jsou známa dosavadní data z letošního roku:

typ objedn.	*malá obj.*	*střední obj.*	*velká obj.*	*high priority obj.*	*celkem*
telefon	1021	216	109	14	1360
email	86	371	308	49	814
formulář	1497	230	86	13	1826
celkem	2604	817	503	76	4000

Určete pst toho, že náhodně přišlá-obdržená objednávka bude emailem nebo prioritní.

[4] Příklad je vzat z učebnice (Robová, Hála, Calda, 2013), ale v trochu jiných souvislostech, s jinou otázkou v zadání.

Řešení: jedná se vlastně o pst statistickou, odhadnutou pomocí naměřených dat – takto určit náhodnost v tomto případě je nejrozumnější[5].

Dále budeme při výpočtu $P(E \cup H)$, protože logická spojka NEBO je v množinové symbolice vyjádřená sjednocením, používat klasický model psti, protože každý klient svou volbou „hlasuje“ pro jeden typ i jeden způsob objednávky. Je vidět, že průnik obou jevů je neprázdný, tj. existují klienti, kteří si objednali emailem, a současně s nejvyšší prioritou, tj. použijeme větu 3.1:

$$P(E \cup H) = P(E) + P(H) - P(E \cap H) = \frac{814 + 76 - 49}{4000} = \frac{841}{4000} \doteq 0{,}21025.$$

Věta 3.2 *(věta o součtu pstí n náhodných jevů) Pro libovolné náhodné jevy A_1, A_2, atd. až A_n platí:*

$$\begin{aligned}
&P(A_1 \cup A_2 \cup \cdots \cup A_n) = P(A_1) + P(A + 2) + \cdots + P(A_n) - \\
&-P(A_1 \cap A_2) - P(A_1 \cap A_3) - \ldots - P(A_{n-1} \cap A_n) + \\
&+P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + \cdots + P(A_{n-2} \cap A_{n-1} \cap A_n) - \\
&\ldots \\
&+(-1)^{n-1} \cdot P(A_1 \cap A_2 \cap \cdots \cap A_n).
\end{aligned}$$

Důkaz bychom museli vést pro $n = 3$ pomocí Vennova diagramu pro tři množiny v obecné poloze (se všemi možnými neprázdnými průniky), pro $n \geq 4$ už Vennovy množiny selhávají, protože už nedokážou postihnout obecnou polohu čtyř a více překrývajících se množin (u čtyř množin bychom jejich vzájemný vztah museli znázornit pomocí čtyř koulí z nichž každá má střed v jiném vrcholu čtyřstěnu, a všechny mají poloměr o něco větší, než je polovina délky hrany čtyřstěnu; už to je náročné, tj. důkaz bychom vedli spíše pomocí charakteristického obecného prvku x).

Ale logika důkazu je celkem jasná: při sečtení pstí jednotlivých množin jsme průniky každých dvou započítali víckrát, takže jejich psti musíme na druhém řádku pravé strany odečíst – ovšem průniky každých tří množin jsme z psti odečetli tolikrát, že tam zase vůbec nejsou započítány, takže jejich psti musíme na třetím řádku pravé strany zase přičíst, atd. až u průniku všech n množin vlastně nevíme, zda jej máme přičíst nebo odečíst, ale to lze snadno zjistit pomocí $(-1)^{n-1}$, protože tento člen ctí střídání sčítání a odčítání pstí na každém dalším řádku pravé strany.

Příklad 3.2 *Elektrikář vystiskl štítky na zvonky bytového domu se čtyřmi byty a zapojuje je náhodně, protože jeho parťák onemocněl a nepřinesl seznam (a nebere telefon). Jaká je pst, že*

a) *Všechny zvonky zapojí ke správným bytům?*

[5] Nebudeme se nyní zabývat výpočtem limity hodnot v tabulce v situaci, kdy počet zákazníků nebude 4000, ale bude se blížit k nekonečnu. Taková data nemáme totiž k dispozici. Namísto toho se spokojíme s tím, že pst výsledku bude určená relativní četností, kterou bychom mohli pro rostoucí počet měření-zákazníků neustále zpřesňovat. Také je pravdou, že vzorek čtyř tisíc zákazníků už o klientele firmy něco vypovídá, tj. má smysl zhruba pst odhadnout pomocí relativní četnosti.

b) *Aspoň jeden zvonek zapojí správně?*

c) *Žádný zvonek nezapojí správně?*

Řešení: je asi nabíledni, že jedna z otázek a), b), c) bude řešena pomocí právě uvedené věty 3.2, a ty ostatní dvě otázky nějak jinak. Kdo se v tom má vyznat?

Začneme tím, že si označíme náhodné jevy, které by možná mohly vést k výsledku v každé otázce – v tom je matematika právě pomocná, že dobré označení či popis situace už je prvním krokem k řešení.

A_1 ... zvonek k bytu č. 1 bude zapojen správně;

A_2 ... zvonek k bytu č. 2 bude zapojen správně;

A_3 ... zvonek k bytu č. 3 bude zapojen správně;

A_4 ... zvonek k bytu č. 4 bude zapojen správně.

No a nyní vyjádříme slovní popis jevů a,b,c pomocí náhodných jevů A_1, A_2, A_3, A_4, snad to bude možné (aniž jsme si tedy vyjádřili množinu Ω všech možných výsledků – brzy se k ní dostaneme):

ad a) $A_1 \cap A_2 \cap A_3 \cap A_4$ ... každý ze zvonků bude zapojen správně, tj. dílčí jevy nastanou současně.

ad b) „Aspoň jeden zvonek bude zpojen ke správnému bytu“ lze ekvivalentně vyjádřit výrokem „k 1. bytu bude zvonek zapojen správně NEBO ke 2. bytu bude zvonek zapojen správně NEBO ke třetímu bytu bude zvonek zapojen správně NEBO ke 4. bytu bude zvonek zapojen správně“, což je právě $A_1 \cup A_2 \cup A_3 \cup A_4$.

ad c) Po krátkém přemýšlení (ať logickém,nebo množinovém) lze dospět k tomu, že jev c) je opačným jevem k jevu b), neboli jev c) lze vyjádřit jako

$$\overline{A_1 \cup A_2 \cup A_3 \cup A_4}.$$

Je tedy jasné že $P(c) = 1 - P(b)$.

Začněme výpočtem psti jevu (a): Nyní se musíme zamyslet nad tvarem množiny Ω: chceme modelovat přiřazení čtyř zvonků k bytům, tj. množina všech možných výsledků by mohla obsahovat všechny možné čtveřice čísel 1, 2, 3, 4 (ale na to už jsme skoro odborníci – to se pozná, když tyto čtveřice budeme nazývat permutacemi), a přitom pokud např. 2 bude na druhém místě této čtveřice, bude to znamenat, že zvonek 2 je správně zapojen k bytu číslo 2. Nyní pst průniku vypočteme klasickým modelem, protože počet možných čtveřic je konečný a každá při náhodném zapojení má stejnou šanci nastat:

$$P(a) = P(A_1 \cap A_2 \cap A_3 \cap A_4) = \frac{1}{24} = 0{,}416666\ldots,$$

protože příznivý případ správného zapojení všech zvonků je jediný a všech možných zapojení neboli permutací čtyřprvkové množiny je 24.

Dostáváme se k výpočtu (b), kde využijeme větu 3.2 pro $n = 4$:

$$\begin{aligned}&P(A_1 \cup A_2 \cup A_3 \cup A_4) = P(A_1) + P(A_2) + P(A_3) + P(A_4) -\\&-P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_1 \cap A_4) - P(A_2 \cap A_3) - P(A_2 \cap A_4) - P(A_3 \cap A_4) +\\&+P(A_1 \cap A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_4) + P(A_1 \cap A_3 \cap A_4) + P(A_2 \cap A_3 \cap A_4) -\\&-P(A_1 \cap A_2 \cap A_3 \cap A_4).\end{aligned}$$

Návratem k permutacím a jejich počtům dostaneme výsledek

$$\frac{1}{24}(3! + 3! + 3! + 3! - 2! - 2! - 2! - 2! - 2! - 2! + 1 + 1 + 1 + 1 - 1) = 0{,}625.$$

No a poslední výsledek c) dostaneme už přes pst opačného jevu: $p(c) = 1 - 0{,}625 = 0{,}375$.

Čili hlavním zájmem našich úvah bylo vypočítat část (b) použitím věty o součtu – doporučuji si pamatovat, že **ve větě o součtu se počítá pst sjednocení**[6], i když se v ní vyskytují průniky jevů! Ovšem sjednocení je v této větě jejím hlavním zájmem.

V další části tohoto oddílu se skutečně zaměříme více na pst průniku jevů – takové psti jsme schopni počítat např. pomocí klasického modelu psti, ale řekneme si výpočtům psti průniku ještě několik důležitých věcí, z nichž nejdůležitější je rozdíl při výpočtu psti průniku jevů podmíněných a jevů nezávislých. Tento rozdíl bude vysvětlen na následujících dvou příkladech.

Příklad 3.3 *Házeme třikrát hrací kostkou. Označme náhodné jevy*

A *… při prvním hodu padne jednička;*

B *… při druhém hodu padne dvojka;*

C *… při třetím hodu padne trojka;*

Jaká je pst náhodného jevu $A \cap B \cap C$?

Odpověď na tuto otázku získáme podobně, jako získáme počet konfigurací vynásobením dílčích konfigurací při kombinatorickém principu součinu – vynásobíme dílčí psti. Jinými slovy,

$$P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C) = \frac{1}{6} \cdot \frac{1}{6} \cdot \frac{1}{6} = 0{,}004629.$$

Situace při výpočtu psti průniku nebude ovšem takto jednoduchá vždy: v příkladu 3.3 totiž dílčí jevy jsou na sobě navzájem náhodně nezávislé, neboli pst padnutí dvojky při druhém hodu není ovlivněna tím, jaký byl výsledek prvního hodu, a pst padnutí trojky

[6]Pokud bychom například přehodili všechny symboly sjednocení za symboly průniku a symboly průniku za symboly sjednocení, za prvé dostaneme nesmysl, a za druhé u zkoušky bychom za tento nesmysl nedostali žádné body.

při třetím hodu, pokud kostkou házeme nestranně, stále stejným stylem, nezávisí na tom, co na kostce padlo při hodu prvním a hodu druhém. Tato skutečnost, že pst průniku je rovna součinu dílčích pstí, nám poslouží jako definice nezávislých jevů[7]:

Definice 3.1 • *Dva náhodné jevy $A_1, A_2 \in \mathcal{A}$ se nazývají stochasticky nezávislé (náhodně nezávislé), když $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2)$.*

- *n náhodných jevů $A_1, A_2, \dots, A_n \in \mathcal{A}$ se nazývají stochasticky nezávislé (= náhodně nezávislé), když*

  (i) $P(A_i \cap A_j) = P(A_i) \cdot P(A_j)$ pro $1 \le i < j \le n$,

  (ii) $P(A_i \cap A_j \cap A_k) = P(A_i) \cdot P(A_j) \cdot P(A_k)$ pro $1 \le i < j < k \le n$,

  (iii) ...

  (iv) $P(A_1 \cap A_2 \cap A_3 \cap \dots \cap A_n) = P(A_1) \cdot P(A_2) \cdot P(A_3) \cdot \dots \cdot P(A_n)$

(tedy ověřujeme $2^n - n - 1$ vztahů).

Třeba v příkladu 3.3 jsou jevy A, B nezávislé, protože platí $P(A \cap B) = P(A) \cdot P(B)$. A také jsou jevy A, B, C (stochasticky) nezávislé, protože platí $P(A \cap B) = P(A) \cdot P(B)$, $P(A \cap C) = P(A) \cdot P(C)$, $P(B \cap C) = P(B) \cdot P(C)$, $P(A \cap B \cap C) = P(A) \cdot P(B) \cdot P(C)$ (nezávislost více než dvou jevů je třeba ověřit platností rovností, kde zkoumáme průniky jakýchkoli množin, které lze z daných množin sestrojit). Čtenář by si mohl možná všimnout, že pst průniku $A \cap B \cap C$ v příkladu 3.3 je kladná (náhodný jev může nastat), a přesto je tato trojice jevů stochasticky nezávislá. Pojem nezávislosti jevů přímo tedy nesouvisí s tím, zda průniky utvářené z těchto jevů jsou neprázdné nebo prázdné[8], ale spíše jakým způsobem se pravděpodobnost těchto průniků vypočte. Tedy i při neprázdných průnicích jevů jsou tyto dílčí jevy někdy závislé, někdy nezávislé – závislost jevů při neprázdném průniku uvidíme v příkladu následujícím.

Příklad 3.4 *Ze sedmi telefonů daného typu na skladu prodejny jsou tři nekvalitní a čtyři kvalitní. Pro zákazníka prodejce náhodně z daných sedmi vybírá tři telefony, které on chce koupit pro svou rodinu. Označme náhodné jevy*

K_1 *... první náhodně vybraný telefon je kvalitní;*

K_2 *... druhý náhodně vybraný telefon je kvalitní;*

N_3 *... třetí náhodně vybraný telefon je NEkvalitní.*

Určete pst toho, že nastanou tyto tři náhodné jevy současně při náhodném výběru tří mobilních telefonů.

[7] Viz např. Budíková, Králová, Martoš, 2009, str. 59.

[8] Náhodné jevy, jejichž průnik je prázdný, se nazývají **neslučitelné**, nikoli **nezávislé** – prosím tyto dva pojmy nezaměňujte. Ovšem vztah logické implikace mezi těmito dvěma pojmy skutečně existuje: pokud například $A \cap B \cap C = \emptyset$, a přitom $P(A) > 0$, $P(B) > 0$, $P(C) > 0$, tak jsou náhodné jevy A, B, C stochasticky závislé, protože $P(A \cap B \cap C) = P(\emptyset) = 0$, ale $P(A) \cdot P(B) \cdot P(C) > 0$. Tedy jestliže neslučitelné jevy s kladnými pstmi mají prázdný průnik, pak jsou stochasticky závislé.

Podobně budeme počítat pst průniku náhodných jevů, ale nyní podle trochu jiného vzorce:

$$P(K_1 \cap K_2 \cap N_3) = P(K_1) \cdot P(K_2|K_1) \cdot P(N_3|K_1 \cap K_2).$$

Nyní totiž pst výběru druhého kvalitního telefonu bude záviset na tom, jaký byl vybrán první telefon. Označení svislé čáry v zápise $P(K_2|K_1)$ vyjadřuje jakousi navazující pst – uvažujeme situaci K_1, na kterou bude navazovat situace $K2$ v tom smyslu, že i když K_1 je náhodný jev, který obecně nastat nemusí, v tomto NAVAZOVACÍM MÓDU budeme předpokládat, že nastal.

$P(K_1) = \frac{4}{7}$, a nyní se ptáme, jaká je pst, že i druhý vybraný mobil bude kvalitní, pokud ten první byl kvalitní: vypočteme $P(K_2|K_1) = \frac{3}{6}$ (při prvním výběru byl odebrán kvalitní, tj. zbývají 3 příznivé případy, všech možných je už jen 6).

Definice 3.2 $P(K_2|K_1)$ *označuje pst jevu* K_2 *vázanou podmínkou, že nastal rovněž jev* K_1 ... **podmíněná pst jevu K_2 za podmínky, že nastal také jev K_1**.

V našem příkladu $P(K_2)$ má jinou hodnotu než $P(K_2|K_1) = \frac{3}{6} = 0{,}5$, neboť K_2 je situace, kdy druhý vybraný mobil je kvalitní, aniž je řečeno něco o prvním vybraném. Při prvním výběru mohou nastat dvě situace, které obě musíme uvažovat, protože obě hrají roli pro výběr druhého kvalitního. Tedy

$$P(K_2) = P(K_1) \cdot P(K_2|K_1) + P(\overline{K_1}) \cdot P(K_2|\overline{K_1}) = \frac{4}{7} \cdot \frac{3}{6} + \frac{3}{7} \cdot \frac{4}{6} \doteq 0{,}57143.$$

Tedy vidíme, že $P(K_2)$, bez ohledu na to, co se stalo-stane při prvním výběru, je větší než $P(K_2|K_1)$. Jedná se tedy o dva různé koncepty a tomu $P(K_2|K_1)$ se říká podmíněná pst.

Dokončeme výpočet našeho příkladu 3.4:

$$P(K_1) \cdot P(K_2|K_1) \cdot P(N_3|K_1 \cap K_2) = \frac{4}{7} \cdot \frac{3}{6} \cdot \frac{3}{5} = 0{,}17143.$$

Vzorec i označení právě vysvětlené a použité ještě zapišme do věty:

Věta 3.3 *Pokud* $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) \neq 0$*, tak pro průnik náhodných jevů* A_1*,* A_2*, až* A_n *platí:*

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2) \cdot \cdots \cdot P(A_n|A_1 \cap A_2 \cap \cdots \cap A_{n-1}).$$

Pro $n = 2$ přechází věta do vzorce $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2|A_1)$, odkud lze získat logicky ekvivalentní představu o tom, co je to podmíněná pst jevu A_2 za podmínky A_1:

$$P(A_2|A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)}, \tag{3.1}$$

tj. podíl psti průniku $A_2 \cap A_1$ ku psti jevu A_1, který předpokládáme, že nastal. Podmíněné psti v příkladu 3.4 jsme vypočetli pomocí klasického modelu přímo, ale v některých jiných příkladech, kde si nebudeme tak jisti, se nám vzorec 3.1 může hodit, protože platí i v situacích, kdy se nejedná o model klasické psti.

Abychom nyní uzavřeli diskusi nad pstí průniku, vidíme, že existují dva různé vzorce: jednak vzorec z příkladu 3.3 pro nezávislé jevy (výpočet pomocí součinu dílčích pstí), z příkladu 3.4 pro jevy závislé (výpočet pomocí součinu, ve kterém „nabalujeme" u každého dalšího jevu podmínku, že nastaly-nastanou současně i všechny jevy dosud uvažované). Porovnáním těchto dvou vzorců je vidět, například na dvou jevech A_1, A_2, že pokud tyto dva náhodné jevy jsou nezávislé, platí $P(A_2|A_1) = P(A_2)$, tj. podmínka A_1 nezávislá na jevu A_2 neovlivní pst jevu A_2 – ta je pořád stejná, ať už A_1 nastane nebo ne.

Závěrem ještě dva příklady na procvičení řečeného v této kapitole.

Příklad 3.5 *Házíme čtyřstěnem, přičemž PADNE ta strana, na kterou celý čtyřstěn dopadne jako na základnu. Přitom*

stěna A: *obarvena červeně;*

stěna B: *obarvena zeleně;*

stěna C: *obarvena modře;*

stěna D: *rozdělena na tři části trojúhelníky, jeden z nich obarven červeně, druhý zeleně, třetí modře.*

Označme náhodné jevy

R *… padne alespoň část stěny červené ($R = \{A, D\}$);*

G *… padne alespoň část stěny zelené ($G = \{B, D\}$);*

B *… padne alespoň část stěny modré ($B = \{C, D\}$).*

Jsou náhodné jevy R, G, B stochasticky nezávislé? ($\Omega = \{A, B, C, D\}$)

Řešení: Musíme ověřit platnost čtyř rovností:

$$\begin{aligned}
P(R \cap G) = P(R) \cdot P(G): &\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2},\\
P(R \cap B) = P(R) \cdot P(B): &\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2},\\
P(G \cap B) = P(G) \cdot P(B): &\quad \frac{1}{4} = \frac{1}{2} \cdot \frac{1}{2},\\
P(R \cap G \cap B) = P(R) \cdot P(G) \cdot P(B): &\quad \frac{1}{4} \neq \frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{8}.
\end{aligned}$$

Z rovností a nerovnosti plyne: každé dva z jevů R, G, B jsou stochasticky nezávislé, kdežto všechny tři současně jsou stochasticky závislé. Tj. závislost tří jevů je trochu tajemnější a může nastat, i když každá dvojice z daných tří je nezávislá.

Příklad 3.6 *Ze sady 100 výrobků je 10 zmetků. Při kontrole jakosti vybereme náhodně tři výrobky z této sady. Určete pst, že*

a) *první dva vybrané výrobky budou kvalitní a třetí vybraný bude zmetek* $(K_1 \cap K_2 \cap \overline{K_3})$*;*

b) *ze tří vybraných budou dva kvalitní a jeden zmetek (nezáleží na pořadí, ale vypočtěte tak, že všechna možná příznivá pořadí projdete);*

c) *totéž jako (b), ale počítejte jiným způsobem – tak nějak najednou, když nezáleží na pořadí;*

d) *druhý výrobek ze tří vybíraných bude kvalitní.*

Řešení: Použijeme v tomto příkladu vše, co bylo v tomto oddílu řečeno:

Ad a) $P(K_1 \cap K_2 \cap \overline{K_3}) = \frac{90}{100} \cdot \frac{89}{99} \cdot \frac{10}{98} \doteq 0{,}0826.$

Ad b) Pokud nezáleží na pořadí zmetku ze tří kontrolovaných, mohou nastat tři situace, které jedna druhou vylučuje, tj. jejich psti se podle axiomu (P3) sečtou:

$$P(K_1 \cap K_2 \cap \overline{K_3}) + P(K_1 \cap \overline{K_2} \cap K_3) + P(\overline{K_1} \cap K_2 \cap K_3) =$$
$$= \frac{90}{100} \cdot \frac{89}{99} \cdot \frac{10}{98} + \frac{90}{100} \cdot \frac{10}{99} \cdot \frac{89}{98} + \frac{10}{100} \cdot \frac{90}{99} \cdot \frac{89}{98} \doteq 0{,}2478.$$

Ad c) užijeme klasické psti, tedy podílu příznivých případů úkolu b) ku počtu všech možných:

$$P = \frac{\binom{10}{1} \cdot \binom{90}{2}}{\binom{100}{3}} \doteq 0{,}2478.$$

Ad d) Tuto pst jsme už počítali v příkladu 3.4, pouze s jinými hodnotami:

$$P(K_2) = P(K_1) \cdot P(K_2|K_1) + P(\overline{K_1}) \cdot P(K_2|\overline{K_1}) = \frac{90}{100} \cdot \frac{89}{99} + \frac{10}{100} \cdot \frac{90}{99} = 0{,}9$$

(výsledek je paradoxní: bez ohledu na pořadí vybíraného výrobku, pst, že vytáhneme výrobek kvalitní, je pořád stejná a rovná $\frac{90}{100}$ ... pokud ji tedy nepodmíníme žádnou informací-požadavkem o tom, jaká je kvalita dříve vybraných výrobků).

3.1 Shrnutí

Kromě axiomu (P3), který stanovuje-popisuje pravidlo pro výpočet sjednocení disjunktních jevů, se teorie psti snaží rozšířit práci s náhodnými jevy i na jevy, které mají neprázdný průnik. Nejprve jsme se v této kapitole věnovali výpočtu psti sjednocení jevů, pro něž průniky kterýchkoli z nich mohou být neprázdné – tuto situaci popisuje **věta o součtu pstí** 3.1 nebo spíše její rozšíření pro tři a více jevů, 3.2. Tato věta je pro konečné množiny u klasické psti vlastně jen obdobou principu inkluze a exkluze, kterým počítáme počet prvků sjednocení množin (obdobou získanou z principu inkluze a exkluze vydělením celé rovnosti počtem prvků množiny Ω).

Už v těchto větách o součtu se vyskytuje výpočet psti průniku náhodných jevů. Tuto pst průniku lze počítat různými způsoby, například vyjádřit i pst průniku jevů podle vzorce pro klasickou pst, pokud je to možné, a kombinatoricky vyjádřit počet možných konfigurací příznivých a vydělit počtem konfigurací všech. Ovšem často výpočet průniku jevů nelze určit přímo, ale je možné vyjádřit ji jako součin pstí jistých dílčích jevů – buď přímo součin pstí jevů, jejichž průnik počítáme, jak je tomu v příkladu 3.3, nebo jako součin pstí jevů podmíněných určitými situacemi, jako je tomu v příkladě 3.4. Obecně tedy pro pst průniku lze užít vzorec věty 3.3, ten můžeme použít vždy. Vzorec z příkladu 3.3 lze užít jen pro průnik jevů stochasticky nezávislých. Jak je vidět v definici —refnezav-jevy, stochastickou nezávislost definujeme právě pomocí vlastnosti, že pst průniku jakéhokoli počtu z těchto jevů je roven součinu pstí těchto dílčích jevů.

U jevů stochasticky (= náhodně) závislých lze definovat jistou podmíněnou pst pomocí vzorce 3.1, když $P(A_1) \neq 0$ pro jev A_1, o kterém mluvíme jako o podmínce a jehož nastoupení-výskyt je předpokládán při výpočtu $P(A_2|A_1)$. Respektive tuto podmíněnou pst lze definovat u jevů jakýchkoli, jen pro nezávislé jevy zjistíme, že nic nového nezískáme, protože pro ně $P(A_2|A_1) = P(A_2)$.

3.2 Otázky k opakování

U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.

Otázka 3.1 *Podmíněná pst vyjadřuje pst náhodného jevu za předpokladu, že je splněna jistá podmínka.*

Otázka 3.2 *Podmíněná pst $P(B|A)$ nemůže být rovna nule.*

Otázka 3.3 *Pst, že žádný ze čtyř bytů nebude při náhodném zapojení zvonků připojen ke správnému jménu, lze vyjádřit jako $P(\overline{A_1} \cap \overline{A_2} \cap \overline{A_3} \cap \overline{A_4})$.*

Otázka 3.4 *Pro stochasticky nezávislé jevy A, B, pro které $P(A) > 0$, $P(B) > 0$, platí $P(B) = P(B|A)$.*

Otázka 3.5 *Vzorec $P(A_1) \cdot P(A_2) \cdot P(A_3) = P(A_1) \cdot P(A_2|A_1) \cdot P(A_3|A_1 \cap A_2)$ platí jen někdy.*

Otázka 3.6 *Vzorec $P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$ neplatí vždy.*

Otázka 3.7 *Vždy platí $P(A_2|A_1) = \frac{P(A_2 \cap A_1)}{P(A_1)}$, protože to plyne ze vzorce $P(A_1 \cap A_2) = P(A_1) \cdot P(A_2|A_1)$.*

Otázka 3.8 *Jevy R, G, B mohou být stochasticky závislé, i když každé dva z nich jsou stochasticky nezávislé.*

Odpovědi na otázky viz 13.3.

4 Bernoulliovy pravděpodobnosti, úplná pravděpodobnost, Bayesův vzorec

4.1 Bernoulliovy pravděpodobnosti

Uvažujme experiment takové povahy, že mohou nastat jen dva různé výsledky, které se navzájem vylučují (nemůže k nim dojít současně): „úspěch“ a „neúspěch“ („úspěch“ nemusí znamenat nic světoborného; označuje se tímto termínem proto, že se jedná o ten ze dvou možných výsledků, na který se ve svých úvahách chceme zaměřit).

Pravděpodobnost úspěchu je p, pravděpodobnost neúspěchu $1-p$. Náhodná veličina X, která udává počet výskytů úspěchu při N nezávislých opakováních experimentu, má tzv. binomické rozdělení pravděpodobnosti (s parametry N, p) a nabývá hodnot z množiny $\{0, 1, 2, \ldots, N\}$ s pravděpodobností

$$P(X = r) = \binom{N}{r} \cdot p^r \cdot (1-p)^{N-r}.$$

Mluví se zde o nezávislých opakováních experimentu. Slovo „nezávislých“ znamená, že výskyt úspěchu při prvním opakování experimentu nemá vliv na to, zda při druhém a dalších opakováních nastane úspěch nebo ne. Skutečnost, že veličina X má binomické rozdělení s parametry N, p, budeme označovat

$$X \sim Bi(N, p).$$

Podívejme se nyní na konkrétní příklady.

Příklad 4.1 *Házeme čtyřikrát kostkou. Veličina X udává, kolikrát přitom padne šestka. Jaké je rozdělení pravděpodobnosti veličiny X?*

Řešení: Pravděpodobnost, že při jednom hodu padne šestka, je rovna $p = \frac{1}{6}$. Hody jsou navzájem nezávislé, tj. pokud v prvním hodu padla šestka, nemá to vliv na to, zda ve druhém hodu padne nebo ne. Tedy veličina X, která měří počet šestek při čtyřech hodech, má binomické rozdělení pravděpodobnosti s parametry $N = 4$, $p = \frac{1}{6}$.

$$\begin{aligned}
P(X=0) &= P(ne\ 6)\cdot P(ne\ 6)\cdot P(ne\ 6)\cdot P(ne\ 6) = \frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} = 0{,}482;\\
P(X=1) &= P(jednou\ padne\ 6,\ jinak\ něco\ jiného\ než\ 6) =\\
&= P(6\ padne\ jako\ první,\ jinak\ ne) + P(6\ padne\ druhá,\ jinak\ ne) +\\
&\quad + P(6\ padne\ jako\ třetí,\ jinak\ ne) + P(6\ padne\ čtvrtá,\ jinak\ ne) =\\
&= \frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} + \frac{5}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} + \frac{5}{6}\cdot\frac{5}{6}\cdot\frac{1}{6}\cdot\frac{5}{6} + \frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{1}{6} =\\
&= (všechna\ možná\ pořadí\ výskytu\ jednoho\ úspěchu)\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} =\\
&= \binom{4}{1}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} = 0{,}386;
\end{aligned}$$

$$\begin{aligned}
P(X=2) &= P(\textit{dvakrát padne šestka, jinak ne}) = \\
&= (\textit{všechny možnosti výběru 2 pořadí ze 4}) \cdot \frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} = \\
&= \binom{4}{2}\cdot\frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6}\cdot\frac{5}{6} = 0{,}116; \\
P(X=3) &= \binom{4}{3}\cdot\frac{1}{6}\cdot\frac{1}{6}\cdot\frac{1}{6}\cdot\frac{5}{6} = 0{,}015; \\
P(X=4) &= \binom{4}{4}\cdot\left(\frac{1}{6}\right)^4 = 0{,}001.
\end{aligned}$$

Všimněte si, že součet těchto pěti pravděpodobností je roven jedné. Při výpočtu jsme zaokrouhlovali na tři desetinná místa.

Příklad 4.2 *Senátor Swenson před volbami tvrdí, že pro něj bude hlasovat 70% voličů. Agentura STEN chce provést průzkum u 20 lidí. Náhodná veličina X udává počet Swensonových voličů z dvaceti dotázaných. Určete*

a) *teoretické rozdělení veličiny X (před provedením průzkumu);*

b) *pravděpodobnost, že Swensona bude volit přesně 14 lidí z 20 dotázaných, pokud se lidé budou chovat přesně podle předpovědi;*

c) *pravděpodobnost, že Swensona bude volit maximálně 14 lidí z 20 dotázaných, pokud se lidé budou chovat přesně podle předpovědi.*

Řešení:

ad a) *Dané teoretické rozdělení je binomické s parametry $N = 20$ a $p = 0{,}7$. Veličina X nabývá hodnot z množiny $\{0, 1, 2, \ldots, 20\}$ s pravděpodobností*

$$P(X=r) = \binom{20}{r}\cdot 0{,}7^r \cdot 0{,}3^{20-r}.$$

ad b) *Dosazením do vzorce a) máme*

$$P(X=14) = 0{,}192,$$

pokud zaokrouhlujeme na tři desetinná místa.

ad c) *Zde nejlépe: vypočteme pravděpodobnost opačného jevu a odečteme ji od jedničky:*

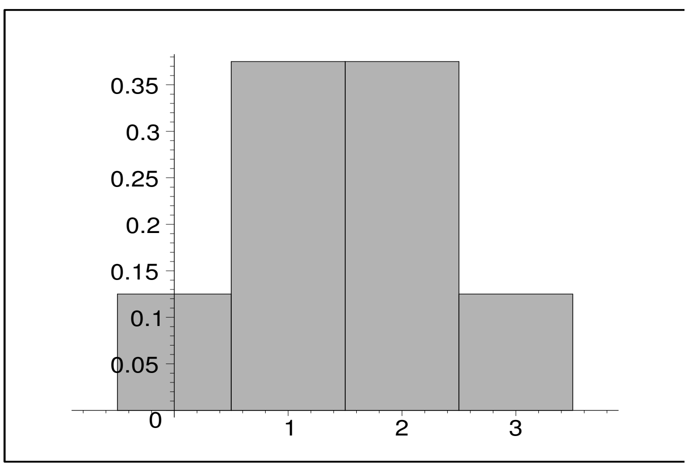
$$\begin{aligned}
P(X \le 14) &= 1 - P(X > 14) = \\
&= 1 - (p(15)+p(16)+p(17)+p(18)+p(19)+p(20)) = \\
&= 1 - (0{,}179 + 0{,}13 + 0{,}072 + 0{,}028 + 0{,}007 + 0{,}001) = 0{,}583.
\end{aligned}$$

Pokud by agentura STEN v předchozím příkladu zjistila, že „pro“ bylo jen 8 lidí z 20, pak některý z teoretických předpokladů nebyl v pořádku:

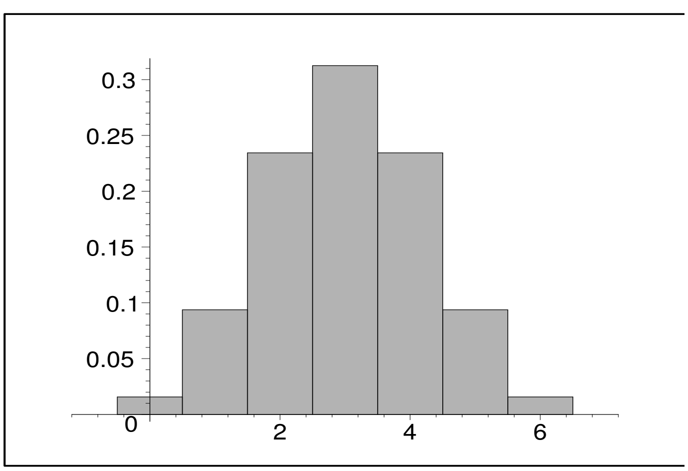
- vzorek dotázaných lidí nebyl náhodný (byl z antiswensonovské oblasti státu);
- odpovědi nebyly nezávislé (odpovídající mezi sebou navzájem diskutovali o Swensonovi);
- STEN pracovala dobře, ale Swenson byl příliš optimistický se svým odhadem (to je nejpravděpodobnější problém).

Ukažme si ještě graficky tvar binomického rozdělení, například pomocí pravděpodobnostního histogramu.

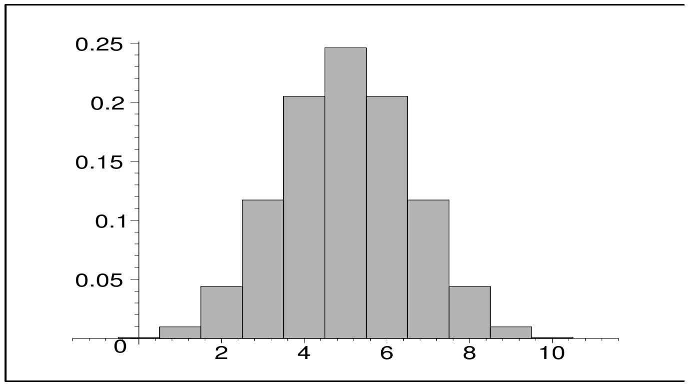
a) Pokud $p = 0{,}5$, rozdělení je vždy symetrické. Například histogram pravděpodobností binomického rozdělení pro $N = 3$, $p = 0{,}5$:



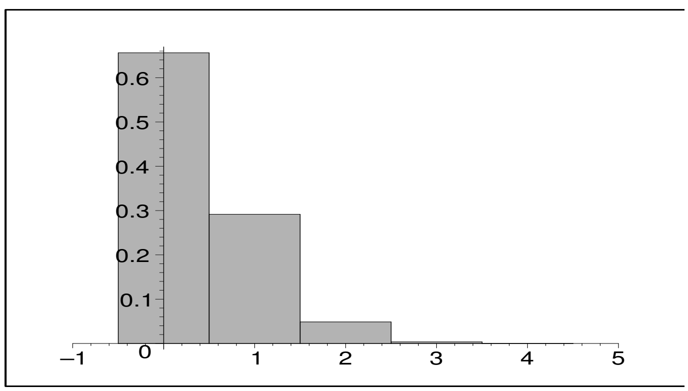
Nebo histogram pravděpodobností binomického rozdělení pro $N = 6$, $p = 0{,}5$:



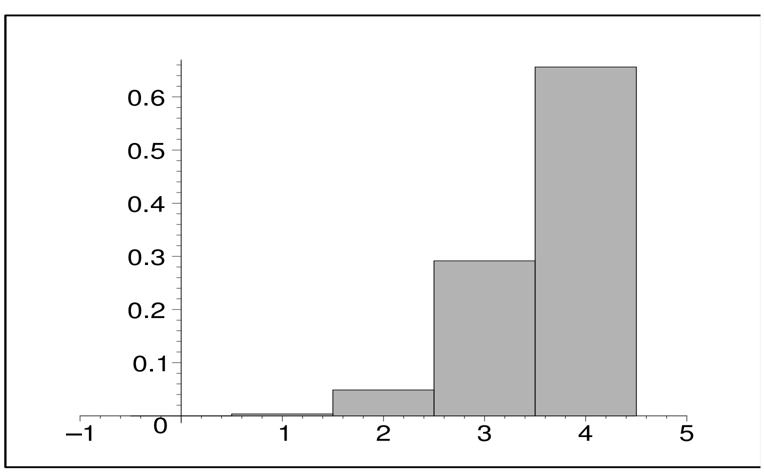
Nebo histogram pravděpodobností binomického rozdělení pro $N = 10$, $p = 0{,}5$:



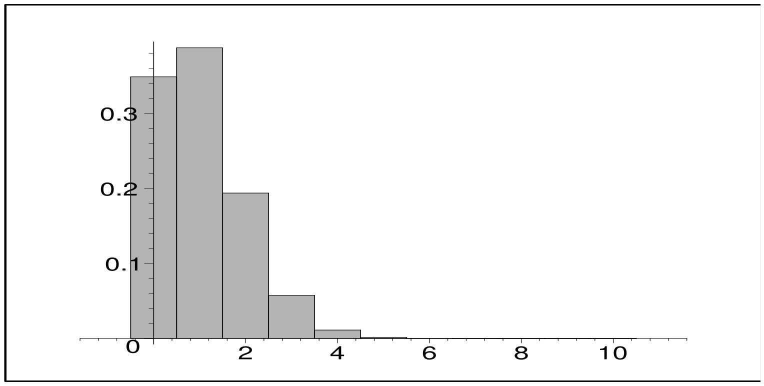
b) Pro $p \neq 0{,}5$ a malé N je rozdělení asymetrické, ale pro rostoucí N se stává více a více symetrickým. Histogram pravděpodobností binomického rozdělení pro $N = 4$, $p = 0{,}1$:



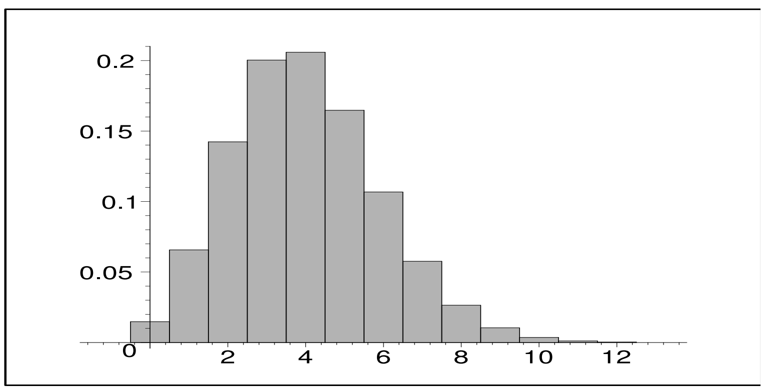
Histogram pravděpodobností binomického rozdělení pro $N = 4$, $p = 0{,}9$:



Histogram pravděpodobností binomického rozdělení pro $N = 10$, $p = 0{,}1$:

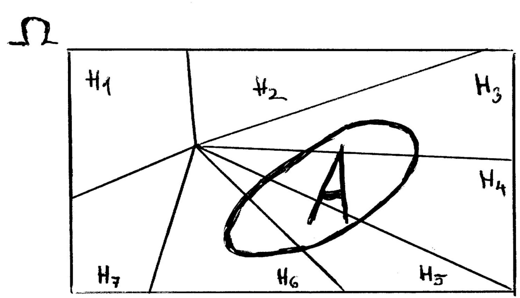


Histogram pravděpodobností binomického rozdělení pro $N = 40$, $p = 0{,}1$:



4.2 Úplná pravděpodobnost

Uvažujme nyní množinu Ω elementárních výsledků experimentu rozloženou (rozklad množiny) na disjunktní podmnožiny H_1, H_2, až H_k – viz obrázek (pro $k = 7$):



Lze vidět z obrázku, že množinu A lze rozložit na sjednocení navzájem disjunktních jevů $H_3 \cap A$, $H_4 \cap A$, $H_5 \cap A$, $H_6 \cap A$. A pak uplatníme axiom (P3) a dostaneme

$$P(A) = P(H_3 \cap A) + P(H_4 \cap A) + P(H_5 \cap A) + P(H_6 \cap A).$$

Užitím vzorce 3.1 pro každý člen na pravé straně dostaneme

$$P(A) = P(H_3) \cdot P(A|H_3) + P(H_4) \cdot P(A|H_4) + P(H_5) \cdot P(A|H_5) + P(H_6) \cdot P(A|H_6).$$

Tomuto právě odvozenému vztahu říkáme věta o úplné psti. Jak je vidět z obrázku, důkaz této rovnosti plyne ze vzájemné disjunktnosti množin $H_3 \cap A$, $H_4 \cap A$, $H_5 \cap A$, $H_6 \cap A$ a z axiomu (P3) o psti sjednocení disjunktních množin.

A ještě poslední věc: kdy tuto větu užijeme? Odpověď je nasnadě: když máme k dispozici (nebo můžeme snadno určit) všechny dílčí psti na pravé straně rovnosti – tehdy $P(A)$ lze určit na základě tohoto vztahu. Tedy pokud neznáme $P(A)$, ale celkem snadno určíme psti $P(A|H_i)$, vzorec se zdarem použijeme.

Věta 4.1 *(Věta o úplné psti) Množinu Ω lze rozložit*[9] *na sjednocení navzájem disjunktních neprázdných podmnožin H_1, H_2, až H_k. A je náhodný jev, tj. $A \subseteq \Omega$. Pak platí:*

$$P(A) = P(H_1) \cdot P(A|H_1) + P(H_2) \cdot P(A|H_2) + \cdots + P(H_k) \cdot P(A|H_k).$$

Příklad 4.3 *V čokoládovně se kompletují bonboniéry na třech výrobních linkách.*

Linka 1: *kompletuje 40% produkce, pokazí (špatně zabalí) 5% bonboniér, které touto linkou procházejí.*

Linka 2: *kompletuje 45% produkce, pokazí (špatně zabalí) 4% bonboniér, které touto linkou procházejí.*

Linka 3: *kompletuje 15% produkce, pokazí (špatně zabalí) 2% bonboniér, které touto linkou procházejí.*

Zkontrolujeme náhodně bonboniéru zabaleno ve skladu (nevíme, ze které linky) – jaká je pst, že nebude v normě (= není správně zabalena)?

Řešení: Když jsou data takhle krásně naservírována, užít věty 4.1 nebude problém – horší budou příklady, kdy tak krásně sestavit data musíme sami při řešení. V každém případě, pokud by se nám nepodařilo najít systém neprázdných podmnožin H_i, který vytváří disjunktní pokrytí množiny Ω, s příkladem bychom se nemohli vypořádat pomocí vzorce na úplnou pst.

Nyní jsou jevy H_1, H_2, H_3 jasné:

H_1 ... náhodně vybraná bonboniéra ze skladu byla vyrobena na lince 1. Odtud $P(H_1) =$ 0,4.

H_2 ... náhodně vybraná bonboniéra ze skladu byla vyrobena na lince 2. Odtud $P(H_2) =$ 0,45.

H_3 ... náhodně vybraná bonboniéra ze skladu byla vyrobena na lince 3. Odtud $P(H_3) =$ 0,15.

[9]Rozklad množiny na podmnožiny: Viz základy matematiky, přednáška 7, rozklad množiny příslušný pojmu relace ekvivalence.

Jako A označíme náhodný jev, jehož pst hledáme: A ... Náhodně vybraná bonboniéra ze skladu je zmetkově zabalena.

Je asi docela srozumitelné, když už vyjádříme výsledek úlohy pomocí

$$P(A) = 0{,}40 \cdot 0{,}05 + 0{,}45 \cdot 0{,}04 + 0{,}15 \cdot 0{,}02 = 0{,}041.$$

Výsledek vlastně udává jakýsi obecný počet procent (4,1%) špatně zabalených bonboniér. A proto se nejedná o úlohu, která by byla příliš vzdálena základní škole – vlastně počítáme jakési procento špatné zabalenosti z celku všech bonboniér.

4.3 Bayesův vzorec

Kdybychom ještě zůstali u stejného obrázku jako v předchozím oddílku, a vlastně i u stejného vzorce 3.1, který jsme využili u úplné psti, lze tento vzorec užít ještě jednou a jinak. Uvažujme tentýž příklad tří linek balicích bonboniéry, i se stejnými údaji o objemech produkce a kvalitách balicích linek – z uvedených dat je možné spočítat ješte psti jiného typu, a sice psti $P(H_1|A)$, $P(H_2|A)$, $P(H_3|A)$.

Aby nedošlo k mýlce, zopakujeme si, co označuje podmíněná pst, ať už oba jevy napíšeme v jakémkoli pořadí:

$P(A|H_1)$ ... pst, že náhodně vybraná bonboniéra ve skladu bude zabalena zmetkovitě za podmínky (= už víme, že nastala situace), že tato bonboniéra byla zabalena na lince 1 ... tuto pst jsme dostali už v zadání příkladu 4.3 a je rovna 0,05.

$P(H_1|A)$ ... pst, že náhodně vybraná bonboniéra pochází z linky 1 za podmínky (= už víme, že nastala situace), že tato bonboniéra byla zabalena nekvalitně. Tato pst v zadání příkladu nebyla uvedena, ale je možné ji vyčíslit pomocí vzorce 3.1, čili

$$P(H_1|A) = \frac{P(H_1 \cap A)}{P(A)}.$$

Pst $P(A) = 0{,}041$ je úplná pst z příkladu 4.3, pst $P(H_1 \cap A)$ vypočteme podle vzorce pro pst průniku:

$$P(H_1|A) = \frac{P(H_1 \cap A)}{P(A)} = \frac{P(H_1) \cdot P(A|H_1)}{P(A)} = \frac{0{,}40 \cdot 0{,}05}{0{,}041} = 0{,}4878.$$

Věta 4.2 *(Bayesův vzorec) Množinu Ω lze rozložit na sjednocení navzájem disjunktních neprázdných podmnožin H_1, H_2, až H_k. A je náhodný jev, tj. $A \subseteq \Omega$. Pak platí:*

$$P(H_i|A) = \frac{P(H_i) \cdot P(A|H_i)}{P(H_1) \cdot P(A|H_1) + P(H_2) \cdot P(A|H_2) + \cdots + P(H_k) \cdot P(A|H_k)}$$

pro každý index $i \in \{1, 2, 3\}$.

Poznámka k tomu, co jsme vlastně spočítali: Pst $P(H_1) = 0{,}40$ je tzv. **apriorní pst** (a priori = předem), neboli pst, jejíž hodnotu známe předem, už v zadání – pst toho, že náhodně zabalená bonboniéra pochází z linky 1.

Pst $P(H_1|A)$ je tzv. **aposteriorní pst** (a posteriori = poté, tj. po měření, po provedení kontroly vybrané bonboniéry), je pst, jejíž hodnotu známe až po provedení nějaké kontroly v balírně bonboniér či ve skladu balírny: náhodně jsme vybrali bonboniéru a zkontrolovali ji a zjistili, že je nekvalitně zabalená. Po této informaci-kontrole se pst změní na $P(H_1|A) = 0{,}4878$ ... pokud kontrolovaná bonboniéra je nekvalitně zabalená, roste šance, že se jedná o bonboniéru balenou na lince 1, protože linka 1 je známá největším procentem zmetkovitého balení, a současně je svým objemem produkce zastoupena celkem silně.

Podobně bychom mohli podle Bayesova vzorce vypočítat aposteriorní psti $P(H_2|A) = 0{,}439$, $P(H_3|A) = 0{,}0732$. Všimněte si v tom například, že apriorní patnáctiprocentní $P(H_3)$ u náhodně kontrolované bonboniéry klesne jen na aposteriorní 7,3%-ní $P(H_3|A)$, protože linka třetí má malou zmetkovitost balení, tj. při špatně zabalené čokoládě klesne šance, že byla balena na lince 3.

Dvě pomůcky, které snad pomohou čtenáři si zapamatovat Bayesův vzorec či zkontrolovat jeho správnost:

- Při výpočtu např. $P(H_1|A)$ se na pravé straně objevuje v čitateli podmíněná pst v opačném pořadí jevů, tj. $P(A|H_1)$ ... tu totiž známe a můžeme ji proto do pravé strany dobře dosadit.

- V čitateli pravé strany se vyskytuje $P(H_1) \cdot P(A|H_1)$, ve jmenovateli také – ve jmenovateli se totiž vyskytuje suma všech možných součinů $P(H_s) \cdot P(A|H_s)$ pro $s \in \{1, 2, \ldots, k\}$. Jinými slovy, v čitateli pravé strany se vyskytuje jeden z členů jmenovatele.

- To už jsme měli zmínit dříve: ve jmenovateli pravé strany je vlastně počítána úplná pst jevu A podle věty o úplné psti.

4.4 Shrnutí

V této kapitole jsme se seznámili s prvním typem rozdělení pravděpodobnosti, které má široké využití v praxi. Veličina X s rozdělením $Bi(N, p)$ nabývá hodnot z množiny $\Omega = \{0, 1, 2, \ldots, N\}$ s pravděpodobností

$$p(k) = P(X = k) = \binom{N}{k} \cdot p^k \cdot (1 - p)^{N-k}. \tag{4.1}$$

Základní témata výpočtu psti rozvíjejí věty 4.1, 4.2. První z nich, větu o úplné pravděpodobnosti, použijeme při možnosti rozložit množinu Ω na několik neprázdných navzájem disjunktních částí H_i takových, že $P(H_i)$ máme k dispozici. Pravděpodobnosti $P(H_i)$ představují jakési váhy, pomocí nichž $P(A)$ pak vypočteme jako vážený průměr pstí $P(A|H_i)$ – a jako u každého dobrého váženého průměru, i pro váhy $P(H_i)$ platí, že jejich součet je roven jedné.

Poslední významnou větu, Bayesův vzorec 4.2, lze využít při výpočtu $P(H_i|A)$ ... tyto „opačné“ podmíněné psti často neznáme, zatímco $P(A|H_i)$ jsou často známy už ze zadání. Pokud dobře víme, že na pořadí podmínky a neznámého jevu při výpočtu podmíněné psti záleží, tak víme, že tyto dvě psti jsou téměř vždy navzájem různé, a jestliže je nezaměníme, tak tu obtížněji vyčíslitelnou právě můžeme vyjádřit pomocí Bayesova vzorce.

Je sice pravda, že pst $P(A \cap H_i)$ v čitateli Bayesova vzorce lze vyjádřit pomocí $P(A) \cdot P(H_i|A)$ i pomocí$P(H_i) \cdot P(A|H_i)$, protože průnik je operace komutativní, tj. větu o psti průniku z kapitoly 3 lze užít „oběma směry“. Ale jen naprostý amatér zde při vybavování vzorce nebo při výpočtu příkladu udělá chybu, neboť $P(H_i|A)$ nemá smysl dosazovat, protože právě to neznáme a chceme určit – v čitateli Bayesova vzorce se použije $P(H_i) \cdot P(A|H_i)$.

4.5 Otázky k opakování

U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.

Otázka 4.1 *Binomické číslo $\binom{N}{k}$ udává, kolika způsoby lze vybrat k prvků z N-prvkové množiny.*

Otázka 4.2 *Pokud $X \sim Bi(N, p)$, tak veličina X může nabývat pouze hodnot z množiny $\{1, 2, \ldots, N\}$.*

Otázka 4.3 *Kromě veličiny X s binomickým rozdělením udávajícím počet výskytů i lze také měřit veličinu $Y = \frac{X}{N}$ relativních četností $\frac{i}{N}$. Přitom platí*

$$P(X = i) = P(Y = \frac{i}{N}).$$

Otázka 4.4 *Věta o úplné psti platí, i pokud existují takové indexy $i, j \in \{1, 2, \ldots, n\}$, že $H_I \cap H_j \neq \emptyset$.*

Otázka 4.5 *Psti $P(H_i)$ ve větě o úplné psti hrají roli tzv. vah, tedy*

$$\sum_1^k P(H_i) = 1.$$

Otázka 4.6 *Psti $P(A|H_i)$ ve větě o úplné psti hrají roli tzv. vah, tedy*

$$\sum_1^k P(A|H_i) = 1.$$

Otázka 4.7 *Pro libovolné náhodné jevy H_i a A platí*

$$P(A) \cdot P(H_i|A) = P(H_i) \cdot P(A|H_i).$$

Otázka 4.8 *Apriorní pstí u Bayesova vzorce je například $P(H_1)$, aposteriorní pstí je například $P(H_1|A)$.*

Otázka 4.9 *Bayesův vzorec bychom mohli teoreticky i upravit na tvar*

$$P(H_i|A) = \frac{P(A) \cdot P(H_i|A)}{P(H_1) \cdot P(A|H_1) + P(H_2) \cdot P(A|H_2) + \cdots + P(H_k) \cdot P(A|H_k)}.$$

Odpovědi na otázky viz 13.4.

5 Procvičování příkladů, příprava na prověrku

Shrnutí a lepší procvičení příkladů z prvních čtyř týdnů semestru na cvičení i přednášce.

6 Prověrka

Prověrka během času přednášky. Počínaje přednáškou 7 bude velmi potřeba, aby studenti absolvovali cvičení až po přednášce, kde bude představena teorie a typy příkladů. To zajistíme tak, že obsah přednášky, kterou bude třeba absolvovat před cvičením 07, probereme na cvičení 06.

Příklady z minulých prověrek byly zahrnuty do textu cvičení, kde najdete tedy i možnost je dobře procvičit.

7 Diskrétní a spojitá náhodná veličina, střední hodnota a rozptyl, distribuční funkce

Viz sken prednaska07.pdf, cviceni07.pdf ... obsáhlé vysvětlení těch základních šesti charakteristik každé náhodné veličiny.

8 Některá význačná rozdělení pravděpodobnosti – diskrétní i spojitá

Viz slajdy 08prednaska, kde najdete přehled základních rozdělení pravděpodobnosti. Obsah přednášky je dobré doplnit i příslušnou kapitolou skript MA08 s využitím Excelu (Fajmon, Nezval, 2021). Podrobnější povídání doporučuji též v textu (Fajmon, Hlavičková, Novák, 2014) – s tím rozdílem, že geometrické rozdělení pravděpodobnosti v tom textu nenajdete, kdežto tam možná najdete navíc představené hypergeometrické rozdělení, které nebylo zařazeno do textu, který čtete.

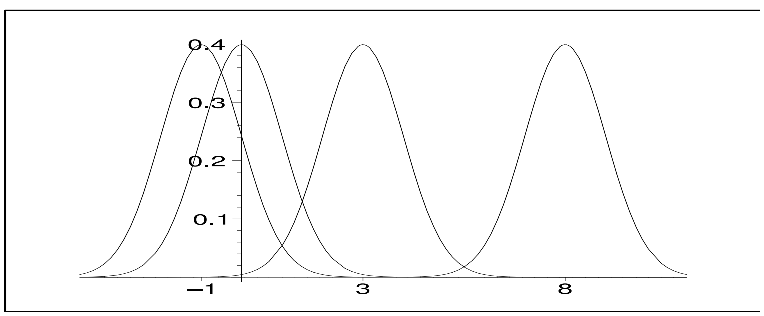
9 Normální rozdělení psti

9.1 Normální rozdělení pravděpodobnosti

Normální rozdělení pravděpodobnosti je rozdělení pro veličiny spojitého typu a má hustotu

$$f(t) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot e^{-\frac{(t-\mu)^2}{2\sigma^2}}.$$

Vzorec této funkce na první pohled nemá příjemný tvar – asi by ji nikdo nechtěl potkat v noci na liduprázdné ulici. Dalo by se spočítat, že střední hodnota veličiny X s rozdělením zadaným touto hustotou je rovna parametru μ, rozptyl je roven parametru σ^2. Proto budeme značit $No(\mu, \sigma^2)$. Na obr. 9.6 jsou uvedeny grafy hustoty pro σ^2 stále rovno jedné a různé střední hodnoty μ, na obr. 9.7 je $\mu = 6$ a mění se hodnoty rozptylu σ^2 (Při malém rozptylu je rameno grafu hustoty vysoké a úzké, pro větší rozptyl hustota nabývá nižších funkčních hodnot, ale interval s hodnotami významně odlišnými od nuly je širší). U všech těchto grafů hustot platí $\int_{-\infty}^{\infty} f(t)dt = 1$.



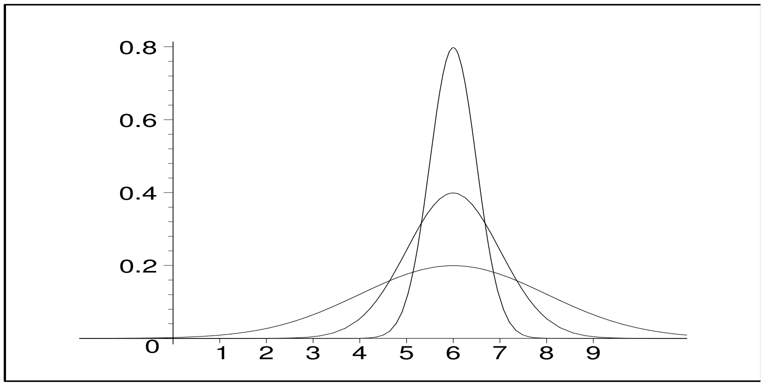
Obrázek 9.6: Hustota normálního rozdělení pro různé střední hodnoty μ.

Normální rozdělení se stalo slavným díky tomu, co říká tzv. **centrální limitní věta**:

Jestliže $X_1, X_2, \ldots, X_N$ jsou navzájem nezávislé veličiny, které mají všechny stejné rozdělení (nemusí být normální, ale libovolné, jeho střední hodnota je $EX_i = \mu$ a rozptyl $DX_i = \sigma^2$), pak součtem těchto veličin je náhodná veličina Y (platí $Y = \sum_1^N X_i$) se střední hodnotou $EY = N \cdot \mu$ a rozptylem $DY = N \cdot \sigma^2$, která má pro dostatečně velké N ($N > 30$) normální rozdělení, tj. platí

$$P(Y \in (a; b)) = \int_a^b \frac{1}{\sqrt{2\pi}\sqrt{N}\sigma} \cdot e^{-\frac{(t-N\mu)^2}{2N\sigma^2}} dt.$$

To, že hodně proměnných lze s velkou přesností popsat pomocí normálního rozdělení, je právě důsledkem centrální limitní věty. Následující dvě situace to dokreslují.



Obrázek 9.7: Hustota normálního rozdělení pro různé rozptyly σ^2.

Příklad 9.1 *Y_1 udává výšku borovic v daném lese (v metrech). Průměrná výška (= μ) je 50 metrů. Vezměme nyní jeden konkrétní strom, jehož výška je 54 metrů. Co způsobilo, že vyrostl o 4 metry nad průměr? Hodně různých vlivů:*

a) *Stromek byl zasazen v obzvlášť příznivém období roku, což způsobilo, že vyrostl o 1m nad průměr.*

b) *Místo, kde strom roste, získává zdroje hnojiva navíc, což vede k růstu o 2,3m nad průměr.*

c) *Nešťastnou náhodou byl stromek při sazení nalomen, což znamená, že narostl o 1,4m nižší, než mohl.*

d) *Strom má dobré místo na slunci, což mu pomohlo vyrůst o 2m nad průměr.*

e) *Skupina příslušníků antagonistického hmyzu si vybrala strom za svůj domov, což mu vzalo šance vyrůst o 0,6m výš než ostatní stromy.*

atd.

Zkrátka a dobře, vychýlení 4m nad průměr je dáno součtem všech těchto možných kladných i záporných vlivů. Protože těchto vlivů je většinou poměrně dost, výslednou výšku stromu danou souštem všech těchto vlivů lze s velkou přesností popsat normálním rodělením.

Příklad 9.2 *Y_2 udává výsledek zkoušky z matematiky. Vezmeme nyní výsledek zkoušky jednoho konkrétního studenta. Co naň mělo vliv?*

a) *Honza měl den před zkouškou chřipku. To snížilo jeho výkon o 5 bodů.*

b) *Honza si něco tipl a náhodou to trefil - přidalo mu to 2 body.*

c) *Honza chyběl na klíčové přednášce a neměl u zkoušky její kopii - přišel o 5 bodů.*

d) *Profesor byl v dobré náladě a při opravování Honzovi 3 body přidal zadarmo.*

atd.

Opět vidíme, že výsledek Honzovy zkoušky je dán součtem většího počtu navzájem nezávislých náhodných vlivů, a tedy jej lze s velkou přesností popsat normálním rozdělením.

Následující příklad by klidně mohl být uveden jako matematická věta, protože se jedná o důležitý důsledek centrální limitní věty (a někdy je také uváděn jako věta - říká se jí Moivre - Laplaceova věta (čti: moávr laplasova)).

Příklad 9.3 Důsledek centrální limitní věty číslo 1. *Specielně i binomické rozdělení lze pro dostatečně velké N dobře popsat (aproximovat, nahradit) normálním rozdělením:*

Uvažujme například veličinu X, která udává počet líců při 100 *hodech korunou. Tato veličina má binomické rozdělení s parametry*

$$N = 100, \;\; p = \frac{1}{2}; \;\; EX = Np = 50; \;\; DX = Np(1-p) = 25.$$

Tuto veličinu lze vyjádřit jako součet veličin $X_1, X_2, \ldots, X_{100}$, kde X_i má binomické rozdělení s parametry $N = 1$, $p = \frac{1}{2}$, tj. udává počet líců v jediném hodu mincí (pro $N = 1$ se binomické rozdělení někdy nazývá alternativní rozdělení, protože veličina může zde nabývat pouze dvou alternativ: 0 *(= číselné vyjádření alternativy „neúspěch“) nebo* 1 *(= číselné vyjádření alternativy „úspěch“)).*

Jako součet stejně rozdělených nezávislých veličin lze tedy X s velkou přesností popsat normálním rozdělením s parametry (pro $N = 100$)

$$\mu = EX = N \cdot EX_i = Np = 50, \quad \sigma^2 = DX = N \cdot DX_i = Np(1-p) = 25.$$

Čili pro dostatečně velké N lze binomické rozdělení s velkou přesností aproximovat normálním rozdělením se stejnou střední hodnotou a rozptylem.

Příklad 9.4 Důsledek centrální limitní věty číslo 2. *Tento důsledek budeme potřebovat v následující kapitole, když se budeme snažit popsat rozdělení průměru ze stejně rozdělených veličin: Jestliže X_1, X_2, ..., X_N jsou stejně rozdělené veličiny, ne nutně normálně rozdělené, a každá z nich má stejnou střední hodnotu μ a stejný rozptyl σ^2, Tak jejich průměr*

$$\overline{X} := \frac{1}{N} \sum_{i=1}^{N} X_i$$

má normální rozdělení se střední hodnotou μ a rozptylem $\frac{\sigma^2}{N}$. Skutečně, lze dokázat výpočtem střední hodnoty a rozptylu této veličiny. Od tvrzení původní centrální limitní věty se tato věta liší pouze tím, že celý součet veličin je ještě vydělený konstantou N, díky tomu tedy jiná střední hodnota a rozptyl:

$$E\left(\frac{1}{N}\sum X_i\right) = \frac{1}{N} \cdot N \cdot \mu = \mu,$$

$$D\left(\frac{1}{N}\sum X_i\right) = \frac{1}{N^2}\cdot N\cdot\sigma^2 = \frac{\sigma^2}{N}$$

(při výpočtu rozptylu využíváme toho, že se počítá jako druhá mocnina odchylky v argumentu – tedy při vytýkání konstanty $\frac{1}{N}$ násobené sumou náhodných veličin před operátor D tuto konstantu musíme umocnit na druhou).

9.2 U-rozdělení

Uvažujme náhodnou veličinu X udávající výsledky zkoušky z matematiky, kterou lze s velkou přesností popsat normálním rozdělením (viz příklad 9.2)s hustotou $f(t)$ a parametry

$$\mu_x = 75, \quad \sigma_x^2 = 25.$$

Její normované hodnoty (viz př. **??**, **??**, **??**) budeme chápat jako hodnoty veličiny U, kde

$$U = \frac{X-\mu_x}{\sigma_x} = \frac{X-75}{5}$$

a platí

$$\begin{aligned} EU &= \int_{-\infty}^{\infty}\frac{t-\mu_x}{\sigma_x}\cdot f(t)dt = \frac{1}{\sigma_x}\left(\int_{-\infty}^{\infty} t\cdot f(t)dt - \mu_x\cdot\int_{-\infty}^{\infty} f(t)dt\right) = \\ &= \frac{1}{\sigma_x}(\mu_x - \mu_x\cdot 1) = 0; \\ DU &= E(U^2) - E^2U = EU^2 - 0 = \int_{-\infty}^{\infty}\left(\frac{t-\mu_x}{\sigma_x}\right)^2\cdot f(t)dt = \\ &= \frac{1}{\sigma_x^2}\int_{-\infty}^{\infty}(t-\mu_x)^2\cdot f(t)dt = \frac{1}{\sigma_x^2}\cdot\sigma_x^2 = 1. \end{aligned}$$

Zajímá-li nás pravděpodobnost, s jakou student dosáhne výsledku mezi 75 a 77 body, musíme spočítat

$$P(75\le X\le 77) = \int_{75}^{77}\frac{1}{\sqrt{2\pi}\cdot 5}\cdot e^{-\frac{(t-75)^2}{50}}dt,$$

což je obsah vyšrafované plochy na obrázku 9.8.

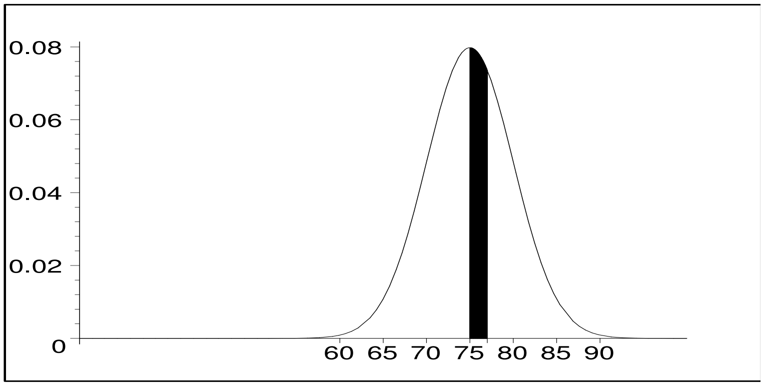
Tato pravděpodobnost je stejná jako pravděpodobnost, že veličina U nabude hodnot z intervalu určeného příslušnými normovanými hodnotami:

$$\begin{aligned} P(75\le X\le 77) &= P\left(\frac{75-75}{5} < \frac{X-75}{5} < \frac{77-75}{5}\right) = \\ &= P(0\le U\le 0.4) = \int_0^{0.4}\frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{u^2}{2}}du, \end{aligned}$$

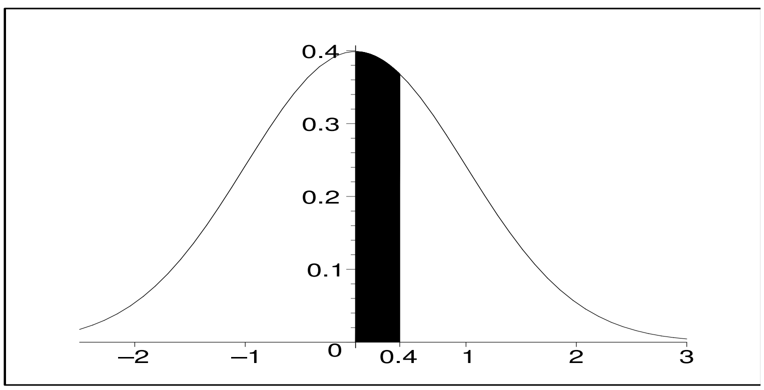
což je obsah šrafované plochy na obrázku 9.9.

Platí tedy

$$\int_{75}^{77}\frac{1}{\sqrt{2\pi}\cdot 5}\cdot e^{-\frac{(t-75)^2}{2\cdot 25}}dt = \int_{\frac{75-75}{5}}^{\frac{77-75}{5}} f(u)du,$$



Obrázek 9.8: Obsah šrafované plochy je roven pravděpodobnosti, že X nabude hodnot z intervalu $< 75; 77 >$.



Obrázek 9.9: Obsah šrafované plochy je roven pravděpodobnosti, že U nabude hodnot z intervalu $< 0; 0.4 >$. Tento obsah je stejný jako obsah šrafované plochy z obr. 9.8.

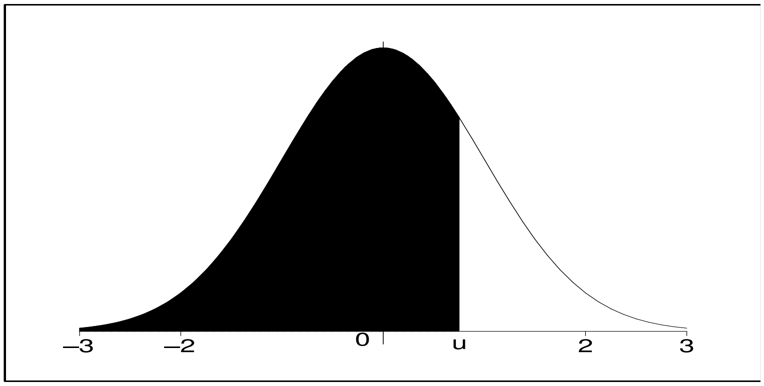
kde $f(u)$ je hustota U-rozdělení, tj. libovolný integrál z hustoty normálního rozdělení lze převést na integrál z hustoty rozdělení U.

Veličina U má tedy normální rozdělení $No(\mu = 0; \sigma^2 = 1)$, které nazýváme **standardizovaným normálním rozdělením** (v anglické literatuře Z-distribution; hodnoty veličiny s tímto rozdělením se nazývají Z-values nebo také Z-scores).

Výpočty uvedených integrálů jsou dosti pracné (buď musíme užít některou z numerických metod, nebo rozvinout exponenciální funkci v nekonečnou řadu a integrovat člen po členu), a proto se s výhodou používá následujícího postupu: pravděpodobnostní výpočty obecného normálního rozdělení se převedou právě popsaným postupem na výpočet integrálu U-rozdělení, pro které byla vypočtena a sestavena tabulka integrálů

$$\Phi(u) = P(U < u) = \int_{-\infty}^{u} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}} dt$$

($\Phi(u)$ je označení distribuční funkce rozdělení U - jako pravděpodobnost má svůj geometrický význam, což znázorňuje obrázek 9.10).



Obrázek 9.10: Obsah šrafované plochy je roven funkční hodnotě distribuční funkce $\Phi(u)$ rozdělení U.

Protože graf funkce $f(u)$ je symetrický vzhledem ke svislé ose (přímce $u = 0$), v tabulce nemusí být uvedeny hodnoty $\Phi(u)$ pro záporná u. Platí totiž pro $u > 0$:

$$\boxed{\Phi(-u) = 1 - \phi(u)}$$

Pravdivost tohoto tvrzení je patrná z toho, že na obou stranách rovnosti v rámečku je obsah téže plochy. Např. $\Phi(-0{,}5) = 1 - \Phi(0{,}5)$, protože (viz obr. 9.11) funkce $f(u)$ je symetrická a celkový obsah plochy pod křivkou je roven jedné:

$$\Phi(-0{,}5) = S(A) = S(B) = 1 - \Phi(0{,}5)$$

Hodnoty funkce $\Phi(u)$ jsou uvedeny v tabulce 9.7 a 9.8.

Příklad 9.5 *Veličinu X udávající výsledek zkoušky lze popsat rozdělením $No(\mu = 75; \sigma^2 = 25)$, S jakou pravděpodobností je výsledek zkoušky*

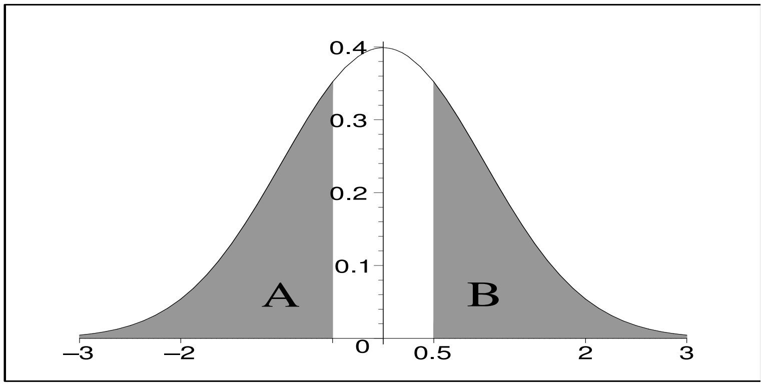
a) *v intervalu* $< 69; 72 >$*?*

b) *menší než* 65*?*

c) *větší než* 80*?*

d) *v intervalu* $< \mu_x - 3\sigma_x; \mu_x + 3\sigma_x >$*?*

Řešení:



Obrázek 9.11: Obsahy ploch A a B jsou stejné.

ad a)

$$
\begin{aligned}
P(69 \leq X \leq 72) &= P\left(\frac{69-\mu_x}{\sigma_x} \leq \frac{X-\mu_x}{\sigma_x} \leq \frac{72-\mu_x}{\sigma_x}\right) = \\
&= P\left(\frac{69-75}{5} \leq U \leq \frac{72-75}{5}\right) = \\
&= P(-1{,}2 \leq U \leq -0{,}6) = \Phi(-0{,}6) - \Phi(-1{,}2) = \\
&= 1 - \Phi(0{,}6) - (1 - \Phi(1{,}2)) = \\
&= \Phi(1{,}2) - \Phi(0{,}6) = 0{,}8849303 - 0{,}7257469 = 0{,}1591834,
\end{aligned}
$$

což je obsah plochy na obrázku 9.12.

Pokud si zvídavý čtenář položil otázku, proč místo některých neostrých nerovností nejsou v tomto odvozování ostré a naopak, pak bych mu rád připomněl, že u spojitých veličin platí

$$P(X = t_0) = 0$$

pro libovolné t_0. Díky tomu nezáleží na tom, zda u normálního rozdělení definujeme distribuční funkci předpisem $F(t) = P(X \leq t)$ nebo $F(t) = P(X < t)$ (tyto dva druhy definice se totiž objevují v matematické literatuře oba, ale žádný velký vliv to nemá – u spojitých veličin to nemá žádný vliv, u diskrétních veličin je schodová distribuční funkce v prvním případě zprava spojitá, ve druhém zleva spojitá, tj. v bodě skoku je v prvním případě funkční hodnota definována na horním schodu, ve druhém případě na dolním).

ad b)

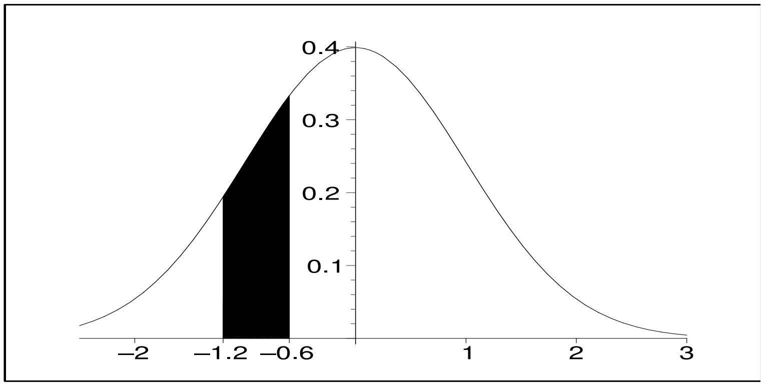
$$
\begin{aligned}
P(X \leq 65) &= P\left(\frac{X-\mu_x}{\sigma_x} \leq \frac{65-\mu_x}{\sigma_x}\right) = P\left(U \leq \frac{65-75}{5}\right) = \\
&= P(U \leq -2) = \Phi(-2) = 1 - \Phi(2) = \\
&= 1 - 0{,}9772499 = 0{,}0227501.
\end{aligned}
$$

Tabulka 9.7: Hodnoty distribuční funkce $\Phi(u)$ - 1.část.

u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$
0,00	0,5000000	0,30	0,6179114	0,60	0,7257469	0,90	0,8159399	1,20	0,8849303
0,01	0,5039894	0,31	0,6217195	0,61	0,7290691	0,91	0,8185887	1,21	0,8868606
0,02	0,5079783	0,32	0,6255158	0,62	0,7323711	0,92	0,8212136	1,22	0,8887676
0,03	0,5119665	0,33	0,6293000	0,63	0,7356527	0,93	0,8238145	1,23	0,8906514
0,04	0,5159534	0,34	0,6330717	0,64	0,7389137	0,94	0,8263912	1,24	0,8925123
0,05	0,5199388	0,35	0,6368307	0,65	0,7421539	0,95	0,8289439	1,25	0,8943502
0,06	0,5239222	0,36	0,6405764	0,66	0,7453731	0,96	0,8314724	1,26	0,8961653
0,07	0,5279032	0,37	0,6443088	0,67	0,7485711	0,97	0,8339768	1,27	0,8979577
0,08	0,5318814	0,38	0,6480273	0,68	0,7517478	0,98	0,8364569	1,28	0,8997274
0,09	0,5358564	0,39	0,6517317	0,69	0,7549029	0,99	0,8389129	1,29	0,9014747
0,10	0,5398278	0,40	0,6554217	0,70	0,7580363	1,00	0,8413447	1,30	0,9031995
0,11	0,5437953	0,41	0,6590970	0,71	0,7611479	1,01	0,8437524	1,31	0,9049021
0,12	0,5477584	0,42	0,6627573	0,72	0,7642375	1,02	0,8461358	1,32	0,9065825
0,13	0,5517168	0,43	0,6664022	0,73	0,7673049	1,03	0,8484950	1,33	0,9082409
0,14	0,5556700	0,44	0,6700314	0,74	0,7703500	1,04	0,8508300	1,34	0,9098773
0,15	0,5596177	0,45	0,6736448	0,75	0,7733726	1,05	0,8531409	1,35	0,9114920
0,16	0,5635595	0,46	0,6772419	0,76	0,7763727	1,06	0,8554277	1,36	0,9130850
0,17	0,5674949	0,47	0,6808225	0,77	0,7793501	1,07	0,8576903	1,37	0,9146565
0,18	0,5714237	0,48	0,6843863	0,78	0,7823046	1,08	0,8599289	1,38	0,9162067
0,19	0,5753454	0,49	0,6879331	0,79	0,7852361	1,09	0,8621434	1,39	0,9177356
0,20	0,5792597	0,50	0,6914625	0,80	0,7881446	1,10	0,8643339	1,40	0,9192433
0,21	0,5831662	0,51	0,6949743	0,81	0,7910299	1,11	0,8665005	1,41	0,9207302
0,22	0,5870604	0,52	0,6984682	0,82	0,7938919	1,12	0,8686431	1,42	0,9221962
0,23	0,5909541	0,53	0,7019440	0,83	0,7967306	1,13	0,8707619	1,43	0,9236415
0,24	0,5948349	0,54	0,7054015	0,84	0,7995458	1,14	0,8728568	1,44	0,9250663
0,25	0,5987063	0,55	0,7088403	0,85	0,8023375	1,15	0,8749281	1,45	0,9264707
0,26	0,6025681	0,56	0,7122603	0,86	0,8051055	1,16	0,8769756	1,46	0,9278550
0,27	0,6064199	0,57	0,7156612	0,87	0,8078498	1,17	0,8789995	1,47	0,9292191
0,28	0,6102612	0,58	0,7190427	0,88	0,8105703	1,18	0,8809999	1,48	0,9305634
0,29	0,6140919	0,59	0,7224047	0,89	0,8132671	1,19	0,8829768	1,49	0,9318879

Tabulka 9.8: **Hodnoty distribuční funkce** $\Phi(u)$ **- 2.část.**

u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$	u	$\Phi(u)$
1,50	0,9331928	1,80	0,9640697	2,10	0,9821356	2,40	0,9918025	4,50	0,9999966
1,51	0,9344783	1,81	0,9648521	2,11	0,9825708	2,41	0,9920237	5,00	0,9999997
1,52	0,9357445	1,82	0,9656205	2,12	0,9829970	2,42	0,9922397	5,50	0,9999999
1,53	0,9369916	1,83	0,9663750	2,13	0,9834142	2,43	0,9924506		
1,54	0,9382198	1,84	0,9671159	2,14	0,9838226	2,44	0,9926564		
1,55	0,9394392	1,85	0,9678432	2,15	0,9842224	2,45	0,9928572		
1,56	0,9406201	1,86	0,9685572	2,16	0,9846137	2,46	0,9930531		
1,57	0,9417924	1,87	0,9692581	2,17	0,9849966	2,47	0,9932443		
1,58	0,9429466	1,88	0,9699460	2,18	0,9853713	2,48	0,9934309		
1,59	0,9440826	1,89	0,9706210	2,19	0,9857379	2,49	0,9936128		
1,60	0,9452007	1,90	0,9712834	2,20	0,9860966	2,50	0,9937903		
1,61	0,9463011	1,91	0,9719334	2,21	0,9864474	2,51	0,9939634		
1,62	0,9473839	1,92	0,9725711	2,22	0,9867906	2,52	0,9941323		
1,63	0,9484493	1,93	0,9731966	2,23	0,9871263	2,53	0,9942969		
1,64	0,9494974	1,94	0,9738102	2,24	0,9874545	2,54	0,9944574		
1,65	0,9505285	1,95	0,9744119	2,25	0,9877755	2,55	0,9946139		
1,66	0,9515428	1,96	0,9750021	2,26	0,9880894	2,56	0,9947664		
1,67	0,9525403	1,97	0,9755808	2,27	0,9883962	2,57	0,9949151		
1,68	0,9535213	1,98	0,9761482	2,28	0,9886962	2,58	0,9950600		
1,69	0,9544860	1,99	0,9767045	2,29	0,9889893	2,59	0,9952012		
1,70	0,9554345	2,00	0,9772499	2,30	0,9892759	2,60	0,9953388		
1,71	0,9563671	2,01	0,9777844	2,31	0,9895559	2,70	0,9965330		
1,72	0,9572838	2,02	0,9783083	2,32	0,9898296	2,80	0,9974449		
1,73	0,9581849	2,03	0,9788217	2,33	0,9900969	2,90	0,9981342		
1,74	0,9590705	2,04	0,9793248	2,34	0,9903581	3,00	0,9986501		
1,75	0,9599408	2,05	0,9798178	2,35	0,9906133	3,20	0,9993129		
1,76	0,9607961	2,06	0,9803007	2,36	0,9908625	3,40	0,9996631		
1,77	0,9616364	2,07	0,9807738	2,37	0,9911060	3,60	0,9998409		
1,78	0,9624620	2,08	0,9812372	2,38	0,9913437	3,80	0,9999277		
1,79	0,9632730	2,09	0,9816911	2,39	0,9915758	4,00	0,9999683		



Obrázek 9.12: K př. 9.5a) - výpočet pravděpodobnosti u normálního rozdělení je roven obsahu šrafované plochy.

ad c)

$$\begin{aligned} P(X \geq 80) &= P\left(U \geq \frac{80-75}{5}\right) = P(U \geq 1) = 1 - P(U < 1) = \\ &= 1 - \Phi(1) = 1 - 0{,}8413447 = 0{,}1586553. \end{aligned}$$

ad d) $P(\mu_x - 3\sigma_x \leq X \leq \mu_x + 3\sigma_x) =$

$$\begin{aligned} &= P\left(\frac{\mu_x - 3\sigma_x - \mu_x}{\sigma_x} \leq U \leq \frac{\mu_x + 3\sigma_x - \mu_x}{\sigma_x}\right) \\ &= P(-3 \leq U \leq 3) = \Phi(3) - \Phi(-3) = \Phi(3) - (1 - \Phi(3)) = \\ &= 2\Phi(3) - 1 = 0{,}9973002 \end{aligned}$$

Většina hodnot veličiny X leží tedy v intervalu $< \mu_x - 3\sigma_x, \mu_x + 3\sigma_x >$. Veličina X nabude hodnoty z tohoto intervalu s pravděpodobností 99,7% (= tzv. **pravidlo tří sigma**).

Příklad 9.6 *Firma vyrábí balíčky ořechů po 200ks, přičemž $\frac{3}{4}$ oříšků jsou burské a $\frac{1}{4}$ lískové, dokonale se promíchají, a pak se teprve sypou do balíčků. Jestliže koupíme jeden balíček ořechů, jaká je pravděpodobnost, že počet lískových ořechů je v intervalu* $< 47; 56 >$*?*

Řešení. Náhodná veličina X udávající počet lískových ořechů v jednom balíčku má rozdělení $Bi(N = 200, p = 0{,}25)$, čili $\mu_x = 50$, $\sigma_x^2 = 37{,}5$. Přímý výpočet

$$\begin{aligned} P &\quad (47 \leq X \leq 56) = P(X = 47) + P(X = 48) + \cdots + P(X = 56) = \\ = &\quad \binom{200}{47} 0{,}25^{47} 0{,}75^{153} + \binom{200}{48} 0{,}25^{48} 0{,}75^{152} + \cdots + \binom{200}{56} 0{,}25^{56} 0{,}75^{144} = \\ = &\quad 0,568 \end{aligned}$$

byl určen pomocí robustní kalkulačky, která má funkci pro obecnou sumu a také funkci pro vyčíslení kombinačních čísel. Při náhradě daného binomického rozdělení normálním rozdělením se stejnou střední hodnotou a rozptylem ($\sigma_x^2 = 37{,}5 \Longrightarrow \sigma_x \doteq 6{,}12$) dostaneme výsledek:

$$\begin{aligned} P(47 \leq X \leq 56) &= P\left(\frac{47-50}{6{,}12} \leq U \leq \frac{56-50}{6{,}12}\right) = \Phi(0{,}98) - \Phi(-0{,}49) = \\ &= \Phi(0{,}98) - (1 - \Phi(0{,}49)) \doteq 0{,}524. \end{aligned}$$

Je vidět, že chyba od přesného výsledku je v řádu procent (druhé desetinné místo). Pokud bychom použili korekce (viz následující příklad 9.7), dostali bychom výsledek $P(46,5 \leq X \leq 56,5) = 0,569$, jehož odchylka od přesného výsledku je v řádu desetin procenta (třetí desetinné místo).

9.3 Aproximace binomického rozdělení normálním s korekcí

Příklad 9.7 *Náhodná veličina X udává počet líců při čtyřech hodech mincí. Vypočteme například pravděpodobnost, že počet líců ve čtyřech hodech bude jeden nebo dva,*

a) *pomocí $Bi(N = 4, p = 0{,}5)$;*

b) *pomocí normálního rozdělení;*

c) *pomocí normálního rozdělení s korekcí.*

Řešení:

ad a) $P(1 \leq X \leq 2) = p_1 + p_2 = 0{,}25 + 0{,}375 = 0{,}625.$

ad b) *Aproximujme binomické rozdělení normálním rozdělením $No(\mu_x = Np = 2, \sigma_x^2 = Np(1-p) = 1)$:*

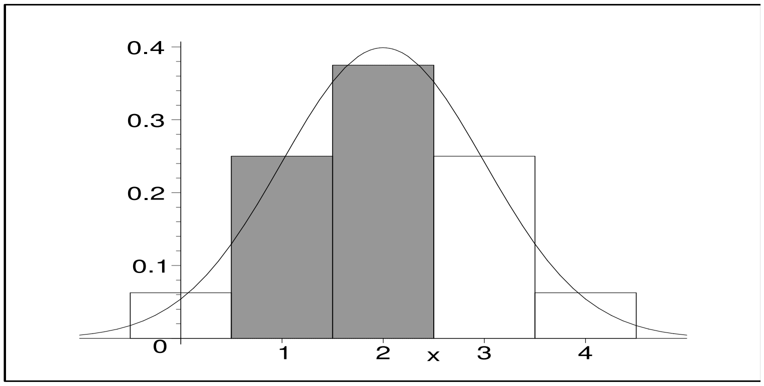
$$\begin{aligned} P(1 \leq X \leq 2) &= P\left(\frac{1-\mu_x}{\sigma_x} \leq U \leq \frac{2-\mu_x}{\sigma_x}\right) = P\left(\frac{1-2}{1} \leq U \leq \frac{2-2}{1}\right) = \\ &= \Phi(0) - \Phi(-1) = 0{,}341. \end{aligned}$$

Hodnota z b) se od hodnoty z a) významně liší!! Kde se udála tak velká chyba? V tom, že obsah plochy dvou obdélníků histogramu na obr.9.13

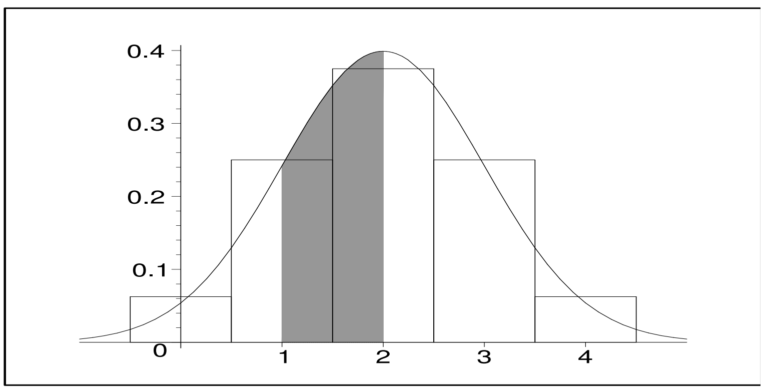
jsme aproximovali pomocí obsahu plochy na obr. 9.14,

nikoliv pomocí šrafované plochy na obr. 9.15.

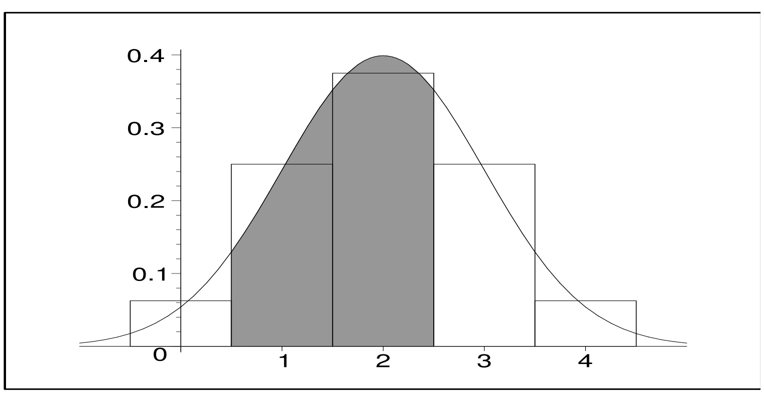
Aproximační chyba se zmenší, pokud výpočet pravděpodobnosti $P(t_1 \leq X \leq t_2)$ pomocí Bi nahradíme obsahem podgrafu hustoty No na intervalu stejné délky, tj. pravděpodobností $P(t_1 - 0{,}5 \leq X \leq t_2 + 0{,}5)$. Toto rozšíření intervalu o 0,5 na obou stranách nazýváme **korekcí**.



Obrázek 9.13: K př. 9.7 - aproximovaná plocha.



Obrázek 9.14: K př. 9.7 - nevhodná aproximace *Bi* pomocí *No*.



Obrázek 9.15: K př. 9.7 - vhodná aproximace *Bi* pomocí *No* užitím korekce.

ad c) *V našem příkladu dostaneme užitím korekce:*

$$\begin{aligned} P(1-0{,}5 \leq X \leq 2+0{,}5) &= P\left(\frac{1-0{,}5-2}{1} \leq U \leq \frac{2+0{,}5-2}{1}\right) = \\ &= \Phi(\frac{1}{2}) - \Phi(-\frac{3}{2}) = 0{,}624, \end{aligned}$$

což je docela dobrá aproximace přesné hodnoty 0,625.

Je vidět, že pomocí korekce lze popsat binomické rozdělení normálním i pro malá N.

10 Úvod do úsudkové statistiky – statistické testy

10.1 Statistické rozhodování – chyba 1. druhu a chyba 2. druhu

Příklad 10.1 Soudní proces jako příklad rozhodovacího procesu. *Uvažujme jednoduchý soudní proces, ve kterém existuje pouze jediný možný trest a soud rozhodne, zda se tomuto trestu obžalovaný podrobí nebo ne. A navíc proti rozhodnutí soudu neexistuje žádné odvolání. Jedná se o jakýsi rozhodovací proces, u kterého mohou nastat čtyři možné výsledky:*

1. *Obžalovaný je vinen a soud jej odsoudí.*
2. *Obžalovaný je nevinen a soud jej osvobodí.*
3. *Obžalovaný je nevinen a soud jej odsoudí. Jedná se o chybné rozhodnutí - tuto chybu budeme označovat jako chybu prvního druhu.*
4. *Obžalovaný je vinen a soud jej osvobodí. Toto rozhodnutí je rovněž chybné - budeme tuto chybu označovat chybou druhého druhu.*

V každém soudním procesu se musí hledat jistá rovnováha mezi tvrdostí a mírností. Jedním extrémem je liberální soudce, který k usvědčení obžalovaného vyžaduje velké množství důkazů. Takový soudce jen zřídka odsoudí nevinného (zřídka se dopustí chyby prvního druhu), ale dosti často osvobodí viníka (chyba druhého druhu). Druhým extrémem je konzervativní soudce, kterému k usvědčení stačí jen několik důkazů. Takový soudce posílá do vězení i jen při stínu podezření, čili častěji odsoudí nevinného (chyba prvního druhu), ale zřídka osvobodí darebáka (= zřídka se dopustí chyby druhého druhu). Slova „konzervativní“ a „liberální“ jsou termíny z politiky. V dnešní době už nikdo neví, co znamenají. Tato jejich „statistická“ definice navrhuje jejich význam, ale také upozorňuje na nebezpečí každého z těchto postojů.

Je otázkou, která z chyb je závažnější - zda chyba prvního druhu, nebo chyba druhého druhu. Všeobecně se má za to, že závažnější je uvěznit nevinného, než osvobodit darebáka. A proto se chybě odsouzení nevinného přisuzuje druh číslo 1 *a věnuje se jí větší pozornost. Ale někde musí být stanovena jistá hranice, po jejímž překročení už soud přistoupí k rozhodnutí „vinen“ a bez skrupulí člověka potrestá.*

Všimněme si jedné věci, která platí jako obecný princip. Pokud se soudce snaží být benevolentní a odsoudí člověka až po nahromadění velkého množství důkazů (snižuje tím možnost výskytu chyby prvního druhu), současně narůstá nebezpečí, že i když je obžalovaný vinen, potřebné množství důkazů se nenajde a soud jej osvobodí (roste možnost výskytu chyby druhého druhu). Není to nic světoborného, ale už jsme dlouho neměli žádný rámeček, a proto jej aspoň uvnitř příkladu můžeme použít:

Snižováním možnosti výskytu chyby prvního druhu roste možnost výskytu chyby druhého druhu - a naopak: pokud zvyšujeme možnost výskytu chyby prvního druhu, snižuje se možnost výskytu chyby druhého druhu.

Z uvedeného rámečku je vidět, že žádnou z chyb není možné naprosto vyrušit: pokud totiž snižujeme možnost výskytu chyby prvního druhu až téměř na nulu, roste tím možnost výskytu chyby druhého druhu do obludných rozměrů a rozhodnutí učiněná tímto stylem jsou nerozumná, až nemoudrá. Strategií v rozhodovacích procesech tohoto typu je tedy zvolit pravděpodobnost výskytu chyby prvního druhu malou, ale ne příliš malou.

Shrňme předchozí úvahy do pěti kroků, které popisují celý soudní proces:

1. *Stojí proti sobě dvě možná rozhodnutí soudu:*

   $$H_0 \text{ ... } \textit{obžalovaný je nevinen}$$
   $$H_1 \text{ ... } \textit{obžalovaný je vinen}$$

   Soud musí rozhodnout právě jednu z těchto variant a toto rozhodnutí je nezvratné, neexistuje proti němu odvolání.

2. *Vystoupí žalobce, který předloží nashromážděné důkazy pro platnost H_1.*

3. *Vystoupí obhájce a vysvětlí všechny souvislosti za předpokladu, že platí H_0. Snaží se vidět a vysvětlit všechny argumenty obžaloby ve světle toho, že obžalovaný je nevinen.*

4. *Porota soudu se odebere k rokování. Bere v úvahu jak množství důkazů a jejich závažnost, tak i argumenty obhajoby a možnost, že tyto důkazy neznamenají nutně vinu obžalovaného, ale v jeho neprospěch hrají jen náhodou.*

5. *Porota se vrací a vyslovuje svůj verdikt: pokud byla překročena míra závažnosti důkazů pro platnost H_1, obžalovaný je vinen. pokud ne, obžalovaný je osvobozen. Toto rozhodnutí soudu je nezvratné.*

Právě uvedených pět kroků v příkladu 10.1 se vyskytuje v mnoha rozhodovacích procesech, které nazýváme **statistické testy**. Tyto principy platí obecně, vyslovme je tedy obecně, už oproštěni od příkladu soudce a obžalovaného (ovšem analogie se soudním procesem zde existuje velice přímá):

(K1) Statistický test obyčejně rozhoduje o tom, zda platí hypotéza H_0 (tzv. **nulová hypotéza**) nebo H_1 (tzv. **alternativní hypotéza**). Tyto dvě hypotézy přitom stojí ve vzájemném rozporu. Ve většině testů H_0 tvrdí, že jistá veličina *nezávisí* na hodnotách určité další veličiny, kdežto H_1 tvrdí, že naopak *závisí* (pro ty, kdo by si chtěli udržet souvislost mezi statistickým testem a soudním procesem, což doporučuji, pomůcka k zapamatování: H_0 testu říká **ne**závisí, a H_0 soudního procesu **ne**vinen).

(K2) Stanovíme **kritérium** (zpravidla určitou funkci), které ukazuje na míru platnosti alternativní hypotézy H_1 (určuje „závažnost důkazů“ pro H_1). Pak provedeme experiment, ve kterém změříme data potřebná pro dosazení hodnot do našeho kritéria.

(K3) Kritériem bývá jistá funkce, která při různých měřeních nabývá různých hodnot, je to tedy náhodná veličina. Určíme **teoretické rozdělení kritéria** za předpokladu, že platí hypotéza H_0. Jinými slovy, popíšeme vlastnosti kriterijní veličiny ve světle toho, že platí H_0.

(K4) Na základě teoretického rozdělení kriterijní veličiny stanovíme určitý interval hodnot, kam když padne empirická hodnota kritéria, tak nezviklá naše přesvědčení o platnosti H_0, ale eventuelní dopad hodnoty kritéria mimo tento interval nás povede k názoru, že byla překročena jistá **kritická míra**, takže usoudíme, že H_0 neplatí. Kritickou míru zpravidla určujeme tak, aby pravděpodobnost výskytu chyby prvního druhu (tj. že rozhodneme, že H_0 neplatí, když ve skutečnosti H_0 platí) byla dostatečně malá, např rovna 0.05 (to se chyby prvního druhu dopustíme nejvýše v pěti procentech případů), ale ne příliš malá, aby nerostla možnost výskytu chyby druhého druhu (tj. že rozhodneme, že H_0 platí, když ve skutečnosti H_0 neplatí) do nerozumných rozměrů.

(K5) Porovnáme empirickou hodnotu kritéria s kritickou mírou. Pokud je kritická míra překročena (hodnota kritéria leží mimo interval nalezený v bodě 4), zamítáme hypotézu H_0 ve prospěch alternativní hypotézy H_1. Pokud není kritická míra překročena, hypotézu H_0 nezamítáme.

Nyní ještě jednou definice chyby prvního a druhého druhu – pozor, je to důležité, protože je potřeba si tyto pojmy pamatovat nejen v příkladu o soudci, ale také v termínech zamítnutí nebo nezamítnutí H_0:

Tabulka 10.9: Čtyři možné výsledky statistického testu.

	skutečnost: H_0 platí	skutečnost: H_1 platí
rozhodnutí: H_0 nezamítáme	O.K.	chyba 2.druhu
rozhodnutí: H_0 zamítáme	chyba 1.druhu	O.K.

Další standardní označení se používá pro pravděpodobnost výskytu chyby 1.druhu (značí se α) a pravděpodobnost výskytu chyby 2.druhu (značíme β).

10.2 Statistický test střední hodnoty binomického rozdělení

V tomto oddílku se pustíme do prvního typu statistického testu. Celý postup bude vysvětlen na příkladu:

Příklad 10.2 *Exper prodejce tvrdil, že o nový výrobek šamponu bude mít zájem* 20 % *zákazníků. Při průzkumu u* 600 *zákazníků jich o nový šampon projevilo zájem* 135. *Na hladině významnosti* $\alpha = 0{,}05$ *testujte hypotézu* H_0*že zájem o výrobek je zhruba stejný, jak expert předpokládal.*

Řešení: Hladina významnosti $\alpha = 0{,}05$ je právě pst výskytu chyby 1. druhu – tuto pst volíme na začátku, před provedením úsudku. Jedná se o jakési nastavení přísnosti-benevolence našeho úsudku; α se obvykle volí $\leq 0{,}05$. Projdeme pět kroků našeho usuzování, jak byly naznačeny v předchozím oddílku: nejprve ovšem provedeme nejdůležitější krok našeho postupu – označíme X = počet zájemců o nový šampón z každých nezávisle dotázaných 600 lidí.

(K1) Budeme rozhodovat mezi hypotézami

- H_0 (= nulová hypotéza): odhad experta se shoduje zhruba s měřením, a tedy řečeno pomocí střední hodnoty: $EX = 120$.
- H_1 (= alternativní hypotéza): skutečný zájem o šampon se od odhadu experta liší, tj. $EX \neq 120$.

Je okamžitě vidět, že v anketě $X = 135$, tj. zájem je „trochu větší než" dvacet procent ze 600 lidí, otázkou ovšem je kritická mez: můžeme pořád tvrdit, že zájem je zhruba 20 procent? Je-není průměrný zájem významně větší než 20 procent? O tom rozhodne tento statistický test.

(K2) Kritériem našeho rozhodování je veličina X = počet zájemců o nový výrobek ze 600 lidí.

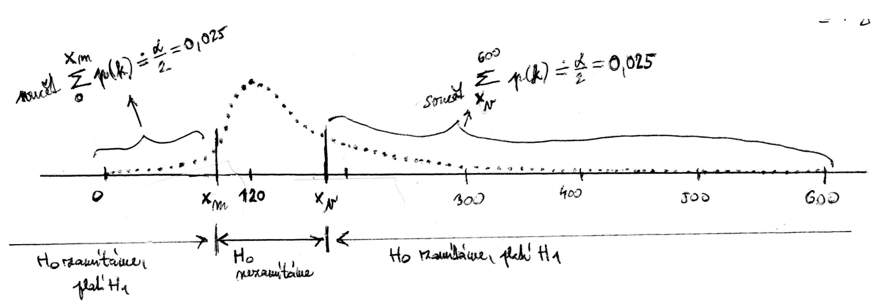
V našem jediném měření je $X = 135$, ale bude tomu tak vždy? Je průměrný zájem skutečně větší než 120 ze 600 lidí?

(K3) Popišme chování veličiny X za předpokladu, že platí H_0 – pokud platí H_0, veličina X má zhruba rozdělení $Bi(N = 600, p = 0{,}20)$.

To znamená, že X může nabývat hodnot $0, 1, 2, \ldots, 600$ s pstí

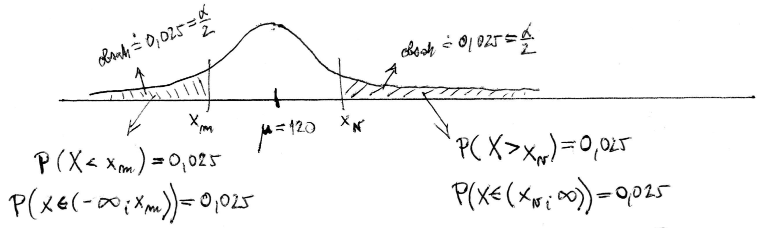
$$p(k) = \binom{600}{k} \cdot 0{,}2^k \cdot 0{,}8^{600-k}.$$

(K4) Pro předem zvolené $\alpha = 0{,}05$ najdeme kritické meze našeho rozhodování: Když počet zájemců bude podstatně větší nebo podstatně menší než 120, zamítneme H_0 a prohlásíme, že platí H_1. Vyjdeme z grafu pstní funkce $p(k)$: máme zde 601 pstí, jejichž součet je roven jedné. „Usekneme" na obou stranách tolik hodnot k, aby pst, že naměřená hodnota veličiny X ležela v intervalu $\langle x_m; x_v \rangle$, se rovnala hodnotě $(1 - \alpha)$, tedy 0,95:



Nyní máme dvě možnosti: a) pracovat s binomickým rozdělení – při hledání x_v musíme sečíst sumu asi čtyř set hodnot z obrázku, při hledání x_m asi sto hodnot. Počítači je to jedno, najde x_m jako $0,025$-kvantil dané binomicky rozdělené veličiny, x_v najde jako 0,975-kvantil.

b) Pokud bych neměli k dispozici počítač, ale jen tabulku distribuční funkce U-rozdělení, využijeme skutečnosti z minulé přednášky, že totiž rozdělení binomické lze dobře aproximovat pomocí rozdělení normálního. To provedeme nyní: namísto diskrétních 601 hodnot pstní funkce budeme pracovat se spojitou hustotou, kde najdeme x_m, x_v pomocí integrací, u kterých známe výsledky:



Pro výpočty budeme potřebovat dosadit střední hodnotu binomického rozdělení $Np = 120$ za μ, a rozptyl binomického rozdělení $Np(1-p)) = 96$ dosadit za σ^2. Způsobem popsaným u normálního rozdělení v minulé přednášce převedeme psti na psti vyjádřené rozdělením U a využijeme tabulky:

$$\Phi\left(\frac{x_m - 120}{\sqrt{96}}\right) = 0{,}025 \implies \frac{x_m - 120}{\sqrt{96}} = -1{,}96 \implies x_m = 120 - 1{,}96 \cdot \sqrt{96} \doteq 100{,}8;$$

$$\Phi\left(\frac{x_v - 120}{\sqrt{96}}\right) = 0{,}975 \implies \frac{x_v - 120}{\sqrt{96}} = +1{,}96 \implies x_v = 120 + 1{,}96 \cdot \sqrt{96} \doteq 139{,}2.$$

(K5) Rozhodnutí našeho statistického testu:

$$X = 135 \in (x_m, x_v) = (100{,}8; 139{,}2),$$

tedy H_0 nezamítáme. Zájem o nový výrobek je zatím zhruba u 20% zákazníků. Neprokázalo se, že by zájem o nový výrobek byl statisticky významně jiný než u 20% zákazníků.

10.3 Statistický test střední hodnoty průměru z normálního rozdělení

Příklad 10.3 *Je známo, že počet bodů získaných souhrnně na testech z matematiky v průběhu prvního pololetí maturitního ročníku má normální rozdělení pro* $\mu = 500$ *bodů a směrodatnou odchylku* $\sigma = 100$ *bodů.*

Firma KAPPA vyvinula program INTEL, jehož cílem je zlepšit znalosti matematiky u středoškolských studentů, zejména pak zlepšit výsledky testů v maturitním ročníku.

Chtějí svůj program INTEL otestovat, a proto náhodně vybrali 25 studentů z ČR a program zaslali každému z nich. Po provedení testu z matematiky se ukázalo, že průměr ohodnocení daných 25 studentů je $\overline{x} = 540$. Otázka zní: lze nyní říct, že program INTEL zlepšuje výkon v testu, nebo se jen náhodou vybralo 25 studentů s vyšším výkonnostním průměrem v matematice? Jedná se o „skutečný“ výsledek (= lze jej zobecnit pro celou populaci?), nebo bylo vyššího průměru dosaženo jen díky náhodným faktorům? Tyto otázky nás přivádějí ke statistickému testu, který rozhodne.

(K1) H_0**:** $\mu = 500$ (program intel nemá vliv na zlepšení matematických schopností, tj. střední hodnota bodového ohodnocení testu celé populace studentů i po rozšíření programu všem (celé populaci) zůstane stejná).

H_1**:** $\mu > 500$ (jednostranný test – můžeme předpokládat, že program znalosti matematiky nezhoršuje).

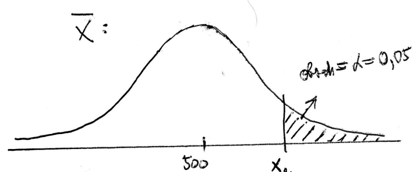
(K2) Kritériem volíme právě veličinu $\overline{X}$, která teoreticky popisuje průměr hodnot (= průměr náhodných naměřených hodnot).

(K3) Za předpokladu platnosti H_0 má veličina $\overline{X}$ parametry

$$\mu_{\overline{X}} = 500, \quad \sigma^2_{\overline{X}} = \frac{\sigma^2}{N} = 400 \Longrightarrow \sigma_{\overline{X}} = 20.$$

(K4) Stanovená kritická U-hodnota je pro $\alpha = 0{,}05$ rovna $u_{0,95} = 1{,}64$. Odtud kritická hodnota v rozměru veličiny $\overline{X}$ je

$$\overline{X}_k = \mu_{\overline{X}} + \sigma_{\overline{X}} \cdot 1{,}64 = 532{,}8;$$



obsah $= \alpha = 0{,}05$ u jednostranného testu „usekáváme“ pouze na té straně od konstanty 500, na kterou „směřuje“ nerovnost:

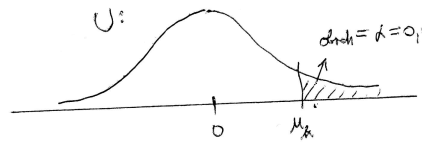
$H_1: E\overline{X} > 500$

$\mathbb{P}(\overline{X} > x_k) = 0{,}05$

$P(\overline{x} \in (x_k; \infty)) = 0{,}05$

$\underbrace{F(\infty)}_{=1} - F(x_k) = 0{,}05$

1 ... v nekonečnu už máme nakumulovány všechny hodnoty veličiny, které lze naměřit



obsah $= \alpha = 0{,}05$

$0{,}95 = F(x_k) = \Phi\Big(\underbrace{\frac{x_k - 500}{20}}_{= u_k}\Big)$

$0{,}95 = \Phi(u_k)$

$\underline{\underline{1{,}64 \doteq u_k}}$

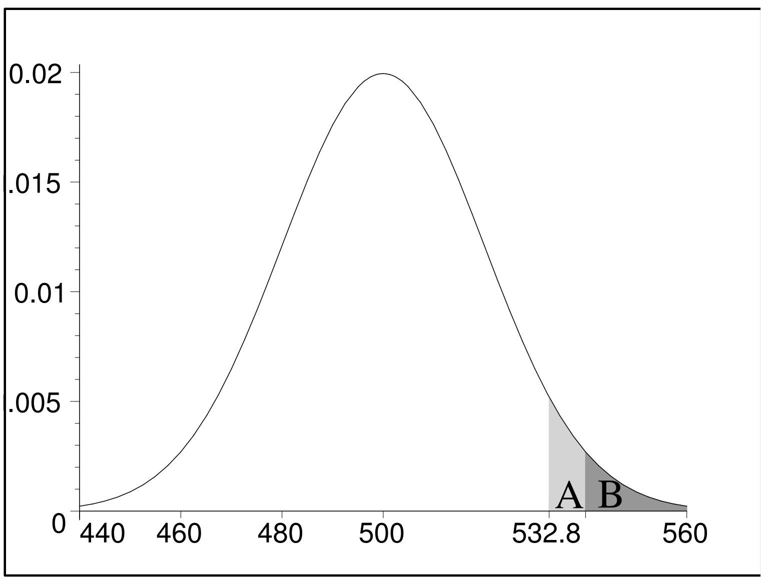
$x_k = 500 + 1{,}64 \cdot 20 \doteq \underline{\underline{532{,}8}}$

(K5) Rozhodnutí testu: pokud příslušná U-hodnota průměru je $\geq 1{,}64$, zamítáme H_0 na hladině významnosti α. V našem případě náhodná veličina $\overline{X}$ nabyla při měření hodnoty $\overline{x} = 540$, tedy příslušná U-hodnota je $u = \frac{540-500}{20} = 2 > 1{,}64$. Proto zamítáme H_0 a uzavíráme, že program „skutečně“ zlepšuje matematické schopnosti studentů.

Poznámka. Souvislost statistického testu s pojmem podmíněné pravděpodobnosti: V průběhu právě dokončeného statistického testu jsme vlastně počítali podmíněnou pravděpodobnost $P(\overline{X} \geq 540|H_0 \text{ platí})$ (čti: pravděpodobnost, že $\overline{X}$ nabude hodnoty větší nebo rovny 540, pokud H_0 platí; tomu, co v uvedeném zápisu následuje za svislou čarou, se říká **podmínka**; **podmíněná pravděpodobnost** je pak pravděpodobnost události zaznamenané před svislou čarou vypočtená za předpokladu, že platí podmínka. Protože $\alpha = 0{,}05 = P(\overline{X} \geq 532{,}8|H_0 \text{ platí})$, je očividné, že

$$P(\overline{X} \geq 540|H_0 \text{ platí}) < \alpha;$$

přesněji (viz obr. 10.16)



Obrázek 10.16: Ad př. 10.3 - hustota rozdělení veličiny $\overline{X}$ za předpokladu, že platí H_0.

$$\alpha = 0{,}05 = P(532{,}8 \leq \overline{X} \leq 540|H_0 \text{ platí}) + P(\overline{X} \geq 540|H_0 \text{ platí}) = S(A) + S(B).$$

Protože podmíněná pravděpodobnost $P(\overline{X} \geq 540|H_0 \text{ platí}) = S(B)$ je menší než naše $\alpha = 0.05 = S(A) + S(B)$, uzavíráme, že něco z našich výchozích předpokladů nebylo správné - to „něco“ je hypotéza H_0. Samozřejmě, že kromě H_0 jsme měli i další výchozí předpoklady, např. naše data mohla být ovlivněna tím, že

a) Náš vzorek 25 studentů nebyl náhodný (byl z výběrových škol).

b) Kolega při opisování dat omylem zapsal některá ohodnocení vyšší než ve skutečnosti.

Ale vlivy typu a),b) mohou být vyloučeny správným naplánováním a provedením měření, takže se v podobných případech většinou uzavírá, že nízká pravděpodobnost $P(\overline{X} \geq 540|H_0$ platí) je důsledkem toho, že nesprávný byl předpoklad platnosti H_0.

Příklad 10.4 *Ředitel firmy KAPPA (data i situace viz předchozí příklad 10.3) zjistil, že konkurenční softwarová firma DELTA rovněž vyvinula program pro výuku matematiky (s názvem KILL). Zavolal si proto svého firemního psychologa a požádal ho, aby zjistil, který z obou konkurenčních programů INTEL a KILL je lepší, tj. který více zvyšuje úroveň matematických znalostí.*

Psycholog získal kopie obou programů. První z nich předal 25 náhodně vybraným studentům, druhou jiným 30 náhodně vybraným studentům. Po provedení testu z matematiky získal od těchto studentů výsledky jejich ohodnocení a spočetl průměry příslušných hodnot. U programu INTEL $\overline{x}_1 = 600$, u programu KILL $\overline{x}_2 = 575$. Aby zjistil, do jaké míry je jeho měření reprezentativní a zda rozdíl průměrů není pouze náhodný (tj. způsobený např. tím, že program INTEL byl rozdán mezi studenty, kteří byli náhodou chytřejší, ale ne tím, že by INTEL byl lepší než KILL), sáhne ke statistickému testu.

(K1) H_0**:** $\mu_1 = \mu_2$ (kdyby se oba programy distribuovaly celé populaci, výsledná střední hodnota ohodnocení by byla u obou stejná).

H_1**:** $\mu_1 \neq \mu_2$ (musíme použít oboustranný test, protože nevíme, který z programů je lepší).

(K2) Testovým kritériem bude rozdíl náhodných veličin $\overline{X}_1 - \overline{X}_2$ s konkrétní naměřenou hodnotou $x_1 - x_2 = 600 - 575 = 25$.

(K3) Za předpokladu platnosti H_0 je rozdělení kritéria $\overline{X_1} - \overline{X_2}$ normální, vypočteme jeho střední hodnotu a rozptyl:

$$E(\overline{X_1} - \overline{X_2}) = E\overline{X_1} - E\overline{X_2} = \mu_1 - \mu_2 = 0,$$

Dále

$$D(\overline{X_1} - \overline{X_2}) \quad = \quad D\overline{X_1} + D\overline{X_2} = D\left(\frac{1}{25} \cdot \sum_1^{25} X_i\right) + D\left(\frac{1}{30} \cdot \sum_1^{30} X_i'\right) = \frac{10000}{25} + \frac{10000}{30} = 733{,}333,$$

a nás bude zajímat směrodatná odchylka $\sqrt{733{,}333} \doteq 27{,}08$. Při výpočtu jsme využili důležitý fakt rozptylu rozdílu dvou nezávislých veličin: rozptyly dílčích veličin sečteme, nikdy je neodečítáme; plyne to z faktu, že rozptyl funguje jako kvadratická odchylka, tedy konstanta (-1) vyjadřující rozdíl se při vytýkání před operátor rozptylu umocní na druhou.

(K4) Pro $\alpha = 0{,}05$ jsou kritické U-hodnoty oboustranného testu stejné jako u oboustranného testu v kapitole **??**: $u_m = -1{,}96$, $u_v = 1{,}96$.

(K5) Rozhodnutí testu: příslušná U-hodnota

$$\frac{\overline{x_1} - \overline{x_2} - 0}{25} = \frac{600 - 575 - 0}{27} = 0{,}92 \in (-1{,}96; 1{,}96) \quad \Longrightarrow \quad \text{NEzamítáme } H_0,$$

programy vyjdou svou kvalitou zhruba nastejno. Nenašlo se dost důkazů pro to, že by oba programy byly odlišné svou kvalitou.

Test v příkladu se liší od předchozího testu pouze krokem (K3), kde jsme museli určit rozdělení rozdílu dvou náhodných veličin.

11 t-test, intervaly spolehlivosti

K přednášce 11 najdete v IS slajdy, ty vás provedou naprostým jádrem – text, kterrý následuje, obsahuje trochu víc materiálu.

11.1 Nestranný a konzistentní odhad parametru rozdělení

Ve statistických testech v minulé kapitole jsme tiše předpokládali, že rozptyl σ^2 je známý. To ale ve skutečnosti většinou není pravda a my jej musíme odhadnout (= přibližně určit). Proto se nyní pustíme do trochy teorie a praxe v odhadování parametrů.

Příklad 11.1 *Pět sad součástek o dvaceti kusech bylo podrobeno zkouškám extrémních teplot. U každé sady je v tabulce uveden počet součástek z daných dvaceti, které v teplotní zkoušce obstály:*

z 20 obstálo x_i	$x_i - \overline{x}$	$(x_i - \overline{x})^2$
13	*0*	*0*
11	*-2*	*4*
12	*-1*	*1*
15	*2*	*4*
14	*1*	*1*

V tabulce už byla využita hodnota průměru $\frac{1}{5}\sum x_i = 13$*. Ve třetím sloupci tabulky jsou uvedeny čtverce odchylek od průměru, odkud spočteme empirický rozptyl (= průměr čtverců odchylek od průměru ... :-)):*

$$s^2 = \frac{1}{5}\sum(x_i - \overline{x})^2 = \frac{10}{5} = 2.$$

Jedná se o měření hodnot náhodné veličiny, kterou je možné popsat střední hodnotou μ *a rozptylem* σ^2*. Ovšem tyto hodnoty neznáme - pokusíme se je odhadnout. Otázka zní: Jak dobrým odhadem pro* μ *je průměr* $\overline{x}$*? Jak dobrým odhadem pro* σ^2 *je empirický rozptyl* s^2*?*

Hodnoty $\overline{x}$, s^2 jsou různé pro různé soubory měření, při jejich popisu užíváme náhodné veličiny $X_1, X_2, \ldots, X_N$ označované jako náhodný výběr. Zde v teorii odhadů je potřeba tento i následující pojmy uvést přesně.

Definice 11.1 *Říkáme, že veličiny* $X_1, X_2, \ldots, X_N$ *tvoří* **náhodný výběr rozsahu** N *z rozdělení pravděpodobnosti o distribuční funkci* $F(x)$*, pokud*

a) *jsou navzájem nezávislé;*

b) *mají stejné rozdělení pravděpodobnosti zadané distribuční funkcí* $F(x)$*.*

Třeba v příkladu 11.1 je $x_1 = 13$, ale stejně dobře jsme mohli naměřit $x_1 = 10$ nebo $x_1 = 17$ – tuto náhodnost prvního měření reprezentuje náhodná veličina X_1, které nepřiřazujeme žádnou konkrétní hodnotu, pouze jsme si vědomi, že pod (velkým písmenem) X_1 se mohou skrývat různé hodnoty. Podobně se mohou skrývat různé hodnoty pod veličinami $X_2, \ldots, X_N$.

Definice 11.2 *Libovolnou funkci* $T_N := T(X_1, X_2, \ldots, X_N)$ *nad náhodným výběrem* X_1, X_2, …, X_N *nazveme* **statistikou**. *Specielně*

a) *Statistiku*

$$\overline{X} = \frac{1}{N} \cdot \sum_{i=1}^{N} X_i \tag{11.1}$$

nazveme **výběrovým průměrem**;

b) *Statistiku*

$$\overline{S^2} = \frac{1}{N-1} \cdot \sum_{i=1}^{N} (X_i - \overline{X})^2 \tag{11.2}$$

nazveme **výběrovým rozptylem**.

Pokud do statistiky T_N dosadíme konkrétní naměřené hodnoty x_1, x_2, …, x_N, dostaneme hodnotu $t_N = t(x_1, x_2, \ldots, x_N)$, která se nazývá **realizací statistiky** T_N.

Například pokud dosadíme do vzorců 11.1, 11.2 konkrétní hodnoty měření x_i, dostaneme realizaci $\overline{x}$ výběrového průměru $\overline{X}$ a realizaci $\overline{s^2}$ výběrového rozptylu $\overline{S^2}$.

No a nyní nás bude zajímat, jak dobrým odhadem neznámé střední hodnoty μ veličiny X je realizace $\overline{x}$ výběrového průměru $\overline{X}$ (respektive jak dobrým odhadem neznámého rozptylu σ jsou realizace s^2, $\overline{s^2}$ veličin S^2, $\overline{S^2}$).

Definice 11.3 *Statistiku* T_N *nazveme* **nestranným odhadem** *parametru* γ, *pokud* $ET_N = \gamma$ *(střední hodnota veličiny* T_N *je rovna hodnotě parametru* γ*).*

Vysvětlení pojmu nestrannosti pomocí konkrétních hodnot měření: pokud budeme opakovaně měřit hodnoty $x_{1,i}$, $x_{2,i}$, …, $x_{N,i}$ a opakovaně počítat realizace $t_{N,i} = t(x_{1,i}, x_{2,i}, \ldots, x_{N,i})$ nestranného odhadu T_N parametru γ pro $i = 1, 2, \ldots, n, \ldots$, bude platit vztah

$$\lim_{n\to\infty} P\left(\left(\frac{1}{n}\sum_{i=1}^{n} t_{N,i}\right) \in (\gamma - \varepsilon; \gamma + \varepsilon)\right) = 1 \tag{11.3}$$

platí pro každé malé pevně zvolené reálné kladné ε.

Laicky řečeno, pokud T_N je nestranným odhadem parametru γ rozdělení veličiny X, tak **pro rostoucí** n **(= rostoucí počet realizací** $t_{N,i}$**)** je průměr realizací $\frac{1}{n}\sum_{i=1}^{n} t_{N,i}$ **skoro jistě (= s pravděpodobností rovnou jedné)** stále blíže hodnotě γ.

Jinými slovy, nestrannost zaručuje takovou konstrukci vzorce pro T_N, která bere v úvahu všechny možné dostupné realizace $t_{N,i}$ a **nestraní** žádné z nich – pro rostoucí počet realizací se aritmetický průměr těchto realizací skoro jistě blíží neznámé hledané hodnotě γ.

Definice 11.4 *Statistiku T_N nazveme* **konzistentním odhadem** *parametru γ, pokud posloupnost náhodných veličin $(T_N)_{N=1}^{\infty}$ konverguje k hodnotě parametru γ podle pravděpodobnosti, tj.*

$$\lim_{N\to\infty} P\left(T_N \in (\gamma-\varepsilon;\gamma+\varepsilon)\right) = 1 \tag{11.4}$$

platí pro každé malé pevně zvolené reálné kladné ε.

Laicky řečeno, pokud T_N je konzistentním odhadem parametru γ rozdělení veličiny X, tak pro rostoucí N (= **rostoucí počet měření pro výpočet jedné realizace**) je hodnota t_N **skoro jistě (= s pravděpodobností rovnou jedné)** stále blíže hodnotě γ.

Jinými slovy, konzistence zaručuje takovou konstrukci vzorce pro T_N, která je **konzistentní (= česky: důsledná)** v tom ohledu, že pro rostoucí počet měření při výpočtu jedné realizace se tato realizace skoro jistě blíží neznámé hledané hodnotě γ.

Uvedené definice nyní osvětlíme konkrétně při hledání odhadu střední hodnoty μ a rozptylu σ^2 veličiny X s normálním rozdělením pravděpodobnosti.

Věta 11.1 *Pokud náhodná veličina X má konečnou střední hodnotu μ, tak výběrový průměr $\overline{X}$ je nestranným a konzistentním odhadem μ.*

Skutečně, v minulé kapitole bylo spočteno, že

a) $\mu_{\overline{X}} = E\overline{X} = E\frac{1}{N}\sum X_i = \mu$, tj. střední hodnota náhodné veličiny $\overline{X}$ je rovna parametru μ; tedy odhad $\overline{X}$ je nestranným odhadem střední hodnoty μ.

b) $\sigma^2_{\overline{X}} = D\overline{X} = D\frac{1}{N}\sum X_i = \frac{\sigma^2}{N}$ a platí

$$\lim_{N\to\infty} D\overline{X} = \lim_{N\to\infty} \frac{\sigma^2}{N} = 0,$$

tj. limita rozptylu odhadu $\overline{X}$ pro rostoucí počet měření je rovna nule – jinými slovy, realizace odhadu $\overline{X}$ se pro rostoucí počet měření skoro jistě blíží hodnotě μ, čili $\overline{X}$ je konzistentním odhadem hodnoty μ.

Čili vzorec pro průměr hodnot $\overline{x}$ funguje přesně tak, jak potřebujeme – „nestraní“ konkrétnímu měření a „skoro jistě“ směřuje přímo k určení střední hodnoty μ, a dále je **konzistentní (= důsledný)** v tom smyslu, že průměr tisíce hodnot je „skoro jistě“ lepší odhadem μ než průměr stovky hodnot.

Otázkou nyní je najít nejvhodnější odhad pro neznámý rozptyl σ^2 veličiny X. Máme k dispozici hodnotu S^2 jako míru vychýlení od průměru – je ona tím nejvhodnějším odhadem hodnoty σ^2?

Věta 11.2 *Pokud náhodná veličina X má konečnou střední hodnotu μ a konečný rozptyl σ^2, tak nestranným a konzistentním odhadem rozptylu σ^2 veličiny X je výběrový rozptyl*

$$\overline{S^2} := \frac{1}{N-1}\sum_{i=1}^{N}(X_i - \overline{X})^2 = \frac{N-1}{N}\cdot S^2.$$

Při vysvětlení obsahu předchozí věty začněme nejprve u odhadu S^2 zopakováním vzorce – při vysvětlování významu rozptylu v popisné statistice bylo (snad) řečeno, že empirický rozptyl s^2 lze vyjádřit buď z definice jako

$$s^2 = \frac{1}{N}\cdot\sum_{i=1}^{N}(x_i - \overline{x})^2, \tag{11.5}$$

nebo po úpravách ve tvaru praktičtějším pro výpočet (= průměr čtverců minus čtverec průměru)

$$s^2 = \left(\frac{1}{N}\sum_{i=1}^{N}x_i^2\right) - \overline{x}^2. \tag{11.6}$$

Při matematickém popisu nyní musíme konkrétní měřené hodnoty ve vzorcích nahradit náhodnými veličinami s velkými písmeny a dostáváme

$$S^2 = \frac{1}{N}\cdot\sum_{i=1}^{N}(X_i - \overline{X})^2, \tag{11.7}$$

respektive

$$S^2 = \left(\frac{1}{N}\sum_{i=1}^{N}X_i^2\right) - \overline{X}^2. \tag{11.8}$$

a) Užijeme nejprve úvah kapitoly předchozí, že totiž $\sigma^2 = EX_i^2 - \mu^2$, a dále platí $\sigma_{\overline{X}}^2 = E\overline{X}^2 - \mu^2$. Vypočtěme střední hodnotu veličiny S^2:

$$\begin{aligned} ES^2 &= E\left(\frac{1}{N}\sum X_i^2 - \overline{X}^2\right) = \frac{1}{N}\left(\sum EX_i^2\right) - E\overline{X}^2 = \\ &= \frac{1}{N}\left(\sum(\sigma^2+\mu^2)\right) - (\sigma_{\overline{X}}^2 + \mu^2) = \\ &= \sigma^2 + \mu^2 - \sigma_{\overline{X}}^2 - \mu^2 = \sigma^2 - \sigma_{\overline{X}}^2 = \sigma^2 - \frac{\sigma^2}{N} = \frac{N-1}{N}\cdot\sigma^2. \end{aligned}$$

Čili střední hodnota statistiky S^2 není rovna odhadovanému parametru σ^2, to znamená, že S^2 není nestranným odhadem rozptylu σ^2. Zkrátka a dobře vzoreček 11.7 není dobře zkonstruován, protože jeho realizace s^2 jsou vždy trochu menší než neznámá hledaná hodnota σ^2

$$\left(s^2 \doteq \frac{N-1}{N}\cdot\sigma^2 < \sigma^2\right),$$

až na patologické případy $\sigma^2 = 0$ a $\sigma^2 = \infty$, které nás nezajímají (matematicky jsou takové náhodné veličiny zkonstruovatelné, ale v praxi měřené veličiny mají vždy konečný kladný rozptyl). Ale k nalezení nestranného odhadu už máme jen krok – můžeme se totiž poučit z výpočtu ES^2. Pokud vynásobíme S^2 konstantou $\frac{N}{N-1}$ (označme novou veličinu $\overline{S^2}$):

$$\overline{S^2} := \frac{N}{N-1} \cdot S^2, \tag{11.9}$$

tak dostaneme

$$E\overline{S^2} = \frac{N}{N-1} \cdot ES^2 = \frac{N}{N-1} \cdot \frac{N-1}{N} \cdot \sigma^2 = \sigma^2,$$

čili $\overline{S^2}$ už je nestranným odhadem hodnoty σ^2.

b) Dá se ukázat, že $\overline{S^2}$ je i konzistentním odhadem rozptylu σ^2, což plyne z faktu, že

$$\lim_{N\to\infty} \sigma^2_{\overline{S^2}} = 0,$$

ale to zde už podrobně provádět nebudeme.

Dále budeme jako nestranný a konzistentní odhad rozptylu σ^2 veličiny X užívat tedy statistiku $\overline{S^2}$. Má tedy ještě nějaký smysl „stará a překonaná“ hodnota S^2? Vrátíme-li se k příkladu 11.1, kde $s^2 = 2$, lze vypočíst $\overline{s^2} = \frac{5}{4} \cdot 2 = 2{,}5$.

1. Hodnota $s^2 = 2$ má svou váhu – vyjadřuje rozptyl souboru uvedených pěti měření. Jedná se o empirický rozptyl – rozptyl naměřených hodnot.
2. Skutečný rozptyl σ^2 veličiny přeživších součástek je větší než rozptyl měření u pěti sad – proto je $\overline{s^2} = 2{,}5$ jeho vhodnějším odhadem.

V příkladu 11.1 můžeme uzavřít, že počet přeživších součástek ze sady dvaceti při extrémní teplotní zátěži lze popsat normálním rozdělením s parametry $\mu \doteq \overline{x} = 13$, $\sigma^2 \doteq \overline{s^2} = 2{,}5$.

Jednoduše řečeno, rozptyl s^2 vypočtený z několika naměřených hodnot je vždy o něco menší než skutečný rozptyl σ^2 měřené veličiny. Proto při odhadu σ^2 musíme vypočtené s^2 „trochu zvětšit“ vynásobením členem $\frac{N}{N-1}$.

Další možný tvar vzorce pro $\overline{S^2}$ lze získat po úpravách s využitím 11.8 a vztahu $\overline{X} = \frac{1}{N} \cdot \sum X_i$:

$$\overline{S^2} = \frac{N}{N-1} \cdot S^2 = \frac{N}{N-1} \cdot \left(\frac{1}{N} \sum X_i^2 - \frac{1}{N^2} \left(\sum X_i \right)^2 \right),$$

a tedy po vykrácení konstantou N a roznásobení závorky dostaneme

$$\overline{S^2} = \left(\frac{1}{N-1} \cdot \sum_{i=1}^{N} X_i^2 \right) - \frac{1}{N(N-1)} \cdot \left(\sum_{i=1}^{N} X_i \right)^2. \tag{11.10}$$

Definice 11.5 *Výraz $ss := \sum (x_i - \overline{x})^2$ budeme nazývat* **součet čtverců** *měření veličiny X.*

Při tomto označení platí

$$S^2 = \frac{1}{N} \cdot SS \tag{11.11}$$

a zejména

$$\overline{S^2} = \frac{1}{N-1} \cdot SS. \tag{11.12}$$

S pojmem součtu čtverců nadále pracuje tzv. analýza rozptylu (ANOVA = ANalysis Of Variance), která se zabývá dalšími statistickými testy zkoumajícími rozptyl (pracuje většinou s tzv. F-rozdělením). Na tu ovšem v tomto kursu už nezbývá prostor.

Podle okolností budeme při výpočtu $\overline{S^2}$ užívat vzorec 11.2, 11.9, 11.10 nebo 11.12.

Zbývá ještě vyjádření k takzvanému počtu stupňů volnosti odhadu.

Příklad 11.2 *Kdybych vám řekl, abyste mi nadiktovali pět reálných čísel, a nedal žádnou další podmínku nebo omezení, mohli byste říct čísla, jaká chcete, například*

$$0, 70, 314, 32, 100.$$

Máte svobodu volby, která čísla vybrat. Jinými slovy, máte pět stupňů volnosti, protože si zcela svobodně a volně vybíráte pět čísel.

Kdyby ale úkol zněl: Nadiktujte mi pět čísel, jejichž průměr je 25, trochu bych vaši volnost omezil – první čtyři čísla byste mohli volit libovolně, například

$$0, 70, 314, 32,$$

ale páté číslo je mým požadavkem jednoznačně určeno. Aby průměr pěti čísel byl 25, jejich součet musí být roven $5 \cdot 25 = 125$, tj. páté číslo musí být rovno $125 - 416 = -291$. Tuto druhou úlohu lze charakterizovat tím, že stupeň její volnosti je 4.

Čili obecně pro N hodnot, u kterých je předem dán jejich průměr, zbývá $N - 1$ stupňů volnosti.

Podobná situace se objevuje i při odhadování rozptylu populace: Uvažujeme-li soubor měření N hodnot jisté veličiny X, z nich lze určit průměr $\overline{x}$. Chceme-li dále odhadnout rozptyl pro tuto konkrétní (už určenou) hodnotu $\overline{x}$, máme už jen $N - 1$ stupňů volnosti (ve vzorci pro $\overline{s^2}$ hodnotu $\overline{x}$ potřebujeme znát, protože při určení $\overline{s^2}$ počítáme totiž míru vychýlení měření právě od hodnoty $\overline{x}$).

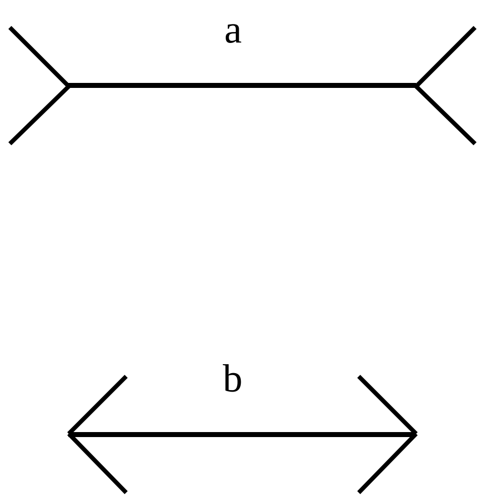
Tedy třeba v příkladu 11.1 má odhad rozptylu $\overline{s^2} = 2{,}5$ čtyři stupně volnosti.

Tento přístup určení počtu stupňů volnosti lze užít i na některé další situace v tomto textu. Obecně platí:

Počet stupňů volnosti odhadu = počet měření minus počet parametrů odhadnutých již dříve.

11.2 t-test typu „μ =konstanta“

Příklad 11.3 *Je známa následující grafická iluze (viz obr. 11.17), že totiž úsečka* **a** *se zdá delší než úsečka* **b** *(počítá se délka bez koncových šipek), i když ve skutečnosti jsou obě úsečky stejně dlouhé.*



Obrázek 11.17: K př. 11.3: Grafická iluze: délky úseček a, b jsou stejné.

Chceme nyní experimentem ověřit, že úsečka typu **a** *se zdá pozorovateli delší sama o sobě, bez porovnání s úsečkou* **b**. *Náhodně vybraným pěti lidem jsme na deset sekund ukázali úsečku* **a** *dlouhou 6 cm. Poté jsme je požádali, aby úsečku dané délky nakreslili, a změřili jsme její délku. Byla získána data $x_1 = 8$, $x_2 = 11$, $x_3 = 9$, $x_4 = 5$, $x_5 = 7$.*

Průměr těchto dat je $\overline{x} = 8$ cm. Chceme testovat hypotézu, že střední hodnota μ celé lidské populace při ohodnocení délky úsečky je významně větší než její skutečná délka 6 cm. V této situaci neznáme rozptyl σ^2 měřené veličiny a musíme jej odhadnout pomocí $\overline{s^2}$. Pak testovacím kritériem nebude

$$\frac{\overline{x} - 6}{\frac{\sigma}{\sqrt{N}}},$$

ale tzv. rozdělení t s hodnotou vypočtenou podle vztahu

$$t := \frac{\overline{x} - 6}{\frac{\overline{S}}{\sqrt{N}}}, \text{ kde } \overline{S} = \sqrt{\overline{S^2}}.$$

Toto t-rozdělení odvodil jistý pan William Sealy Gosset – ovšem příslušný článek uveřejnil nikoli pod svým vlastním jménem, ale pod jménem Student, a rozdělení je tedy známo pod názvem Studentovo t-rozdělení.

Hustota t-rozdělení závisí na počtu ν stupnů volnosti, se kterými se určí odhad rozptylu $\overline{S^2}$, a má tvar

$$f_\nu(x) = \frac{1}{\sqrt{\nu} \cdot \beta(\frac{1}{2}, \frac{\nu}{2}) \cdot \left(1 + \frac{x^2}{\nu}\right)^{\frac{1+\nu}{2}}} = \frac{\Gamma(\frac{1+\nu}{2})}{\sqrt{\pi \cdot \nu} \cdot \Gamma(\frac{\nu}{2}) \cdot \left(1 + \frac{x^2}{\nu}\right)^{\frac{1+\nu}{2}}},$$

podle toho, zda se čtenáři více líbí funkce

$$\beta(p, q) = \int_0^1 u^{1-p} \cdot (1-u)^{1-q} \,\mathrm{du}$$

(tzv. β-funkce), nebo funkce

$$\Gamma(r) = \int_0^\infty u^{r-1} \cdot e^{-u} \,\mathrm{du}$$

(tzv. Γ-funkce). Přechod mezi jednotlivými verzemi vzorce hustoty plyne ze vztahu mezi β-funkcí a Γ-funkcí

$$\beta(p, q) = \frac{\Gamma(p) \cdot \Gamma(q)}{\Gamma(p+q)}$$

a z jedné další drobnosti, že totiž $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ (tato drobnost je vypočtena například v učebnici [**?**], která je více zaměřena na odvozování matematických vztahů v porovnání třeba s učebnicí [**?**]).

Funkce $f_\nu(x)$ je opět jedna z funkcí, kterou by člověk nerad potkal pozdě na ulici při návratu domů, ale ukazuje se, že i takové funkce jsou užitečné.

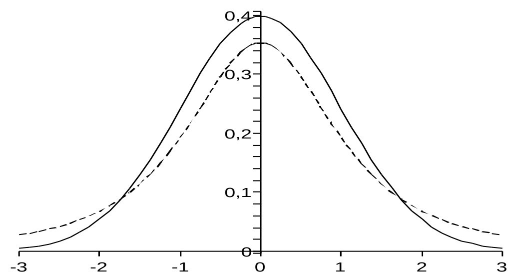
Uveďme zde některé vlastnosti t-rozdělení, které budeme využívat:

a) Hustota f_ν je symetrickou funkcí vzhledem k ose y, tj. je sudá: platí $f_\nu(x) = f_\nu(-x)$.

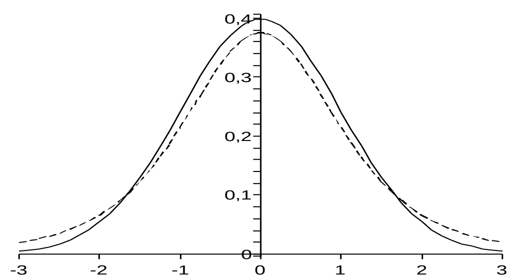
b) Kritická t-hodnota je dále od počátku než kritická U-hodnota, protože t-rozdělení je odvozeno při neznámém rozptylu, tj. zahrnuje větší míru náhodnosti a nejistoty než U-rozdělení (veličina t má větší rozptyl než veličina U). Hustota rozdělení t je „nižší a širší“ než hustota rozdělení U.

c) Čím lepší je náš odhad neznámého rozptylu σ měřené veličiny, tím více se t-rozdělení bude podobat U-rozdělení. A odhad rozptylu bude tím lepší, čím vyšší je počet měření N (tj. čím vyšší je počet stupňů volnosti ν).

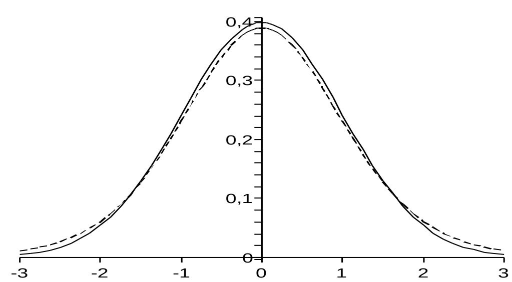
Vlastnosti b), c) lze znázornit graficky porovnáním U-rozdělení s t-rozdělením o různém počtu ν stupňů volnosti – (hustota U-rozdělení je v obrázcích znázorněna plnou čarou, hustota t-rozdělení čárkovanou čarou):



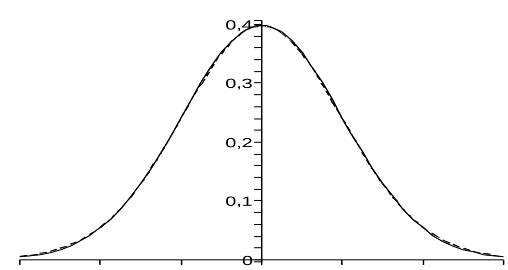
$\nu = 2$



$\nu = 4$



$\nu = 10$



$\nu = 60$

Z obrázků je vidět, že pro rostoucí počet stupňů volnosti se hustota t-rozdělení (v obrázku její graf vyznačen slabě) svým tvarem stále více blíží ke tvaru funkce hustoty U-rozdělení, a pro $\nu = 60$ je hustota t-rozdělení prakticky totožná s hustotou U-rozdělení (omlouvám se za malou tloušťku čar, ale při silnější tloušťce nebyl na obrázku patrný rozdíl mezi čárkovanou a plnou čarou).

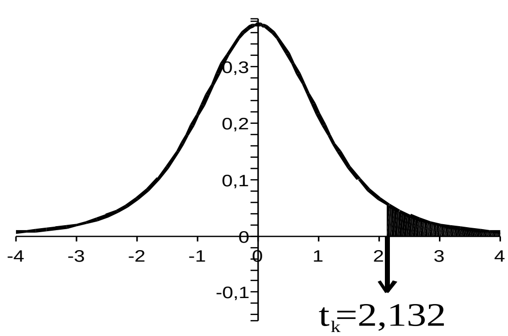
d) Pro určení kritických hodnot t_k budeme potřebovat hodnoty integrálů

$$P_\nu(t \leq x) = F_\nu(x) = \int_{-\infty}^{x} f_\nu(u)\mathrm{d}u, \quad P_\nu(-x \leq t \leq x) = \int_{-x}^{x} f_\nu(u)\mathrm{d}u.$$

Tyto integrály se nepočítají vždy znovu a znovu, poněvadž jejich výpočet je složitý, ale jednou provždy byly spočteny a sestaveny do tabulky. Protože pro jednu hodnotu ν lze sestavit tabulku tak velkou jako je tabulka funkce Φ (BMA3), měli bychom pro 35 různých hodnot ν také 35 různých tabulek. Toto množství dat je zredukováno jen na několik hodnot v závislosti na hladině významnosti α.

A vůbec, pro statistické testy je nejužitečnější místo hodnot distribuční funkce přímo tabulka kritických hodnot pro různé α a ν – jedná se o tabulku 11.10. S touto tabulkou budeme pracovat tak, že vybereme řádek s daným počtem stupňů volnosti ν, sloupec s danou hodnotou $\alpha = q$ u pravostranného ($\alpha = 2q$ u oboustranného) testu. Na průsečíku řádku se sloupcem se pak nachází požadovaná kritická hodnota testu.Tabulku lze volit takto úsporně, protože mezi jednostranným a oboustranným testem je následující vztah:

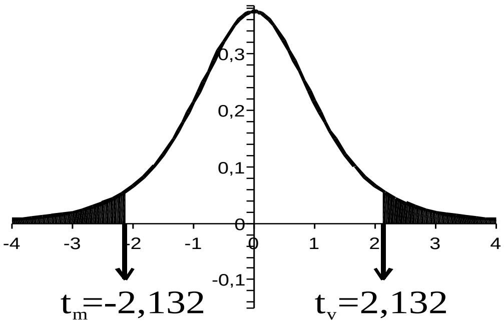
- t_k u levostranného testu se liší od pravostranného pouze znaménkem.
- t_k u pravostraného testu pro $\alpha = q$ je stejné jako $|\pm t_k|$ u oboustranného testu pro $\alpha = 2q$. Je to vidět i na srovnání následujících dvou obrázků:



Pravostranný t-test pro $\alpha = q = 0{,}05$, $\nu = 4$ ... $t_k = 2{,}132$. Šrafovaná část tvoří 5% obsahu celého podgrafu.

Tabulka 11.10: Kritické hodnoty t-testu.

	q=0,4	0,25	0,1	0,05	0,025	0,01	0,005	0,001
ν	2q=0,8	0,5	0,2	0,1	0,05	0,02	0,01	0,002
1	0,325	1,000	3,078	6,314	12,704	31,821	63,657	318,31
2	0,289	0,816	1,886	2,920	4,303	6,965	9,925	22,326
3	0,277	0,765	1,638	2,353	3,182	4,541	5,841	10,213
4	0,271	0,741	1,533	2,132	2,776	3,747	4,604	7,173
5	0,267	0,727	1,476	2,015	2,571	3,365	4,032	5,893
6	0,265	0,718	1,440	1,943	2,447	3,143	3,707	5,208
7	0,263	0,711	1,415	1,895	2,365	2,998	3,499	4,785
8	0,262	0,706	1,397	1,860	2,306	2,896	3,355	4,501
9	0,261	0,703	1,383	1,833	2,262	2,821	3,250	4,297
10	0,260	0,700	1,372	1,812	2,228	2,764	3,169	4,144
11	0,260	0,697	1,363	1,796	2,201	2,718	3,106	4,025
12	0,259	0,695	1,356	1,782	2,179	2,681	3,055	3,930
13	0,259	0,694	1,350	1,771	2,160	2,650	3,012	3,852
14	0,258	0,692	1,345	1,761	2,145	2,624	2,977	3,787
15	0,258	0,691	1,341	1,753	2,131	2,602	2,947	3,733
16	0,258	0,690	1,337	1,746	2,120	2,583	2,921	3,686
17	0,257	0,689	1,333	1,740	2,110	2,567	2,898	3,646
18	0,257	0,688	1,330	1,734	2,101	2,552	2,878	3,610
19	0,257	0,688	1,328	1,729	2,093	2,539	2,861	3,579
20	0,257	0,687	1,325	1,725	2,086	2,528	2,845	3,552
21	0,257	0,686	1,323	1,721	2,080	2,518	2,831	3,527
22	0,256	0,686	1,321	1,717	2,074	2,508	2,819	3,505
23	0,256	0,685	1,319	1,714	2,069	2,500	2,807	3,485
24	0,256	0,685	1,318	1,711	2,064	2,492	2,797	3,467
25	0,256	0,684	1,316	1,708	2,060	2,485	2,787	3,450
26	0,256	0,684	1,315	1,706	2,056	2,479	2,779	3,435
27	0,256	0,684	1,314	1,703	2,052	2,473	2,771	3,421
28	0,256	0,683	1,313	1,701	2,048	2,467	2,763	3,408
29	0,256	0,683	1,311	1,699	2,045	2,462	2,756	3,396
30	0,256	0,683	1,310	1,697	2,042	2,457	2,750	3,385
40	0,255	0,681	1,303	1,684	2,021	2,423	2,704	3,307
60	0,254	0,679	1,296	1,671	2,000	2,390	2,660	3,232
120	0,254	0,677	1,289	1,658	1,98	2,358	2,617	3,160
∞	0,253	0,674	1,282	1,645	1,960	2,326	2,576	3,090



Oboustranný t-test pro $\alpha = 2q = 0{,}1$, $\nu = 4$ ... $t_m = -2{,}132$, $t_v = 2{,}132$.
Šrafovaná část tvoří celkem 10% obsahu celého podgrafu (na každé straně je vyšrafováno 5% obsahu podgrafu).

– Ze symetrie hustoty t-rozdělení je patrné, že pokud budeme provádět levostranný t-test, stačí najít kritickou hodnotu pravostranného testu a změnit její znaménko na záporné.

Vraťme se nyní zpět k příkladu 11.3 a proveďme statistický t-test:

(K1) *H_0: $\mu = 6$ (střední hodnota délky úsečky není ovlivněna iluzí prodloužení).*
H_1: $\mu > 6$.

(K2) *Testovým kritériem bude výběrový průměr $\overline{X}$, jehož realizaci $\overline{x} = 8$ převedeme na t-hodnotu*

$$\frac{8-6}{est\sigma_{\overline{X}}}.$$

(„est“ označuje odhad, z anglického „estimate“ [estimit] – protože v dalším textu budeme odhadovat odchylku i pomocí jiných funkcí než $\overline{X}$, bude výhodné si tímto označením připomenout, u které veličiny vlastně odchylku nebo rozptyl odhadujeme).

(K3) *$\overline{s^2} = 5$, a tedy*

$$est\sigma^2_{\overline{X}} = \frac{\overline{s^2}}{N} = \frac{5}{5} = 1,$$

přičemž počet stupňů volnosti odhadu je $N - 1 = 5 - 1 = 4$, tj. za předpokladu platnosti H_0 má veličina $\frac{\overline{X}-6}{1}$ Studentovo t-rozdělení pro $\nu = 4$.

(K4) *Příslušná kritická hodnota t_k je na průsečíku řádku $\nu = 4$ a sloupce pro $\alpha = 0{,}05$ (u pravostranného testu), tj. $t_k = 2{,}132$.*

(K5) *$\frac{8-6}{1} = 2 < t_k = 2{,}132 \Rightarrow H_0$ nezamítáme, nenašli jsme dostatek důkazů pro potvrzení iluze větší délky.*

11.3 Několik poznámek ke statistickému testu

Poznámka A) Logika formulace „nezamítáme H_0" V právě dokončeném příkladu bylo odpovědí „hypotézu H_0 nezamítáme". Logiku této formulace snad osvětlí následující příklad.

Příklad 11.4 *Když něco někde nenajdu, neznamená to, že to tam není.*

S nástupem lyžařské sezóny pan Loftus začal oprašovat svou výstroj a zjistil, že nemůže najít své lyžařské brýle, i když prohledal celý byt. Potom jej ovšem zděsila představa, že si za dva tisíce bude muset koupit brýle nové, a celý byt prohledal ještě jednou, a to důkladněji. Nakonec brýle našel v zadním rohu své skříně!!

Tento příklad je ilustrací jednoho celkem logického principu: pokud naleznu hledaný předmět, určitě vím, že tam je. Ale pokud jej nenaleznu, může to znamenat buď že tam není, nebo že tam je, ale mé hledání nebylo dost důkladné (nemělo dostatečnou sílu).

Testování hypotéz je také takovým hledáním – hledáme vliv jedné veličiny na druhou veličinu. Pokud nalezneme tento vliv (zamítáme H_0), víme „s jistotou" (na hladině významnosti α), že existuje. Pokud vliv nenalezneme, může to znamenat buď že tento vliv neexistuje, nebo že existuje, ale síla testu nebyla dostatečná pro jeho nalezení.

Výsledek „zamítáme H_0" má docela pevný logický základ (pro $\alpha = 0{,}05$ platí s pravděpodobností 95%). Ale říct v případě, kdy nebyla překročena kritická hodnota, že „H_0 přijímáme" nebo „H_0 platí", je příliš ukvapené, protože při důkladnějším hledání by se mohlo ukázat, že určitý vliv existuje, tj. H_0 neplatí. Ustálilo se tedy rčení „nezamítáme H_0", které odpovídá jisté opatrnosti v učinění konečného závěru.

Fráze „nezamítáme H_0" je tedy opatrným vyjádřením, které je zcela na místě. Znamená to, že říkáme: „Výsledky testu nám neposkytují dostatečné důkazy k závěru, že H_0 neplatí."

K příkladu 11.4: Co by to znamenalo, kdyby pan Loftus provedl tak důkladné hledání, že by obrátil celý byt naruby, ale přesto brýle nenašel? Stále by existovala jistá malá šance, že je přehlédl v nějaké zapadlé škvíře, ale protože je nenašel, bylo by rozumné koupit nové.

Podobně i experiment provedený v úžasné síle a rozsahu, pokud neprokáže vliv nezávislé proměnné na závislou proměnnou, nás vede k závěru, že „je rozumné přijmout H_0".

Příklad 11.5 *Chceme otestovat kvalitu jisté techniky pamatování. Vybereme náhodně dvě skupiny lidí. Oběma skupinám je předložen jistý počet navzájem nesouvisejících slov k zapamatování (například* 20 *slov). Poté jsou lidé z první skupiny okamžitě vyzkoušeni, kolik slov si zapamatovali. Lidé ze druhé skupiny jsou vyzkoušeni až o týden později. Čili nezávislou proměnnou je doba pamatování, závislou proměnnou je počet zapamatovaných slov z daných dvaceti. Stanovíme hypotézy testu:*

H_0**:** *počet zapamatovaných slov nezávisí na době pamatování (technika zapamatování byla tak dobrá, že po týdnu si pamatuji stejně dobře jako po bezprostředním naučení slov).*

H_1**:** *počet zapamatovaných slov závisí na době pamatování (= době držení slov v paměti).*

Pokud v testovaných skupinách bude například jen v každé pět lidí a vliv se neprokáže, můžeme uzavřít, že H_0 platí, tj. že technika učení je vynikající? Asi ne, protože právě zde bychom se mohli dopustit té chyby, že jsme vliv nenašli, i když je možné, že existuje. Na druhé straně pokud by v každé z testovaných skupin bylo 10000 *lidí a stále by se neprokázala existence závislosti, závěr „H_0 platí“ by asi byl docela rozumný.*

Poznámka B) Snížení rozptylu zvyšuje sílu testu V příkladu 11.5 lze vidět jednu celkem přirozenou skutečnost, že totiž výpovědní síla statistického testu se zvyšuje se zvýšením počtu měření. Pro připomenutí **síla testu** je pozitivní pojem – je to pravděpodobnost, se kterou správně zamítneme H_0, pokud platí H_1. Je to tedy jakási síla nalezení vlivu mezi proměnnými testu. Na tuto sílu má především vliv rozptyl $\sigma^2_{\overline{X}}$ neboli

$$\sigma^2_{\overline{X}} = \frac{\sigma^2}{N}. \tag{11.13}$$

Cokoli, co snižuje rozptyl $\sigma^2_{\overline{X}}$, zvyšuje současně i sílu testu. Jedním činitelem je počet měření N – je vidět ze vzorce, že pokud zvyšujeme N, zvyšuje se hodnota jmenovatele ve zlomku, a tím klesá hodnota rozptylu. Další možností, jak snížit rozptyl, je snížit přímo hodnotu rozptylu σ^2 celé populace v čitateli zlomku. Pokud si říkáte, že σ^2 je dáno a nelze je přece měnit, máte pravdu. Ale stejně některé vlivy na rozptyl σ^2 celé populace můžeme vyloučit vhodným naplánováním experimentu.

Příklad 11.6 *Experimentem se má zjistit, zda alkohol v krvi má vliv na reakční dobu řidiče. Náhodně byly vybrány dvě skupiny lidí. První skupině je nabídnut alkohol ve formě ginu s tonikem, druhé skupině pouze tonik. Pak jsou všichni podrobeni měření reakční doby. Testovaný člověk sedí u stolu a před sebou má lampu s červenou žárovkou a tlačítko. Lampa se rozsvěcuje v různých nepravidelných intervalech – jakmile se rozsvítí, je úkolem testovaného co nejrychleji stisknout tlačítko. Je změřena jeho reakční doba (v milisekundách). U každého člověka se měření několikrát opakuje, a pak je vypočten průměr jeho reakční doby.*

Samozřejmě uvnitř každé ze skupin se projeví určitá variabilita v průměrné době reakce. Ta je dána různými faktory, z nichž některé nemůžeme ovlivnit, ale jiné ano. Mezi **neovlivnitelné faktory** *patří:*

1. *Nálada člověka během měření (špatná nálada = delší doba reakce).*
2. *Osobní předpoklady – někdo má prostě schopnost rychlejší reakce než ostatní.*
3. *Postoj člověka vůči experimentu (znuděný postoj = delší doba reakce).*

Mezi **ovlivnitelné faktory** *lze zahrnout*

1. *Teplotu v místnosti, kde se provádí měření (extrémní teploty ⇒ delší doba reakce).*

2. *Vlhkost v místnosti, kde se provádí měření (vyšší vlhkost ⇒ delší doba reakce).*

3. *Pohlaví (u žen ... kratší doba reakce).*

4. *Věk (vyšší věk ... kratší doba reakce).*

5. *Čas dne (měření pozdě odpoledne ... delší doba reakce).*

Některé faktory rozptylu hodnot můžeme významně ovlivnit – jak výběrem sledované populace (omezení se na stejné pohlaví a věk eliminuje rozdíly způsobené těmito faktory), tak zajištěním stejných podmínek měření (stálá vlhkost a teplota místnosti, měření u všech ve stejnou denní dobu). Tímto způsobem plánování a provedení experimentu pak následný statistický test získá větší výpovědní sílu v těch parametrech, které jsou pro nás důležité, a není zkreslen rozdíly v těch ovlivnitelných faktorech, které nás nezajímají.

11.4 Interval spolehlivosti pro střední hodnotu μ

Kromě statistických testů jsou při zpracování měření často užitečnější tzv. **intervaly spolehlivosti**. V této fázi výkladu se seznámíme s intervalem spolehlivosti pro střední hodnotu $E\overline{X} = \mu$ průměru z normálního rozdělení.

11.4.1 Interval spolehlivosti pro μ při známém rozptylu

Střední hodnotu μ měřené veličiny obyčejně neznáme. Kdybychom ji znali, nemusíme provádět ani měření, ani statistický test Určit μ je v podstatě naším cílem.

Jakýmsi odhadem střední hodnoty je výběrový průměr $\overline{X}$. Ovšem tento průměr (zejména při nižším počtu měření N) není přesně roven střední hodnotě μ – vlastně víme, že platí $P(\overline{X} = \mu) = 0$. Potřebovali bychom spíše najít nějaký interval $(\overline{X} - a; \overline{X} + a)$ pro nějaké vhodné a „ne příliš velké“ $a > 0$. A to bude vlastně interval spolehlivosti pro střední hodnotu μ. Přesněji řečeno, **r%-ní interval spolehlivosti pro střední hodnotu μ průměru $\overline{X}$ je takový interval, který obsahuje μ s pravděpodobností $\frac{r}{100}$.**

Příklad 11.7 *Životnost 75-wattové žárovky má normální rozdělení se směrodatnou odchylkou $\sigma = 25$ hodin. U náhodně vybraného vzorku dvaceti žárovek byla naměřena průměrná životnost $\overline{x} = 1024$ hodin. Utvořte 95%-ní interval spolehlivosti pro střední hodnotu μ průměrné životnosti.*

Řešení tohoto příkladu je instruktivní, proto si dovolím vypnout kurzívu (:-)). Každý interval spolehlivosti je velmi úzce svázán se statistickým testem. Budeme se zabývat zejména oboustrannými intervaly spolehlivosti, proto v našem příkladu je zde vazba na oboustranný U-test s hypotézami

H_0: $\mu = \mu_0$ (skutečnou střední hodnotu μ_0 neznáme).

H_1: $\mu \neq \mu_0$.

Pokud hledáme 95%-ní interval spolehlivosti, je příslušná hladina významnosti $\alpha = 0{,}05$. Kritériem testu je výběrový průměr $\overline{X}$. Dále vezmeme v úvahu počet měření pro výpočet průměru:

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{N}} = \frac{25}{\sqrt{20}} = 5{,}59.$$

Jak je dobře známo ze BMA3, příslušné kritické hodnoty v tomto případě jsou

$$u_{\frac{\alpha}{2}} = -1{,}96, \quad u_{1-\frac{\alpha}{2}} = 1{,}96.$$

Obě tyto hodnoty lze převést na kritické hodnoty vzhledem k veličině $\overline{X}$:

$$-1{,}96 = \frac{\overline{x}_m - \mu_0}{5{,}59} \quad \Rightarrow \quad \overline{x}_m = \mu_0 - 1{,}96 \cdot 5{,}59 = \mu_0 - 10{,}96;$$

$$1{,}96 = \frac{\overline{x}_v - \mu_0}{5{,}59} \quad \Rightarrow \quad \overline{x}_v = \mu_0 + 1{,}96 \cdot 5{,}59 = \mu_0 + 10{,}96.$$

Interval pro „nezamítnutí H_0“ je pak

$$\overline{X} \in (\mu_0 - 10{,}96; \mu_0 + 10{,}96)\,. \tag{11.14}$$

Ovšem hodnotu μ_0 neznáme. Ale pokud ze vzorce 11.14 vyjádříme místo $\overline{X}$ hodnotu μ_0, dostaneme vztah pro interval spolehlivosti:

$$\mu_0 \in \left(\overline{X} - 10{,}96; \overline{X} + 10{,}96\right). \tag{11.15}$$

Odtud pro jednu realizaci $\overline{x} = 1024$ dostaneme

$$\mu_0 \in (1024 - 10{,}96; 1024 + 10{,}96) = (1013{,}04; 1034{,}96),$$

se spolehlivostí 0,95, což je odpověď příkladu 11.7. **Obecně s využitím symetrie**

$$u_{\frac{\alpha}{2}} = -u_{1-\frac{\alpha}{2}}$$

lze $(1-\alpha)\cdot 100$%-ní interval spolehlivosti pro střední hodnotu μ při známém rozptylu psát ve tvaru

$$\mu \in \left(\overline{X} - u_{1-\frac{\alpha}{2}} \cdot \sigma_{\overline{X}}; \overline{X} + u_{1-\frac{\alpha}{2}} \cdot \sigma_{\overline{X}}\right). \tag{11.16}$$

V některé literatuře se objevuje kromě pojmu „interval spolehlivosti“ ještě pojem „konfidenční interval“ – to je jen nesprávný (nebo jiný) překlad anglického termínu „confidence interval“, ale jedná se o totéž (překladatelé se zase jednou nedomluvili ... :-)).

Pojem intervalu spolehlivosti úzce souvisí se silou příslušného statistického testu – jak už to vyplývá ze vzorce 11.16, čím větší je síla příslušného statistického testu, tím menší je rozptyl $\sigma^2_{\overline{X}}$, a tím užší je interval spolehlivosti.

11.4.2 Interval spolehlivosti pro μ při neznámém rozptylu

Příklad 11.8 *Uvažujme stejnou situaci jako v příkladu 11.7 ($N = 20$, $\overline{x} = 1024$), pouze odchylka σ není známá a musíme ji odhadnout – tedy z měření životnosti u dvaceti žárovek byl vypočten rozptyl $\overline{s^2} = 625$. Odtud*

$$est\sigma^2_{\overline{X}} = \frac{\overline{s^2}}{N} = \frac{625}{20} = 31{,}25.$$

Tedy $est\sigma_{\overline{X}} = \sqrt{31{,}25} = 5{,}59$ – toto číslo je stejné jako v příkladu 11.7, ovšem nyní se nejedná o přesnou hodnotu, ale odhad, čili zde bude $\nu = N - 1 = 19$ stupňů volnosti odhadu. Nalezněte 95 %-ní interval spolehlivosti pro střední hodnotu μ výběrového průměru $\overline{X}$.

V příslušném statistickém testu použijeme nyní t-rozdělení – nalezneme kritické hodnoty pro oboustanný test ($\alpha = 2q = 0{,}05$) a pro $\nu = 19$ stupňů volnosti (tabulka 11.10): $t_v = 2{,}093$, odtud $t_m = -2{,}093$.

Hypotézy H_0, H_1 a kritérium oboustranného testu zde zůstávají stejné jako v 11.7: $H_0 : \mu = \mu_0$, $H_1 : \mu \neq \mu_0$, kritériem je výběrový průměr $\overline{X}$. Přepočteme nyní kritické hodnoty vzhledem k veličině $\overline{X}$:

$$-2{,}093 = \frac{\overline{x}_m - \mu_0}{5{,}59} \quad \Rightarrow \quad \overline{x}_m = \mu_0 - 2{,}093 \cdot 5{,}59 = \mu_0 - 11{,}7;$$
$$2{,}093 = \frac{\overline{x}_v - \mu_0}{5{,}59} \quad \Rightarrow \quad \overline{x}_v = \mu_0 + 2{,}093 \cdot 5{,}59 = \mu_0 + 11{,}7.$$

Nyní interval pro „nezamítnutí H_0“ je

$$\overline{X} \in (\mu_0 - 11{,}7; \mu_0 + 11{,}7)\,, \tag{11.17}$$

nás ovšem zajímá spíše tvar (rychle jej naň převedem)

$$\mu_0 \in \left(\overline{X} - 11{,}7; \overline{X} + 11{,}7\right). \tag{11.18}$$

Odtud pro příklad 11.8 se spolehlivostí 0,95 a realizaci (= konkrétní měření) průměru $\overline{x} = 1024$ dostaneme

$$\mu_0 \in (1024 - 11{,}7; 1024 + 11{,}7) = (1012{,}3; 1035{,}7),$$

Vidíme, že při větší míře nejasnosti, kdy rozptyl neznáme a musíme jej odhadnout, je interval spolehlivosti širší (ostatní hodnoty v příkladech 11.7 a 11.8 jsou stejné, takže lze opravdu provést porovnání).

Obecně s využitím symetrie

$$t_{\frac{\alpha}{2}} = -t_{1-\frac{\alpha}{2}}$$

lze $(1-\alpha)\cdot 100$**%-ní interval spolehlivosti pro střední hodnotu** μ **při neznámém rozptylu psát ve tvaru**

$$\mu \in \left(\overline{X} - t_{1-\frac{\alpha}{2}} \cdot est\sigma_{\overline{X}}; \overline{X} + t_{1-\frac{\alpha}{2}} \cdot est\sigma_{\overline{X}}\right). \tag{11.19}$$

11.4.3 Několik důležitých poznámek k intervalům spolehlivosti

Následuje několik poznámek k intervalům spolehlivosti, každá z nich je důležitější než ty ostatní ... :-)

a) **Vztah mezi intervalem spolehlivosti a pravděpodobností.** Je potřeba říci něco velmi důležitého k formulaci „95%-ní interval spolehlivosti pro střední hodnotu μ obsahuje tuto hodnotu s pravděpodobností 0,95“. Věta je vyslovena správně ale může být zavádějící – a problém pochopení souvisí s pojmem **statistika** a **realizace statistiky** (viz definice 11.2). Interval spolehlivosti je totiž vlastně intervalem, jehož mezemi jsou náhodné veličiny. A pak pro konkrétní měření dosadíme do vzorců 11.16, 11.19 konkrétní realizaci $\overline{x}$. Třeba v příkladu 11.7 interval $(1013{,}04; 1034,96)$ neznámou střední hodnotu μ buď obsahuje (stoprocentně), nebo neobsahuje (stoprocentně).

   Souvislost s pravděpodobností se projeví pouze při opakovaném měření a opakované konstrukci realizace intervalu spolehlivosti: pokud bychom například tisíckrát zopakovali experiment změření životnosti u dvaceti žárovek, spočetli tisíc různých realizací $\overline{x}_1$, $\overline{x}_2 \ldots$, $\overline{x}_{1000}$ a sestrojili tisíc různých realizací intervalu spolehlivosti podle vzorce 11.16, tak přibližně 95% těchto intervalů (tj. asi 950 z nich) bude neznámou hodnotu μ obsahovat, zbylých pět procent ji obsahovat nebude.

   V dalších výpočtech realizací intervalů spolehlivosti pro $\overline{x}$ budu slovo „realizace“ vynechávat, takže pojem **interval spolehlivosti** se středem v $\overline{X}$ (středem intervalu je nikoli konkrétní hodnota, ale náhodná veličina) bude splývat s pojmem **realizace intervalu spolehlivosti** se středem v $\overline{x}$ (středem intervalu je konkrétní číslo). Matematicky je mezi těmito dvěma pojmy rozdíl, ale v praxi pod intervalem spolehlivosti máme na mysli vždy konkrétní interval vypočtený na základě měření. !!!! :-)

b) **Jednostranné intervaly spolehlivosti.** Pozorný čtenář by možná mohl položit otázku: no dobrá, oboustranný interval spolehlivosti odpovídá jistému oboustrannému statistickému testu (příklad 11.7) – pokud se ale týká jednostranných testů (například test grafické iluze 11.3), existuje také u nich nějaký analogický jednostranný interval spolehlivosti? Odpověď zní „ano, existuje“. Jeho odvození je analogické, proveďme jej například pro data testu grafické iluze 11.3:

$H_0 : \mu = \mu_0$, $H_1 : \mu > \mu_0$, kritériem je $\overline{X}$, kritickou hodnotu použijeme přesně tu stejnou jako v daném pravostranném testu: $t_k = 2{,}132$. Převodem normovaného tvaru kritické hodnoty do tvaru aktuálního vzhledem k veličině $\overline{X}$ dostaneme

$$2{,}132 = \frac{\overline{x}_k - \mu_0}{1} \quad \Rightarrow \overline{x}_k = \mu_0 + 2{,}132 \cdot 1 = \mu_0 + 2,132;$$

tedy interval pro „nezamítnutí H_0" je

$$\overline{X} \in (-\infty; \mu_0 + 2{,}132), \tag{11.20}$$

ale protože nás zajímá spíše interval pro μ_0, tak z tohoto vztahu vyjádříme ohraničení pro μ_0 a index 0 vypustíme:

$$\mu \in (\overline{X} - 2{,}132; \infty) = (5{,}868; \infty) \tag{11.21}$$

a to je hledaný interval se spolehlivostí 0,95. Vidíme, že **oboru 11.20 pravostranného testu odpovídá levostranný interval spolehlivosti 11.21**.

I když jsou tedy jednostranné intervaly spolehlivosti zcela přirozené a možné, v dalším textu se na ně nebudeme příliš zaměřovat – spíše nás bude zajímat vymezení pro μ na intervalu konečné délky. Tedy i když budeme provádět jednostranné statistické testy, intervaly spolehlivosti budeme hledat na základě příslušného oboustranného testu se stejnou hladinou významnosti.

Tak tedy i v příkladu 11.3 lze sestrojit oboustranný interval spolehlivosti: Pro $\nu = 4$ je kritická hodnota rovna

$$t_{1-\frac{0{,}05}{2}}(4) = 2{,}776;$$

Pak (vzorec 11.19):

$$\mu \in (8 - 2{,}776 \cdot 1; 8 + 2{,}776 \cdot 1) = (5{,}224; 10{,}776).$$

c) **Vztah mezi intervalem spolehlivosti a statistickým testem.** Mezi výsledkem statistického testu a intervalem spolehlivosti existuje jednoznačný vztah:

Statistický test zamítne hypotézu $H_0 : \mu = \mu_0$ právě tehdy, když μ_0 nenáleží do příslušného intervalu spolehlivosti

Třeba pokud bychom v situaci příkladu 11.7 testovali hypotézu $H_0 : \mu = 1000$, místo abychom prováděli test, stačí se podívat na příslušný interval spolehlivosti oboustranného testu: $1000 \notin (1013{,}04; 1034{,}96)$, tj. to znamená, že příslušný statistický test zamítne hypotézu H_0.

Nebo podíváme-li se na příklad pravostranného testu 11.3, kde $H_0 : \mu = 6$, vidíme, že $6 \in (5{,}868; \infty)$ (levostranný interval spolehlivosti vypočten v předchozí poznámce b)), tj. to znamená, že hypotéza H_0 příslušného jednostranného testu nebude zamítnuta.

d) Interval spolehlivosti uvádí více informací než statistický test. Omlouvám se za další členění, ale budu prezentovat tuto poznámku ve čtyřech myšlenkách:

1. Výsledek statistického testu není přesně to, co bychom chtěli znát. Ve skutečnosti chceme znát míru platnosti hypotézy H_1, je-li dán výsledek experimentu – ovšem místo toho se ze statistického testu dovídáme pravděpodobnost výsledku experimentu za předpokladu platnosti hypotézy H_0 (a stále neznáme **míru platnosti** H_1, a to ani v případě, kdy je H_0 zamítnuto).
2. Statistický test nám ve svém výsledku nedává informaci o rozsahu studovaného vlivu, kdežto interval spolehlivosti ano. Informace o umístění střední hodnoty μ je ve statistických testech, které jsme dosud prováděli, skryta, ale interval spolehlivosti činí tuto informaci zjevnou a překládá ji do srozumitelného měřítka.

   Například test 11.3 pouze prohlásí, že se nenašlo dostatek důkazů pro vliv grafické iluze na μ. Ale interval spolehlivosti (nyní už spíše ten oboustranný, tj. pro $\alpha = 0{,}05$ byl odvozen v poznámce b)) $(5{,}224; 10{,}776)$ navíc naznačuje, že se spolehlivostí 0,95 by střední hodnota místo šesti mohla být stejně dobře rovna i osmi nebo deseti, ale už ne jedenácti nebo dvanácti.
3. Statistický test zdůrazňuje chybu prvního druhu, ale říká velmi málo o chybě druhého druhu. Na druhé straně interval spolehlivosti naznačuje i chybu druhého druhu – pokud α necháme pevné a zmenšujeme β, tak interval spolehlivosti zmenšuje svou délku.
4. Závěr: Z uvedených důvodů je lepší používat intervaly spolehlivosti spíše než jen pouhé testování hypotéz. Někdy statistické zpracování v literatuře příliš zdůrazňuje testy a zapomíná dodat užitečné informace navíc, které lze snadno vyčíst z intervalů spolehlivosti.

11.5 t-test typu „$\mu_1 = \mu_2$“

11.5.1 Párový test

V tomto oddílu se budeme zabývat statistickými testy při experimentech, kde získáváme dva soubory měření. Zde je potřeba si dát pozor na vztah mezi těmito dvěma soubory (= skupinami) měření, na základě tohoto vztahu rozlišujeme totiž dva typy statistického testu – párový a nepárový test. Párovým testem se budeme zabývat nejdříve – spočívá v tom, že sice získáme dvě skupiny (= dva soubory) měření, ale tyto soubory jsou navzájem těsně svázány v tom smyslu, že ke každé hodnotě v prvním souboru měření lze jednoznačně přiřadit tzv. párovou hodnotu měření ze druhého souboru. Zejména to taky znamená, že počet měření v obou souborech je stejný – a v podstatě bychom mohli říct, že místo dvou souborů měření máme jediný soubor, ve kterém jedna položka je reprezentována uspořádanou dvojicí hodnot.

Párový test tedy užijeme v situaci, kdy sice máme k dispozici dva soubory měření, ale tyto dva soubory měření jsou spolu těsně svázány – obyčejně tak, že v obou skupinách jsou hodnoceni stejní jedinci; nejprve provedeme měření vybrané skupiny jedinců za systému podmínek A, a pak provedeme měření téže skupiny jedinců za systému podmínek B. Proto se tomuto typu experimentů také říká **experiment opakovaného měření**. Další vhodný název je zde experiment typu **„jedna skupina dvakrát“**, protože jedna skupina jedinců je podrobena měření při dvou různých situacích.

Příklad 11.9 *Chceme experimentem zjistit, jak se změní počet úderů srdce člověka za minutu po vypití šálku kávy (studujeme vliv kofeinu na činnost srdce). Kdybychom k tomuto experimentu přistupovali „nepárovým“ přístupem a vybrali náhodně dvě skupiny lidí, z nichž jedna by vypila kávu s kofeinem a druhá kávu bez kofeinu, do měření by byl zanesen jistý rozptyl způsobený tím, že tempo srdečních úderů se liší u různých lidí.*

Mnohem vhodnější je zde experiment opakovaného měření, kdy jedna a táž skupina vybraných lidí je vystavena měření po kávě bez kofeinu, a pak po kávě s kofeinem. Potom se vypočte rozdíl obou hodnot vždy u téhož člověka a testuje se, zda je tento rozdíl významný. Byla získána následující data počtu srdečních úderů za minutu u devíti lidí (na každém řádku jsou hodnoty měření jednoho člověka):

x_{i1} *bez kofeinu*	x_{i2} *s kofeinem*	*rozdíl* $x_{i2} - x_{i1}$
70	*76*	*6*
60	*61*	*1*
49	*52*	*3*
72	*71*	*-1*
70	*81*	*11*
66	*70*	*4*
55	*55*	*0*
54	*61*	*7*
80	*89*	*9*

První dva sloupečky tabulky představují dva soubory měření párového testu. My ovšem budeme dále pracovat jen s jednorozměrným souborem, a sice s vektorem rozdílů měření ve třetím sloupci. To znamená, že párový test vlastně odpovídá situaci jednorozměrného souboru měření (= oddílu 11.2).

Spočteme průměr $\overline{x} = 4{,}44$ a odhadneme rozptyl pomocí $\overline{s^2}$. Aby čtenář neměl pocit, že v oddílu 11.5 nebude vzhledem k 11.2 nic nového, vypočteme $\overline{s^2}$ pomocí pátého možného vzorce, který jsme si ještě neuváděli:

$$\overline{S^2} = \frac{SS}{\nu} \tag{11.22}$$

(ono se vlastně jedná o vzorec 11.12, ovšem ve jmenovateli vzorce je ν místo $N-1$). Dále do čitatele za SS dosadíme ze vzorce 11.11, který zkombinujeme s nejpohodlnějším způsobem výpočtu 11.8:

$$SS = N \cdot S^2 = \left(\sum_{i=1}^{N} X_i^2\right) - \frac{\left(\sum_{i=1}^{N} X_i\right)^2}{N},$$

čili pak pro realizaci ss platí

$$ss = \left(\sum_{i=1}^{N} x_i^2\right) - \frac{\left(\sum_{i=1}^{N} x_i\right)^2}{N}, \tag{11.23}$$

a po dosazení

$$ss = 314 - \frac{40^2}{9} = 136{,}22.$$

Počet stupňů volnosti pro odhad rozptylu je $\nu = N - 1 = 8$, a tedy

$$\overline{s^2} = \frac{ss}{\nu} = 17{,}03.$$

a) Sestrojme nyní 95**%-ní interval spolehlivosti** pro μ: t_k najdeme jako průsečík řádku $\nu = 8$ a sloupce $2q = \alpha = 0{,}05$ (oboustranný interval spolehlivosti vychází z oboustranného testu): $t_k = 2{,}306$. Pak interval pro μ se spolehlivostí 0,95 je

$$\overline{x} \pm \sqrt{\frac{\overline{s^2}}{N}} \cdot t_k(\nu = 8) = 4{,}44 \pm \sqrt{\frac{17{,}03}{9}} \cdot 2{,}306 = (1{,}27; 7{,}61).$$

Z tohoto intervalu spolehlivosti se dovídáme, že kofein zvyšuje činnost srdce, a sice o 1,27 až o 7,61 úderů za minutu.

b) Provedeme i statistický t-test:

(K1) H_0: $\mu = 0$ (rozdíl hodnot je nulový, tj. počet úderů srdce za minutu je stejný s kofeinem i bez kofeinu – srdeční tep nezávisí na kofeinu);
H_1: $\mu \neq 0$.

(K2) Testovým kritériem bude výběrový průměr $\overline{X}$, respektive jeho normovaná hodnota $\frac{\overline{X}-0}{est\ \sigma_{\overline{X}}}$.

(K3) $\overline{s^2} = 17{,}03 \Rightarrow$

$$est\ \sigma^2_{\overline{X}} = \frac{\overline{s^2}}{N} = \frac{17{,}03}{9} = 1{,}89 \Rightarrow est\ \sigma_{\overline{X}} = \sqrt{1{,}89} = 1{,}37.$$

Tj. za předpokladu platnosti H_0 má veličina $\frac{\overline{X}-0}{1{,}37}$ Studentovo t-rozdělení s $\nu = 8$ stupni volnosti.

(K4) $t_k = \pm 2{,}306$ je u oboustranného testu stejné jako u intervalu spolehlivosti a).

(K5) Příslušná t-hodnota je

$$\frac{4{,}44 - 0}{1{,}37} = 3{,}24 > 2{,}306,$$

a tedy zamítáme H_0 o nezávislosti, je potvrzen vliv kofeinu na zvýšení srdečního tepu.

Hypotézy H_0 nejsou pravdivé téměř nikdy (pokud uvažujeme větší počet desetinných míst), tj. zamítnutí H_0 nám nic podstatného neříká. Spíše nás zajímalo, jak velký je vliv kofeinu, a to jsme se dozvěděli z intervalu spolehlivosti.

11.5.2 Nepárový test

Nyní se pojďme věnovat **nepárovému testu** neboli zpracování dat při experimentu typu **„dvě skupiny jednou“**.

Příklad 11.10 *Chceme zjistit kvalitu jisté techniky pamatování. Náhodně jsme vybrali deset lidí a rozdělili do dvou skupin po pěti lidech. Skupina 1 (tzv. experimentální skupina) se naučila 100 zadaných slov novou technikou, skupina 2 (kontrolní skupina … zažitý nesprávný překlad anglického „control group“, správný význam překladu je „řízená skupina“, protože „control“= řídit, vést, nikoli kontrolovat) použila obyčejnou klasickou techniku zapamatování. Po týdnu se vyzkouší, kolik si kdo pamatuje z daných 100 slov – jsou získána data*

experimentální skupina	*kontrolní skupina*
43	*16*
37	*22*
51	*24*
27	*30*
32	*18*

Nyní data na jednom řádku spolu nijak nesouvisí, jedná se o měření u dvou různých lidí. Zdá se, že experimentální skupina má lepší výsledky, ale musíme statisticky prokázat, že to není způsobeno pouze náhodnými vlivy.

V každé ze skupin vypočteme průměr a odhadneme rozptyl.

skupina 1: $\overline{x}_1 = 38$, podle vzorce 11.23 $ss_1 = 352$, $\nu_1 = 4$, a tedy podle 11.22

$$est_1\ \sigma^2 = \overline{s_1^2} = \frac{352}{4} = 88.$$

skupina 2: $\overline{x}_2 = 22$, podle vzorce 11.23 $ss_2 = 120$, $\nu_2 = 4$, a tedy podle 11.22

$$est_2\ \sigma^2 = \overline{s_2^2} = \frac{120}{4} = 30.$$

No a teď přicházíme k úvahám, které jsou pro další vývoj (i další kapitolu) důležité. V situaci experimentu typu „dvě skupiny jednou“ je potřeba se vypořádat se dvěma typy rozptylu, kterými jsou – **vnitřní rozptyl** a **vnější rozptyl**.

vnitřní rozptyl: Jedná se o rozptyl představující vzájemnou rozdílnost jedinců v celé populaci (např. rozdílnost lidí, rozdílnost různých součástek stejného typu, apod.). V tomto textu se budeme převážně zabývat situacemi, kdy je splněn tzv. **princip homogenního vnitřního rozptylu: jakýkoliv experiment nemá vliv na rozptyl rozdělení celé populace, z níž byla náhodně vybrána skupina jedinců pro měření**.

Slovy našeho příkladu, rozdílnost výsledků est_1 způsobená růzností lidí ve skupině 1 je přibližně stejná jako rozdílnost výsledků est_2 způsobená růzností lidí ve skupině 2. Jinak řečeno, ať už měříte daný soubor měření za jakékoli podmínky, „rozmanitost“ těchto měření je v dané skupině přibližně stejná.

Z Tohoto principu homogenního rozptylu tedy plyne, že odhady est_1, est_2 jsou odhady jednoho a stejného vnitřního rozptylu σ^2 (díky tomu jsem u písmenka σ o pár řádků výše už nepsal žádný index), který vypočteme jako aritmetický průměr obou odhadů:

$$est\ \sigma^2 = \frac{est_1\ \sigma^2 + est_2\ \sigma^2}{2} = \frac{88 + 30}{2} = 59.$$

Tedy nejlepší možný odhad vnitřního rozptylu σ^2 je roven 59 a v dalším budeme pracovat s ním. Počet stupňů volnosti je $\nu = 4 + 4 = 8$, protože jsme tento odhad získali pomocí dvou jiných odhadů o počtu stupňů volnosti 4 – stupně volnosti ve výsledném odhadu sečítáme.

vnější rozptyl: Jedná se o **rozptyl vyjadřující rozdílnost mezi danými podmínkami měření**. Tento rozptyl v případě dvou souborů měření lze odhadnout na základě průměrů $\overline{x}_1 = 38$, $\overline{x}_2 = 22$. Způsob je následující: Vypočteme rozptyl těchto dvou průměrů podle vzorců 11.23 a 11.22: Protože $\nu = 2 - 1 = 1$, máme

$$est\ r_N = \frac{ss}{1} = \frac{38^2 + 22^2 - \frac{60^2}{2}}{1} = 128.$$

Ale to ještě není všechno – tento odhad je odhadem rozptylu průměru pěti hodnot ($N = 5$). Abychom získali odhad rozptylu jediné hodnoty měření, musíme v souladu se vzorcem $est\ r_N = \frac{est\ r}{N}$ (analogie vzorce 11.13) počítat

$$est\ r = N \cdot est\ r_N = 5 \cdot 128 = 640.$$

celkový rozptyl: Aby byla plejáda přehledu rozptylů úplná, je možné si položit následující otázku – co se vlastně spočítá, když budeme považovat všech deset hodnot za měření jediné veličiny a vypočteme $\overline{s^2}$ ze všech deseti měření? Podle logiky výpočtu by to měl být jakýsi celkový rozptyl – a skutečně je tomu tak.

Průměr všech deseti hodnot je $\overline{x} = 30$, což je mimochodem aritmetický průměr hodnot $\overline{x}_1$, $\overline{x}_2$. Bude tedy podobně hodnota $\overline{s^2}$ průměrem hodnot $\overline{s_1^2}$, $\overline{s_2^2}$?

Podle vzorce 11.23 $ss = 1112$, a tedy podle 11.22

$$\overline{s^2} = \frac{ss}{\nu} = \frac{1112}{9} \doteq 123{,}56,$$

což není hodnota rovná součtu $\overline{s_1^2} + \overline{s_2^2} = 88 + 30 = 118$. Ještě méně se zdá, že by odhad celkového rozptylu 123,56 byl součtem odhadu vnitřního rozptylu 59 a odhadu vnějšího rozptylu 640 – ale přesto zde platí jisté součtové vzorce:

a) Celkový počet stupňů volnosti 9 u odhadu celkového rozptylu je součtem volnosti 8 u odhadu vnitřního rozptylu a volnosti 1 u odhadu vnějšího rozptylu.

b) Celkový součet čtverců 1112 je součtem součtu čtverců 472 vnitřního rozptylu (který vznikl součtem $ss_1 = 352$ a $ss_2 = 120$) a 640 u vnějšího rozptylu.

Tolik přehled a jemné intro do problematiky rozptylu – více na to téma nebude prostor, ale čtenáři by se s tématem setkali při už zmiňované analýze rozptylu (ANOVA), která se používá až tehdy, když skupiny měření jsou tři a více. Nyní se vraťme k řešení našeho příkladu.

Odhad vnitřního rozptylu je est $\sigma^2 = 59$, *odtud odhad rozptylu průměru*

$$est\sigma^2_{\overline{X}} = \frac{59}{5} = 11{,}8.$$

Nyní lze užitím 11.19 určit např. 95 %-ní interval spolehlivosti pro střední hodnotu průměru v každé ze skupin měření: Pro $t_k(\nu = 8; \alpha = 2q = 0{,}05) = 2{,}306$

$$\mu_1 \in 38 \pm \sqrt{11{,}8} \cdot 2{,}306 \doteq (30{,}08; 45{,}92),$$

$$\mu_2 \in 22 \pm \sqrt{11{,}8} \cdot 2{,}306 \doteq (14{,}08; 29{,}92).$$

V souvislosti se statistickým testem tohoto příkladu nás ovšem spíše zajímá interval spolehlivosti pro střední hodnotu veličiny $\overline{X_1} - \overline{X_2}$. *Střed intervalů spolehlivosti bude v tomto případě* $\overline{x_1} - \overline{x_2} = 38 - 22 = 16$, *odhad rozptylu je*

$$est\ \sigma^2_{\overline{X_1}-\overline{X_2}} = est\ \sigma^2_{\overline{X_1}} + est\ \sigma^2_{\overline{X_2}} = \frac{59}{5} + \frac{59}{5} = 23{,}6.$$

Tedy pro $t_k(\nu = 8; \alpha = 2q = 0{,}05) = 2{,}306$

$$\mu_1 - \mu_2 \in 16 \pm \sqrt{23{,}6} \cdot 2{,}306 \doteq 16 \pm 11{,}2 = (4{,}8; 27{,}2).$$

Protože výsledný interval spolehlivosti pro rozdíl středních hodnot neobsahuje nulu, víme také, že příslušný statistický test zamítne hypotézu $H_0 : \mu_1 = \mu_2$. A skutečně, pojďte se přesvědčit:

(K1) *H_0: $\mu_1 = \mu_2$ (střední hodnoty obou skupin ohodnocení jsou stejné, tj. nová technika zapamatování ovlivňuje výsledek přibližně stejně jako ta dřívější).*
H_1: $\mu_1 \neq \mu_2$ (nová technika zapamatování má přibližně stejné výsledky jako dřívější technika).

(K2) *Testovým kritériem bude rozdíl $\overline{X_1} - \overline{X_2}$.*

(K3) *Při platnosti H_0 má veličina*

$$\frac{\overline{X_1} - \overline{X_2} - 0}{\sqrt{est\ \sigma^2_{\overline{X_1} - \overline{X_2}}}} = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{23{,}6}}$$

Studentovo t-rozdělení pro $\nu = 8$.

(K4) *Pro $\alpha = 0{,}05$ příslušná kritická hodnota t_k je na průsečíku řádku $\nu = 8$ a sloupce pro $\alpha = 2q = 0{,}05$, tj. $t_k = 2{,}306$.*

(K5) *$\frac{38-22}{4{,}858} = 3{,}29 \notin (-t_k; t_k)$. H_0 tedy zamítáme – nová technika významně zvyšuje úroveň zapamatování.*

Příklad 11.11 *Psychologové fyziologie chtějí experimentem ověřit, že podvěsek mozkový je hlavním řídícím centrem sexuálního chování. Rozhodli se získat dvacet dobrovolníků z řad studentů, kteří by se chtěli podrobit operaci mozku. Protože žádní dobrovolníci se nepřihlásili (zejména do experimentální skupiny), byly náhodně vybrány dvě skupiny po deseti krysách.*

Krysám z experimentální skupiny byl operací odebrán podvěsek mozkový. Krysám z kontrolní skupiny byla pouze otevřena lebka, ale nic nebylo odebráno (aby byl snížen rozptyl způsobený otevřením lebky).

Protože operaci prováděl nezkušený operatér, některé z krys zahynuly přímo na operačním stole. V experimentální skupině přežilo pět krys, v kontrolní skupině sedm. Psycholové byli rozmrzelí nad nezkušeným studentem, ale pokračovali dále v experimentu. Byla získána data o počtu pohlavních spojení v průběhu jistého časového intervalu. V experimentální skupině ($N_1 = 5$): $0, 1, 4, 4, 1$. Odtud $\overline{x_1} = 2$, $ss_1 = \sum x_i^2 - \frac{(\sum x_i)^2}{N_1} = 14$, $\nu_1 = N_1 - 1 = 4$, pak $\overline{s_1^2} = \frac{ss_1}{\nu_1} = \frac{14}{4} = 3{,}5$.

V kontrolní skupině ($N_2 = 7$): $5, 7, 4, 3, 4, 6, 6$. Odtud $\overline{x_2} = 5$, $ss_2 = \sum x_i^2 - \frac{(\sum x_i)^2}{N_2} = 12$, $\nu_2 = N_2 - 1 = 6$, pak $\overline{s_2^2} = \frac{ss_2}{\nu_2} = \frac{12}{6} = 2$.

Odhad vnitřního rozptylu: Podobně jako u předchozího příkladu (vyjdeme z platnosti principu homogenního rozptylu), i nyní vypočteme jakýsi jeden odhad rozptylu jako průměr odhadů $\overline{s_1^2}$, $\overline{s_2^2}$ – nebude se jednat ovšem o aritmetický průměr, ale o tzv. **vážený průměr**.

Protože odhad $\overline{s_2^2}$ *byl sestaven na základě většího počtu stupňů volnosti (většího počtu měření), budeme mu přikládat větší váhu:*

$$est\ \sigma^2 = \frac{\nu_1}{\nu_1+\nu_2}\cdot\overline{s_1^2} + \frac{\nu_2}{\nu_1+\nu_2}\cdot\overline{s_2^2} \tag{11.24}$$

V našem příkladu $est\ \sigma^2 = \frac{4}{10}\cdot 3{,}5 + \frac{6}{10}\cdot 2 = 2{,}6$. *Pak intervaly spolehlivosti pro jednotlivé střední hodnoty a* $t_k(\nu = 10, \alpha = 2q = 0{,}05) = 2{,}228$ *jsou*

$$\mu_1 \in 2 \pm \sqrt{\frac{2{,}6}{5}}\cdot 2{,}228 \doteq 2 \pm 1{,}61 = (0{,}39; 3{,}61);$$

$$\mu_2 \in 5 \pm \sqrt{\frac{2{,}6}{7}}\cdot 2{,}228 \doteq 2 \pm 1{,}36 = (3{,}64; 6{,}36).$$

Intervaly nemají společný průnik, což znamená, že testovaná hypotéza o rovnosti středních hodnot bude zamítnuta. Dále je možné si všimnout, že interval spolehlivosti pro μ_2 *má menší délku než interval pro* μ_1 *– to je dáno větším počtem měření ve druhé skupině, pak totiž ve druhé skupině při odhadu rozptylu průměru dělíme sedmi, nikoli pěti.*

V testu, který bude následovat, budeme používat rozdělení $\overline{X_1} - \overline{X_2}$. *Pokud bychom chtěli najít 95%-ní interval spolehlivosti pro rozdíl* $\mu_1 - \mu_2$, *použijeme střed* $\overline{x_1} - \overline{x_2} = 2 - 5 = -3$ *a odhad rozptylu*

$$est\ \sigma^2_{\overline{X_1}-\overline{X_2}} = est\ \sigma^2_{\overline{X_1}} + est\ \sigma^2_{\overline{X_2}} = \frac{2{,}6}{5} + \frac{2{,}6}{7} \doteq 0{,}891.$$

Pak

$$\mu_1 - \mu_2 \in -3 \pm \sqrt{0{,}891}\cdot 2{,}228 \doteq -3 \pm 2{,}1 = (-5{,}1; -0{,}9).$$

Příslušný statistický test:

(K1) H_0: $\mu_1 = \mu_2$ *(odstranění podvěsku mozkového nemá vliv na sexuální chování);*
H_1: $\mu_1 < \mu_2$ *(odstranění podvěsku mozkového povede ke snížení sexuální aktivity).*

(K2) *Testovým kritériem bude veličina* $\overline{X_1} - \overline{X_2}$.

(K3) *Při platnosti* H_0 *má veličina*

$$\frac{\overline{X_1} - \overline{X_2} - 0}{\sqrt{est\ \sigma^2_{\overline{X_1}-\overline{X_2}}}} = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{0{,}891}}$$

Studentovo t-rozdělení pro $\nu = 10$.

(K4) *Pro* $\alpha = 0{,}05$ *příslušná kritická hodnota* t_k *je na průsečíku řádku* $\nu = 10$ *a sloupce pro* $\alpha = q = 0{,}05$, *tj.* $t_k = 1{,}812$. **Pozor, intervaly spolehlivosti budeme vždy konstruovat pro $\alpha = 2q$, ovšem u jednostranného statistického testu musíme vzít hodnotu t_k pro $\alpha = q$!!**

(K5) *Odpovídající t-hodnota kritéria je* $\frac{-3}{\sqrt{0{,}891}} = -3{,}178 \notin (-t_k; t_k)$. H_0 *tedy zamítáme.*

11.6 Předpoklady použitelnosti parametrických testů

Při odvozování testů (zejména t-testu – odvození je mimo rámec tohoto kursu) muselo být učiněno několi předpokladů:

1. Naměřené hodnoty x_i jsou navzájem nezávislé (= předpoklad nezávislosti) – například předpoklad nezávislosti v příkladu 11.10 je porušen, pokud členové skupiny podvádějí a opisují jeden od druhého.
2. Měřená veličina má normální rozdělení (= předpoklad normality).
3. Rozptyl uvnitř experimentální skupiny je stejný jako rozptyl uvnitř kontrolní skupiny, tj. oba rozptyly jsou odhadem stejného (vnitřního) rozptylu σ^2 celé populace (= princip homogenního rozptylu).

Testy, které splňují uvedené tři předpoklady, se nazývají **parametrické testy**. Pokud některý z předpokladů není splněn, kritérium použité v testu nemá rozdělení t (nebo U, pokud známe rozptyl σ^2), a tudíž nelze určit kritické hodnoty – pro srovnání středních hodnot se v tom případě užívají tzv. **neparametrické testy**. Na tu v tomto kursu už nebude moc prostoru – jsou to například Mannův-Whitneův test, Friedmanův test, Wilcoxonův test, znaménkový test, Kruskalův-Wallisův test, Spearmanův test korelace a další.

Úplná diskuse použitelnosti parametrických testů U, t (a v následující kapitole testu F) je mimo rámec tohoto kursu. Ovšem následující poznámky mohou být důležitým vodítkem, který pro běžného „uživatele“ statistiky dostačuje:

a) Mějte se na pozoru pouze tehdy, když t-hodnota kritéria je blízká hodnotě t_k. Přepoklady by musely být porušeny velmi hrubě, aby například při jednostanném testu pro $\alpha = 0{,}05$ kritická hodnota t_k „usekla“ ve skutečnosti významně více než 5% obsahu podgrafu hustoty t – i při porušených předpokladech tato kritická hodnota většinou usekne 6 nebo 7 procent, a nikoli třeba 15 procent obsahu. Proto **pokud t-hodnota kritéria přesáhne hodnotu t_k výrazně, je rozhodnutí zamítnout H_0 celkem bezpečné**.

b) Zkontrolujte své rozdělení. Nákres histogramu rozdělení je užitečný jak pro ověření předpokladu normality, tak ověření principu homogenního rozptylu.

c) Porovnejte $\overline{s_1^2}$ a $\overline{s_2^2}$. pokud se hodnoty obou odhadů liší výrazně, znamená to porušení předpokladu homogenního rozptylu. Ale pokud podíl těchto odhadů je menší než 4 při přibližně stejné délce obou souborů měření, není třeba dělat paniku.

d) Porušení předpokladů nemá takový dopad, pokud $N_i \geq 20$ a $N_1 \doteq N_2$.

e) Pozor na porušení měřítka. Některé proměnné jsou bezproblémové, např. X = průměrná rychlost (v km za hodinu), protože rozsah stupnice 10 až 20 km má stejnou váhu jako rozsah 70km až 80km.

Ale existují pochybné proměnné v tom smyslu, že měřítko porušují – například $Y =$ ohodnocení otázky v dotazníku počtem bodů ze stupnice 1 až 7. Zde totiž rozsah 3 až 4 body nemusí být ekvivalentní rozsahu stupnice 6 až 7 bodů. Taková stupnice je hodně subjektivní, nemá pevně vymezený absolutní přírůstek, a proto může vést ke zkreslenému tvaru rozdělení.

f) Pokud některý z předpokladů je porušen do té míry, že U-test nebo t-test nelze použít, je možné zkusit neparametrické testy.

11.7 Shrnutí

Tato kapitola je klíčovou kapitolou z kapitol 10,11,12, protože obsahuje nejčastěji prováděný test porovnání dvou souborů měření – t-test – a také odvozuje a ilustruje význam intervalů spolehlivosti.

Úvodem v oddílu 11.1 jsou zopakovány některé vzorce, zejména vzorce pro $\overline{X}$, S^2 – a dále je vysvětlen a odvozen vzorec pro $\overline{S^2}$ jako nejlepší nestranný odhad neznámého rozptylu σ^2 normálního rozdělení.

t-test používáme v situacích analogických U- testu (= u veličiny s normálním rozdělením) s tím jediným rozdílem, že totiž není znám rozptyl $\sigma^2_{\overline{X}}$ a musíme jej odhadnout pomocí $\frac{\overline{S^2}}{N}$. Pro odhad rozptylu průměru je důležitý zejména vzorec 11.13, na který je potřeba nezapomenout.

Celá kapitola předpokládá, že čtenář už ví, co je to statistický test – přesto ty důležité informace ohledně statistických testů opakuje (poznámky 11.3 objasňují terminologii a otázku přístupu ke statistickému testu).

Protože téměř všechno se vším souvisí, zamítne statistický test, který je dostatečně silný, hypotézu H_0 prakticky vždy – z toho důvodu je někdy lepší hledat informace nikoli ve statistickém testu, ale v intervalu spolehlivosti (blíže viz poslední poznámku v 11.4.3).

Vyvrcholením kapitoly jsou testy typu $\mu_1 = \mu_2$, které opět nemohou pozorného studenta překvapit, protože test podobného typu byl probrán v předchozí kapitole. Ovšem věcí novou je představení rozdílu mezi párovým a nepárovým testem.

11.8 Otázky k opakování

U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.

Otázka 11.1 *Výběrový průměr $\overline{X}$ je nestranným a konzistentním odhadem parametru μ měření veličiny s normálním rozdělením.*

Otázka 11.2 *Statistika S^2 je nestranným a konzistentním odhadem parametru σ^2 měření veličiny s normálním rozdělením.*

Otázka 11.3 *Kritická t- hodnota je blíže počátku než odpovídající (= pro stejnou hladinu významnosti sestrojená) kritická U-hodnota.*

Otázka 11.4 *Pro rostoucí počet stupňů volnosti graf hustoty t-rozdělení stále více splývá s grafem hustoty normovaného normálního rozdělení.*

Otázka 11.5 *Je možné, že interval spolehlivosti pro střední hodnotu* μ *průměru* $\overline{X}$ *sestrojený na základě konkrétní realizace* $\overline{x}$ *tuto střední hodnotu vůbec neobsahuje.*

Otázka 11.6 *Pokud hodnota* μ_0 *nepadne do oboustranného intervalu spolehlivosti pro střední hodnotu* μ *průměru* $\overline{X}$*, znamená to, že oboustranný test pro* $H_0 : \mu = \mu_0$ *tuto hypotézu* H_0 *nezamítne.*

Otázka 11.7 *Párovým testem je každý test typu* $\mu_1 = \mu_2$*, který porovnává střední hodnoty dvou náhodných veličin.*

Otázka 11.8 *Vnitřní rozptyl popisuje rozmanitost konkrétních jedinců v populaci – tato rozmanitost přitom nezávisí na podmínkách, při kterých je měření prováděno (je stejná u kontrolní skupiny i experimentální skupiny).*

Otázka 11.9 *Celkový rozptyl je součtem vnitřního rozptylu a vnějšího rozptylu.*

Otázka 11.10 *Při nestejných délkách obou souborů měření lze vnitřní rozptyl odhadnout jako aritmetický průměr rozptylů* $\overline{s_1^2}$*,* $\overline{s_2^2}$ *jednotlivých souborů měření.*

Otázka 11.11 *Existují situace, kdy neplatí princip homogenního rozptylu (rozptyl měření v kontrolní skupině je nesrovnatelně větší než rozptyl měření v experimentální skupině).*

Odpovědi na otázky viz 13.5.

12 Testy χ^2 (chí kvadrát), Mannův-Whitneyův test

12.1 Vlastnosti rozdělení χ^2

Uvažujme veličinu X s normálním rozdělením pravděpodobnosti $No(\mu, \sigma^2)$. Víme, že veličina $U = \frac{X-\mu}{\sigma}$ má normované normální rozdělení $No(0;1)$. Jaké rozdělení má veličina

$$U^2 = \frac{(X-\mu)^2}{\sigma^2}?$$

Především můžeme říci, že hodnoty veličiny U^2 jsou nezáporné. Dále víme, že asi 68% veličiny U leží v intervalu $\langle -1; 1\rangle$, tj. asi 68% veličiny U^2 leží v intervalu $\langle 0; 1\rangle$. Zkrátka a dobře, vlastnosti veličiny U^2 lze odvodit z vlastností veličiny U. A protože se veličina U^2 ve statistice hojně používá, dostala i své jméno: $\chi^2(1)$ ... čti: chí kvadrát o jednom stupni volnosti.

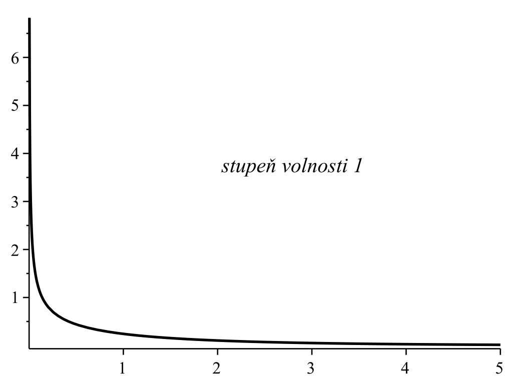
Označení $\chi^2(1)$ naznačuje, že může nastat i více stupňů volnosti než jeden. A skutečně, pokud dvě hodnoty U_1, U_2 veličiny U umocníme na druhou mocninu a sečteme, dostaneme obecně větší hodnotu než $\chi^2(1)$, označujeme ji $\chi^2(2)$ (veličina „chí kvadrát“ se dvěma stupni volnosti):

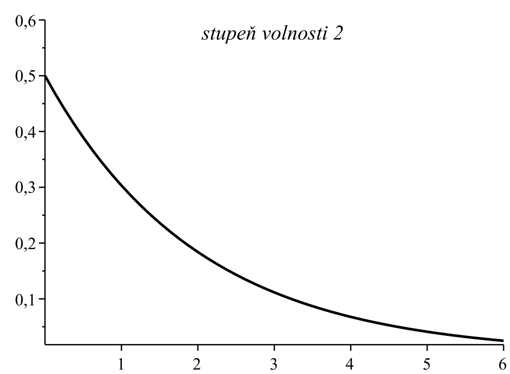
$$\chi^2(2) = U_1^2 + U_2^2.$$

Čili střední hodnota veličiny $\chi^2(2)$ je větší než střední hodnota veličiny $\chi^2(1)$. Obecně lze pak vyjádřit veličinu o n stupních volnosti:

$$\chi^2(n) = U_1^2 + U_2^2 + \cdots + U_n^2.$$

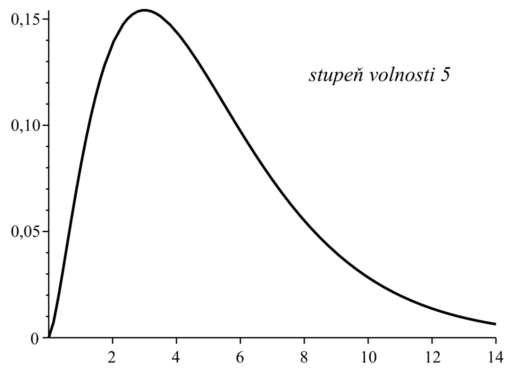
Příklady χ^2 o různých stupních volnosti jsou na obrázku:

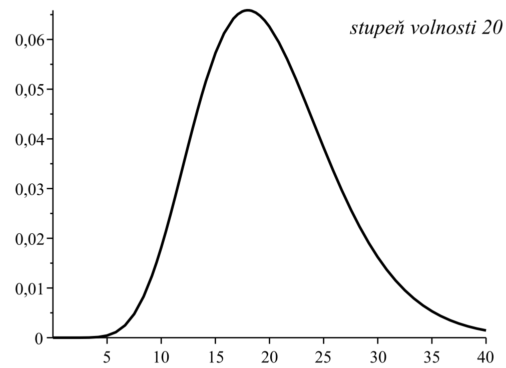




Obecný vzorec hustoty rozdělení $\chi^2(n)$ zní

$$f(x) = \begin{cases} 0 & \dots \quad x \le 0; \\ \frac{1}{2^{\frac{n}{2}} \cdot \Gamma\left(\frac{n}{2}\right)} \cdot x^{\left(\frac{n}{2}-1\right)} \cdot e^{-\frac{x}{2}} & \dots \quad x > 0, \end{cases}$$





kde Γ je tzv. gama-funkce. Co se týká střední hodnoty a rozptylu rozdělení χ^2, vypočteme jen střední hodnotu, rozptyl uvedeme bez důkazu:

$$E\left\{\chi^2(n)\right\} = E(U_1^2 + U_2^2 + \cdots + U_n^2) = 1 + 1 + \cdots + 1 = n.$$

Opravdu platí $E(U_i^2) = 1$:

$$\begin{aligned} E\left(\frac{X-\mu}{\sigma}\right)^2 &= \frac{1}{\sigma^2} \cdot E(X-\mu)^2 = \frac{1}{\sigma^2}(EX^2 - 2\mu EX + \mu^2) = \\ &= \frac{1}{\sigma^2}(EX^2 - 2\mu^2 + \mu^2) = \frac{1}{\sigma^2}(\sigma^2 + \mu^2 - \mu^2) = 1. \end{aligned}$$

Pro výpočet EX^2 jsme využili faktu, že $\sigma^2 = DX = EX^2 - E^2X = EX^2 - \mu^2$, tj. $EX^2 = \sigma^2 + \mu^2$.

$$D\left\{\chi^2(n)\right\} = 2n.$$

Další důležitou vlastností je tzv. aditivita rozdělení χ^2, že totiž

$$\chi^2(n_1) + \chi^2(n_2) = \chi^2(n_1 + n_2),$$

která v podstatě plyne z toho, že $\chi^2(n) = \sum_1^n U_i^2$.

Podobně jako u jiných rozdělení, i zde byly kritické hodnoty rozdělení spočteny jednou provždy a seřazeny do tabulky, takže při konkrétním užívání kritických hodnot nemusíme počítat žádné nechutné integrály (kritické hodnoty viz tabulky 12.11, 12.12).

Například pokud chceme určit kritickou hodnotu h_k tak, že

$$P(\chi^2(10) \geq h_k) = 0{,}05,$$

tak ve druhé části tabulky najdeme hodnotu 18,307 na průsečíku řádku 10 a sloupce pro 0,05.

Pokud hledáme d_k tak, aby

$$P(\chi^2(10) \leq d_k) = 0{,}05,$$

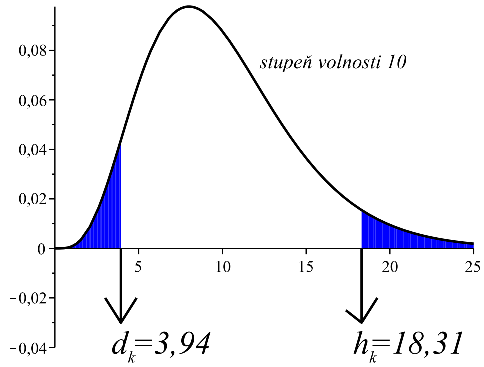
tak v první části tabulky na průsečíku řádku 10 a sloupce 0,95 (protože v tabulce jsou kritické hodnoty uvedené pro pravou část podgrafu) najdeme hodnotu 3,94 (viz též obrázek):

Tabulka 12.11: Kritické hodnoty jednostranného testu χ^2 – část 1.

$q \rightarrow$ $\nu \downarrow$	0,995	0,990	0,975	0,950	0,900	0,750	0,500
1	$392{,}704 \cdot 10^{-10}$	$157{,}088 \cdot 10^{-9}$	$982{,}069 \cdot 10^{-9}$	$393{,}214 \cdot 10^{-8}$	0,0157908	0,1015308	0,454937
2	0,0100251	0,0201007	0,0506356	0,102587	0,210720	0,575364	1,38629
3	0,0717212	0,114832	0,215795	0,351846	0,584375	1,212534	2,36597
4	0,206990	0,297110	0,484419	0,710721	1,063623	1,92255	3,35670
5	0,411740	0,554300	0,831211	1,145476	1,61031	2,67460	4.35146
6	0,675727	0,872085	1,237347	1,63539	2,20413	3,45460	5,34812
7	0,989265	1,239043	1,68987	2,16735	2,83311	4,25485	6,34581
8	1,344419	1,646482	2,17973	2,73264	3,48954	5,07064	7,34412
9	1,734926	2,087912	2,70039	3,32511	4,16816	5,89883	8,34283
10	2,15585	2,55821	3,24697	3,94030	4,86518	6,73720	9,34182
11	2,60321	3,05347	3,81575	4,57481	5,57779	7,58412	10,3410
12	3,07382	3,57056	4,40379	5,22603	6,30380	8,43842	11,3403
13	3,56503	4,10691	5,00874	5,89186	7,04150	9,29906	12,3398
14	4,07468	4,66043	5,62872	6,57063	7,78953	10,1653	13,3393
15	4,60094	5,22935	6,26214	7,26094	8,54675	11,0365	14,3389
16	5,14224	5,81221	6,90766	7,96164	9,31223	11,9122	15,3385
17	5,69724	6,40776	7,56418	8,67176	10,0852	12,7919	16,3381
18	6,26481	7,01491	8,23075	9,39046	10,8649	13,6753	17,3379
19	6,84398	7,63273	8,90655	10,1170	11,6509	14,5620	18,3376
20	7,43386	8,26040	9,59083	10,8508	12,4426	15,4518	19,3374
21	8,03366	8,89720	10,28293	11,5913	13,2396	16,3444	20,3372
22	8,64272	9,54249	10,9823	12,3380	14,0415	17,2396	21,3370
23	9,26042	10,19567	11,6885	13,0905	14,8479	18,1373	22,3369
24	9,88623	10,8564	12,4011	13,8484	15,6587	19,0372	23,3367
25	10,5197	11,5240	13,1197	14,6114	16,4734	19,9393	24,3366
26	11,1603	12,1981	13,8439	15,3791	17,2919	20,8434	25,3364
27	11,8076	12,8786	14,5733	16,1513	18,1138	21,7494	26,3363
28	12,4613	13,5648	15,3079	16,9279	18,9392	22,6572	27,3363
29	13,1211	14,2565	16,0471	17,7083	19,7677	23,5666	28,3362
30	13,7867	14,9535	16,7908	18,4926	20,5992	24,4776	29,3360
40	20,7065	22,1643	24,4331	26,5093	29,0505	33,6603	39,3354
50	27,9907	29,7067	32,3574	34,7642	37,6886	42,9421	49,3349
60	35,5346	37,4848	40,4817	43,1879	46,4589	52,2938	59,3347
70	43,2752	45,4418	48,7576	51,7393	55,3290	61,6983	69,3344
80	51,1720	53,5400	57,1532	60,3915	64,2778	71,1445	79,3343
90	59,1963	61,7541	65,6466	69,1260	73,2912	80,6247	89,3342
100	67,3276	70,0648	74,2219	77,9295	82,3581	90,1332	99,3341

Tabulka 12.12: Kritické hodnoty jednostranného testu χ^2 – část 2.

$q \rightarrow$ $\nu \downarrow$	0,250	0,100	0,050	0,025	0,010	0,005	0,001
1	1,32330	2,70554	3,84146	5,02389	6,63490	7,87944	10,828
2	2,77259	4,60517	5,99147	7,37776	9,21034	10,5966	13,816
3	4,10835	6,25139	7,81473	9,34840	11,3449	12,8381	16,266
4	5,38527	7,77944	9,48773	11,1433	13,2767	14,8602	18,467
5	6,62568	9,23635	11,0705	12,8325	15,0863	16,7496	20,515
6	7,84080	10,6446	12,5916	14,4494	16,8119	18,5476	22,458
7	9,03715	12,0170	14,0671	16,0128	18,4753	20,2777	24,322
8	10,2188	13,3616	15,5073	17,5346	20,0902	21,9550	26,125
9	11,3887	14,6837	16,9190	19,0228	21,6660	23,5893	27,877
10	12,5489	15,9871	18,3070	20,4831	23,2093	25,1882	29,588
11	13,7007	17,2750	19,6751	21,9200	24,7250	26,7569	31,264
12	14,8454	18,5494	21,0261	23,3367	26,2170	28,2995	32,909
13	15,9839	19,8119	22,3621	24,7356	27,6883	29,8194	34,528
14	17,1170	21,0642	23,6848	26,1190	29,1413	31,3193	36,123
15	18,2451	22,3072	24,9958	27,4884	30,5779	32,8013	37,697
16	19,3688	23,5418	26,2962	28,8454	31,9999	34,2672	39,252
17	20,4887	24,7690	27,5871	30,1910	33,4087	35,7185	40,790
18	21,6049	25,9894	28,8693	31,5264	34,8053	37,1564	42,312
19	22,7178	27,2036	30,1435	32,8523	36,1908	38,5822	43,820
20	23,8277	28,4128	31,4104	34,1696	37,5662	39,9968	45,315
21	24,9348	29,6151	32,6705	35,4789	38,9321	41,4010	46,797
22	26,0393	30,8133	33,9244	36,7807	40,2894	42,7956	48,268
23	27,1413	32,0069	35,1725	38,0757	41,6384	44,1813	49,728
24	28,2412	33,1963	36,4151	39,3641	42,9798	45,5585	51,179
25	29,3389	34,3816	37,6525	40,6465	44,3141	46,9278	52,620
26	30,4345	35,5631	38,8852	41,9232	45,6417	48,2899	54,052
27	31,5284	36,7412	40,1133	43,1944	46,9630	49,6449	55,476
28	32,6205	37,9159	41,3372	44,4607	48,2782	50,9933	56,892
29	33,7109	39,0875	42,5569	45,7222	49,5879	52,3356	58,302
30	34,7998	40,2560	43,7729	46,9792	50,8922	53,6720	59,703
40	45,6160	51,8050	55,7585	59,3417	63,6907	66,7659	73,402
50	56,3336	63,1671	67,5048	71,4202	76,1539	79,4900	86,661
60	66,9814	74,3970	79,0819	83,2976	88,3794	91,9517	99,607
70	77,5766	85,5271	90,5312	95,0231	100,425	104,215	112,317
80	88,1303	96,5782	101,879	106,629	112,329	116,321	124,839
90	98,6499	107,565	113,145	118,136	124,116	128,299	137,208
100	109,141	118,498	124,342	129,561	135,807	140,169	149,449



12.2 Využití rozdělení χ^2

12.2.1 Testování hypotézy $\sigma^2 = konst$

Příklad 12.1 *Zajímá nás, zda jistý výukový program je schopen naučit jisté partie aritmetiky při výuce na ZŠ. Může se totiž stát, že nadaní studenti budou chápat instrukce počítače, kdežto průměrní žáci budou mít potíže s programem komunikovat a naučí se méně, než by pochopili z výkladu živého učitele.*

Je známo, že rozptyl výsledků testu z aritmetiky při klasické výuce je $\sigma_0^2 = 25$. *Pokud by naše obava byla oprávněná, rozptyl výsledků znalostí by byl při použití programu větší (nadaní žáci by byli lepší, průměrní žáci horší než obvykle).*

Byl proveden experiment, kdy deset žáků se podrobilo programové výuce. Výsledky znalostí (bodové ohodnocení) byly: 68, 90, 70, 91, 72, 80, 85, 82, 91, 95. *Odtud*

$$\sum_1^{10} (x_i - \overline{x})^2 = 846{,}4$$

a odhad est $\sigma^2 = 94{,}04$ *populačního roptylu je mnohem větší než* $\sigma_0^2 = 25$. *Proveďte statistický test, který by potvrdil, že ke zvýšení rozptylu nedošlo náhodou.*

K1: H_0: rozptyl výsledků znalostí σ^2 nabytých počítačovu výukou je stejný jako rozptyl $\sigma_0^2 = 25$ při klasické výuce (tj. $\sigma^2 = 25$).

H_1: $\sigma^2 \neq 25$ (oboustranný test).

K2: Ukazuje se (uvidíme v bodě K3), že vhodným kritériem testu je $\frac{1}{\sigma_0^2} \cdot \sum_1^{10} (x_i - \overline{x})^2$.

K3: Za předpokladu platnosti H_0 má výraz $\frac{1}{\sigma_0^2} \cdot \sum_1^{10} (x_i - \overline{x})^2$ rozdělení $\chi^2(9)$.

Skutečně, dokažme tento fakt.

Víme, že $\frac{1}{\sigma_0^2} \cdot \sum_1^{n} (x_i - \mu)^2$ má rozdělení $\chi^2(n)$, protože se jedná o součet n čtverců rozdělení U. Ovšem μ neznáme – jaký je tedy vztah mezi $\frac{1}{\sigma_0^2} \cdot \sum_1^{n} (x_i - \mu)^2$ a $\frac{1}{\sigma_0^2} \cdot \sum_1^{10} (x_i - \overline{x})^2$?

Platí

$$\sum (x_i - \mu)^2 = \sum (x_i - \overline{x} + \overline{x} - \mu) = \sum (x_i - \mu)^2 + 2 \cdot \sum (x_i - \overline{x}) \cdot (\overline{x} - \mu) + \sum (\overline{x} - \mu)^2.$$

A protože

$$2 \cdot \sum (x_i - \overline{x}) \cdot (\overline{x} - \mu) = 2(\overline{x} - \mu) \cdot \underbrace{\sum (x_i - \overline{x})}_{=0} = 0,$$

dostaneme

$$\sum (x_i - \overline{x})^2 = \sum (x_i - \overline{x})^2 + \sum (\overline{x} - \mu)^2.$$

Máme tedy rovnici

$$\sum (x_i - \overline{x})^2 = \sum (x_i - \overline{x})^2 + n \cdot (\overline{x} - \mu)^2,$$

kterou vynásobením $\frac{1}{\sigma_0^2}$ lze převést na tvar

$$\underbrace{\frac{\sum (x_i - \mu)^2}{\sigma_0^2}}_{=\chi^2(n)} = \frac{\sum (x_i - \overline{x})}{\sigma_0^2} + \underbrace{\frac{(\overline{x} - \mu)^2}{\frac{\sigma_0^2}{n}}}_{=\chi^2(1)}.$$

To, že člen na levé straně má rozdělení $\chi^2(n)$, už bylo řečeno, člen na pravé straně má rozdělení $\chi^2(1)$, protože se jedná o druhou mocninu normovaného rozdělení průměru $\overline{x}$ se střední hodnotou μ a rozptylem $\frac{\sigma_0^2}{n}$. Dohromady tedy dostáváme to, co jsme chtěli spočítat a dokázat:

$$\frac{x_i - \overline{x}}{\sigma_0^2} = \chi^2(n) - \chi^2(1) = \chi^2(n-1).$$

K4: Pro $\alpha = 0{,}05$ usekáváme na obou stranách hodnotu 0,025 z obsahu hustoty, tj. $d_k(9) = 2{,}70039$ (na průsečíku řádku 9 a sloupce 0,975) a $h_k(9) = 19{,}0228$ (na průsečíku řádku 9 a sloupce 0,025).

K5: Po dosazení naměřených hodnot

$$\frac{x_i - \overline{x}}{\sigma_0^2} = \frac{846{,}4}{25} = 33{,}86 \notin (2{,}70039; 19{,}0228) \quad \Rightarrow \quad \text{zamítáme } H_0,$$

rozptyl je (bohužel) významně odlišný od 25, v našem případě významně větší. A to je tragédie počítačové výuky :-)

V případě oboustranného testu – kdybychom neměli teoretický podklad, že rozptyl poroste, ale bylo by stejně možné, že třeba i klesne – bychom našli dvě kritické hodnoty na řádku 9, a sice $\chi_m = 2{,}7$ (ve sloupci 0,975) a $\chi_v = 19{,}02$ (ve sloupci 0,025) a H_0 bychom zamítli na hladině významnosti $\alpha = 0{,}05$, pokud by hodnota kritéria ležela mimo interval $(2{,}7; 19{,}02)$.

12.2.2 Test druhu rozdělení – χ^2 test dobré shody

Příklad 12.2 *Rodičům se narodily už čtyři dcery, a přesto by si přáli i syna. Chystají se mít páté dítě s nadějí, že pravděpodobnost narození chlapce je* 0,5 *(tj. že rození dětí má podobný charakter – co se týká pohlaví dítěte – jako házení korunou). Ale přece jen si vyhledali data o* 1024 *rodinách s pěti dětmi a zjistili, v kolika rodinách se narodilo kolik chlapců:*

0 *chlapců*	...	40 *rodin*
1 *kluk*	...	184 *rodin*
2 *kluci*	...	300 *rodin*
3 *kluci*	...	268 *rodin*
4 *kluci*	...	196 *rodin*
5 *kluků*	...	36 *rodin*

Pokud tato data o pohlaví při rození dětí mají stejný charakter jako házení korunou, pak je lze dobře popsat binomickým rozdělením s parametry $N = 1024$, $p = 0{,}5$. *Teoretické rozdělení pravděpodobnosti a rozdělení četnosti mají následující průběh:*

počet chlapců x_i	p_i	*četnost* f_i
0	*1/32*	32
1	*5/32*	160
2	*10/32*	320
3	*10/32*	320
4	*5/32*	160
5	*1/32*	32

Otázka zní: do jaké míry se shoduje empirické rozdělení četnosti a teoretické rozdělení četnosti, čili: lze nashromážděná data dobře popsat binomickým rozdělením s uvedenými parametry?

Provedeme tzv. **test dobré shody** *(v angličtině: good fit test ... GFT).*

K1: H_0: počet chlapců z pěti narozených dětí lze dobře popsat binomickým rozdělením pro $p = 0{,}5$.

H_1: Nelze.

K2: Jaké kritérium zvolit? Vezmeme součet čtverců rozdílů normalizovaných odchylek naměřené a teoretické četnosti $\sum_1^k \frac{(f_t - f_m)^2}{f_t}$:

počet chlapců	naměřená četnost	teoretická četnost	$(f_t - f_m)^2$	$\frac{(f_t-f_m)^2}{f_t}$
0	40	32	64	2
1	184	160	576	3,6
2	300	320	400	1,25
3	268	320	2704	8,45
4	196	160	1296	8,1
5	36	32	16	0,5

K3: Za předpokladu platnosti H_0 má kritérium z bodu K2 rozdělení $\chi^2(k-1)$. Tento fakt nebudeme dokazovat.

K4: Kritérium je vhodným reprezentantem míry platnosti H_0: pokud H_0 platí, očekáváme, že hodnota kritéria bude malá, pokud neplatí, bude velká. Jedná se tedy o jednostranný test, kde k je počet skupin četnosti, tj. v našem případě $k = 6$. Tedy pro $\alpha = 0{,}05$ je $\chi_k(5) = 11{,}0705$ („usekáváme" přet procent obsahu hustoty psti zprava).

K5:

$$\sum_1^6 \frac{(f_m - f_t)^2}{f_t} = 23{,}9 > 11{,}07 \Rightarrow \text{ zamítáme } H_0,$$

binomické rozdělení není příliš dobré pro popis našich dat (tj. pohlaví chlapců při narození se nechová jako počet líců při hodu korunou).

12.2.3 Testování nezávislosti v kontingenční tabulce

Příklad 12.3 *Zajímá nás, zda existuje vztah mezi názory na jadernou energii a politickou příslušností. Proto jsme se dotázali nezávisle vybraných* 200 *lidí, jaký mají názor na jadernou elektrárnu Temelín (neměla by být v provozu – nezajímá mě to – měla by být v provozu), a dále které ze stran ODS, ČSSD dávají větší přednost. Výsledky průzkumu byly sestaveny do kontingenční tabulky (čísla v závorkách vyjadřují empirické pravděpodobnosti = četnosti vydělené číslem* 200*):*

	elektr. by neměla být	nevím	elektr. by měla být	
ČSSD	40 (0,2)	70 (0,35)	40 (0,2)	150 (0,75)
ODS	35 (0,175)	5 (0,025)	10 (0,05)	50 (0,25)
	75 (0,375)	75 (0,375)	50 (0,25)	200 (1,000)

Při testování, zda existuje závislost mezi oběma proměnnými, můžeme použít test χ^2:

K1: H_0: Obě proměnné se chovají nezávisle, čili empirické rozdělení je hodně blízké následujícímu teoretickému rozdělení, kde poslední řádek a poslední sloupec jsou stejné jako v předchozí tabulce (udávající výsledky průzkumu), ovšem ostatní pravděpodobnosti jsou získány vynásobením příslušných pravděpodobností v posledním řádku a sloupci; označme $A_1 \ldots$ ČSSD ($P(A_1) = 0{,}75$); $A_2 \ldots$ ODS ($P(A_2) = 0{,}25$); $B_1 \ldots$ elektrárna NE ($P(B_1) = 0{,}375$); $B_2 \ldots$ nevím ($P(B_2) = 0{,}375$); $B_3 \ldots$ elektrárna ANO ($P(B_3) = 0{,}25$). Nyní pokud A_i, B_j jsou nezávislé, tak

$$P(A_i \cap B_j) = P(A_i) \cdot P(B_j)$$

(například $P(A_1 \cap B_1) = P(A_1) \cdot P(B_1) = 0{,}281$, atd.). Dostáváme tedy tabulku teoretických pravděpodobností (uvedených v závorce), četnosti jsou získány vynásobením příslušné pravděpodobnosti číslem 200 (četnosti tedy nemusí být celočíselné):

	elektr. by neměla být	nevím	elektr. by měla být	
ČSSD	56,2 (0,281)	56,2 (0,281)	37,6 (0,188)	150 (0,75)
ODS	18,8 (0,094)	18,8 (0,094)	12,4 (0,062)	50 (0,25)
	75 (0,375)	75 (0,375)	50 (0,25)	200 (1,000)

H_1: Obě proměnné jsou závislé, čili nelze je dost dobře popsat příslušným teoretickým rozdělením.

K2: Kritériem bude $\sum_1^k \frac{(f_t-f_m)^2}{f_t}$, kde k je počet různých tříd četností (v našem případě počet různých oken tabulky kromě posledního řádku a posledního sloupce: $k = 2\cdot 3 = 6$), f_t jsou příslušné teoretické četnosti (=četnosti z poslední tabulky), f_m příslušné naměřené četnosti (z předposlední tabulky).

K3: Za předpokladu platnosti H_0 má kriteriální funkce rozdělení $\chi^2((J-1)(K-1))$, kde J je počet podmínek veličiny A, K je počet podmínek veličiny B (tento fakt nebudeme dokazovat). V našem případě $\chi^2(1\cdot 2) = \chi^2(2)$.

K4: Pro $\alpha = 0{,}05$ kritická hodnota $\chi_k(2) \doteq 5{,}99$.

K5:

$$\sum_1^6 \frac{(f_t-f_m)^2}{f_t} = \frac{(56{,}2-40)^2}{56{,}2} + \frac{(56{,}2-70)^2}{56{,}2} + \cdots + \frac{(12{,}4-10)^2}{12{,}4} = 32{,}76;$$

toto číslo zdaleka přesahuje kritickou hodnotu testu 5,99, tedy H_0 zamítáme, prokázala se závislost obou veličin.

12.3 Neparametrický Mannův-Whitneyův test podle pořadí

Příklad 12.4 *Vraťme se k příkladu 12.1, kde jsme chtěli zjistit, jaký vliv na studenty bude mít počítačová výuka matematiky. Proveďme nyní experiment jiného rázu: náhodně vybraných* 24 *studentů rozdělíme na dvě skupiny po dvanácti lidech. Jedna skupina se účastnila výuky pod dohledem učitele, druhá příslušné počítačové výuky. Potom se žáci podrobili písemce, kde se měřil čas potřebný na vyřešení zadaných úloh. Získala se data:*

$$1 \ldots\ \textit{běžná výuka: } 43, 12, 21, 41, 39, 23, 27, 37, 35, 31, 33, 29;$$

$$2 \ldots\ \textit{počítačová výuka: } 3, 13, 1, 5, 11, 10, 9, 8, 6, 2, 4, 7;$$

Pomocí všech naměřených hodnot odhadneme rozptyl:

$$est\ \sigma^2 = \frac{SS_1 + SS_2}{V_1 + V_2} = \frac{908{,}92 + 154{,}92}{11 + 11} = 48{,}36$$

A nyní můžeme provést t-test:

K1: H_0: $\mu_1 = \mu_2$ (doba potřebná na vyřešení písemky z aritmetiky má stejnou střední hodnotu u obou skupin);
H_1: $\mu_1 \neq \mu_2$.

K2: Kritériem je $\frac{\overline{x}_1-\overline{x}_2-\mu_{1-2}}{est\ \sigma_{1-2}}$, kde $\mu_{1-2}=0$,

$$est\ \sigma_{1-2}=\sqrt{est\ \sigma_1^2+est\ \sigma_2^2}=\sqrt{\frac{48{,}36}{12}+\frac{48{,}36}{12}}=2{,}84.$$

K3: Za předpokladu platnosti H_0 má naše kritérium rozdělení $t(22)$, protože $V_1+V_2=22$.

K4: Pro $\alpha=0{,}05$ je $t_k(22)=\pm 2{,}09$.

K5: Hodnota kritéria: $\frac{30{,}91-6{,}58}{2{,}84}=8{,}57$, a tak H_0 zamítáme.

Ovšem právě provedený t-test nesplňoval předpoklad rovnosti rozptylů v obou skupinách: $est_1\ \sigma^2=\frac{SS_1}{V_1}=82{,}63$, kdežto $est_2\ \sigma^2=\frac{SS_2}{V_2}=14{,}08$. První odhad je přibližně šestinásobkem druhého, což už přesahuje rozumnou (= čtyřnásobnou) míru. Toto hrubé porušení předpokladu t-testu je důvodem k detailnějšímu prozkoumání dat pomocí neparametrického testu. Vhodným kandidátem nyní je Mannův-Whitneyův test.

Vraťme se datům z právě opuštěného příkladu 12.4 a podrobme jej Mannovu-Whitneyovu testu:

K1: Uspořádejme všechna měření (z obou souborů dohromady) podle velikosti a zaznamenejme přitom

a) pořadí dané hodnoty;

b) příslušnost dané hodnoty k experimentální skupině (B ... běžná výuka, P ... počítačová výuka).

hodnota	1	2	3	4	5	6	7	8	9	10	11	12
pořadí	1	2	3	4	5	6	7	8	9	10	11	12
skupina	P	P	P	P	P	P	P	P	P	P	P	B

13	21	23	27	29	31	33	35	37	39	41	43
13	14	15	16	17	18	19	20	21	22	23	24
P	B	B	B	B	B	B	B	B	B	B	B

K2: Nyní vypočteme jakési míry U_1, U_2:

$$U_1=\sum_{\text{děti ze skupiny P}}\text{(počet dětí z B, které mají horší skóre než dítě } i\text{).}$$

Tedy pro dítě s hodnotou 1 na prvním místě pořadí má všech dvanáct dětí z B horší skóre. Pro dítě s hodnotou 2 na druhém místě pořadí má opět všech dvanáct dětí z B horší skóre, atd. až pro dítě s hodnotou 11 na jedenáctém místě má stále všech dvanáct dětí z B horší skóre. A konečně změna, pro dítě s hodnotou 13 na třináctém místě má už jen jedenáct dětí z B horší skóre. Dohromady

$$U_1 = \underbrace{12 + 12 + \cdots + 12}_{\text{jedenáctkrát}} + 11 = 143.$$

Podobně

$$U_2 = \sum_{\text{děti ze skupiny B}} (\text{počet dětí z P, které mají horší skóre než dítě } i).$$

Zde pouze pro dítě s hodnotou 12 na dvanáctém místě má jedno dítě z P horší skóre, tj.

$$U_1 = 1 + \underbrace{0 + 0 + \cdots + 0}_{\text{jedenáctkrát}} = 1.$$

Z konstrukce U_1, U_2 je vidět, že tyto míry zachycují závažnost, s jakou má jedna skupina lepší pořadí než druhá. Samozřejmě existují poněkud pohodlnější vzorce pro jejich výpočet:

$$U_1 = n_1 \cdot n_2 + \frac{n_2 \cdot (n_2 + 1)}{2} - \sum (\text{pořadí hodnot ze skupiny 2}),$$

$$U_2 = n_1 \cdot n_2 + \frac{n_1 \cdot (n_1 + 1)}{2} - \sum (\text{pořadí hodnot ze skupiny 1})$$

(mimochodem, platí také $U_1 + U_2 = n_1 \cdot n_2$ – v našem příkladu tedy $143 + 1 = 12 \cdot 12$). Kritériem testu bude nyní menší z obou vypočtených hodnot: $U = min(U_1, U_2)$. V našem příkladu $U = min(143, 1) = 1$.

K3: Kritická hodnota testu:

a) pro $n_1, n_2 \leq 20$: Byly sestaveny tabulky kritických hodnot našeho oboustranného testu, a to pro $\alpha = 0{,}05$ tabulka 12.13, pro $\alpha = 0{,}01$ tabulka 12.14.

b) pro $n_1, n_2 > 20$: Rozdělení kritéria U je normální se střední hodnotou $\mu = \frac{n_1 \cdot n_2}{2}$ a směrodatnou odchylkou

$$\sigma = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}},$$

čili normovanou hodnotu

$$\frac{U - \frac{n_1 \cdot n_2}{2}}{\sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}}$$

testujeme s běžnými kritickými hodnotami $\pm 1{,}96$ normálního rozdělení pro $\alpha = 0{,}05$.

K4+K5: V našem příkladu tedy stanovíme kritickou hodnotu

a) z tabulky pro $\alpha = 0{,}05$, $n_1 = n_2 = 12$: $U_k = 37 > U = 1$, tj. H_0 zamítáme ve prospěch alternativní hypotézy H_1.

Všimněte si, že na rozdíl od většiny testů v této přednášce zde zamítnutí H_0 nastane tehdy, když hodnota kritéria kritickou hodnotu NEPŘEKROČÍ. Je to dáno konstrukcí kriterijní funkce – pokud převáží malé hodnoty pořadí v jednom souboru, tak hodnota kritéria je malá, ale současně to znamená jasný a zřetelný rozdíl mezi skupinami.

b) pomocí aproximace normálním rozdělením (nyní méně přesné, ale výpočetně jednodušší):

$$\frac{U - \frac{n_1 \cdot n_2}{2}}{\sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}} = -4{,}1 < -1{,}96 \Rightarrow H_0 \text{ zamítáme}$$

ve prospěch alternativní hypotézy H_1.

Vidíme, že výsledky neparametrického testu jsou stejné jako výsledky příslušného parametrického testu 12.4 s porušenými předpoklady. Z toho je vidět, že navzdory porušeným předpokladům t-testu je i při silném rozdílu obou souborů výsledek testu správný, To ovšem nemusí být pravdou, pokud rozdíly mezi oběma skupinami nebudou tak jednoznačné, jak to dokumentují další situace v této kapitole.

12.4 Shrnutí

Cílem této kapitoly bylo představit další užitečné rozdělení pravděpodobnosti, a sice rozdělení χ^2. rozdělení $\chi^2(n)$ (o n stupních volnosti) se definuje jako součet čtverců n normovaných normálních rozdělení U^2. Ukazuje se, že mnohé veličiny v praxi jsou rozděleny jako χ^2, a tedy toto rozdělení se vyskytuje v mnoha statistických testech:

1. V příkladu 12.1 jsme viděli, že rozdělení χ^2 lze užít v testu typu $\sigma^2 = konst$. Pokud měření jsme získali z populace, jejíž rozptyl je σ_0^2, pak kritérium

$$\frac{\sum (x_i - \overline{x})^2}{\sigma_0^2}$$

má rozdělení χ^2 o počtu stupňů volnosti $(n-1)$.

2. V příkladu 12.2 jsme viděli, že χ^2 lze užít v testu, zda naměřená data odpovídají jistému teoretickému rozdělení pravděpodobnosti (toto je asi nejčastější a nejznámější využití rozdělení χ^2, kterému se říká **test dobré shody**). Pokud máme konkrétně n naměřených (či pozorovaných) četností f_m, a jim odpovídající teoretické četnosti f_t, pak kritérium

$$\frac{\sum (f_t - f_m)^2}{f_t}$$

má rozdělení χ^2 o počtu stupňů volnosti $(n-1)$.

Tabulka 12.13: Kritické hodnoty Mannova-Whitneyova testu pro $\alpha = 0{,}05$.

	n_2																	
n_1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	0	0	0	0	1	1	1	1	1	2	2	2	2
3	–	–	0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4		0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	13
5			2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6				5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7					8	10	12	14	16	18	20	22	24	26	28	30	32	34
8						13	15	17	19	22	24	26	29	31	34	36	38	41
9							17	20	23	26	28	31	34	37	39	42	45	48
10								23	26	29	33	36	39	42	45	48	52	55
11									30	33	37	40	44	47	51	55	58	62
12										37	41	45	49	53	57	61	65	69
13											45	50	54	59	63	67	72	76
14												55	59	64	67	74	78	83
15													64	70	75	80	85	90
16														75	81	86	92	98
17															87	93	99	105
18																99	106	112
19																	113	119
20																		127

Tabulka 12.14: Kritické hodnoty Mannova-Whitneyova testu pro $\alpha = 0{,}01$.

								n_2										
n_1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–
2	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	–	0	0
3	–	–	–	–	–	–	0	0	0	1	1	1	2	2	2	2	3	3
4		–	–	0	0	1	1	2	2	3	4	4	5	5	6	6	7	8
5			0	1	2	3	3	4	5	6	7	7	8	9	10	11	12	13
6				2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7					4	6	7	9	10	12	13	15	16	18	19	21	22	24
8						8	10	12	14	16	18	19	21	23	25	27	29	31
9							11	13	16	18	20	22	24	27	29	31	33	36
10								16	18	21	24	26	29	31	34	37	39	42
11									21	24	27	30	33	36	39	42	45	48
12										27	31	34	37	41	44	47	51	54
13											34	38	42	45	49	53	56	60
14												42	46	50	54	58	63	67
15													51	55	60	64	69	73
16														60	65	70	74	79
17															70	75	81	86
18																81	87	92
19																	93	99
20																		105

3. V příkladu 12.3 jsme viděli, že χ^2 lze užít při testu nezávislosti hodnot v kontingenční tabulce o počtu řádků a sloupců $J \times K$ (contingent = závislý, podmíněný ... tj. test si klade otázku: do jaké míry jsou data pozorovaná přibližně stejná jako data vypočtená = nezávislá? tj: je mezi dvěma uvedenými veličinami nějaká závislost?). Tento test je specielním příkladem předchozího testu dobré shody, kdy za teoretické pravděpodobnosti vezmeme ty, co jsou dány součinem součtových pozorovaných pravděpodobností, nikoli měřením. Za této situace má kritérium

$$\frac{\sum (f_m - f_t)^2}{f_t}$$

rozdělení χ^2 o $(J-1)\cdot(K-1)$ stupních volnosti.

V závěru kapitoly byl představen jeden příklad neparametrických testů. U parametrických testů byl obyčejně nějaký parametr rozdělení neznámý, a ten byl podroben danému testu (např. μ, σ). Nyní u neparametrických testů žádný takový parametr není u daného testu k dispozici – odtud název „neparametrické“.

Parametrické testy většinou lépe a rychleji vyvrátí hypotézu H_0, což je jejich cílem, pokud skutečně H_0 má být zamítnuta. Když ovšem nejsou splněny tři důležité předpoklady POUŽITELNOSTI těchto testů, máme k dispozici pouze testy neparametrické.

Tabulka 12.15 uvádí přehled statistických testů neparametrických i parametrických, z nichž k některým jsme se v tomto textu vůbec nedostali, společně s přehledem situací a vhodností jejich použitelnosti.

12.5 Otázky k opakování

U následujících výroků rozhodněte, zda se jedná o výrok pravdivý či nepravdivý.

Otázka 12.1 *Rozdělení χ^2 o n stupních volnosti je součtem n stejně rozdělených veličin U.*

Otázka 12.2 *Rozptyl veličiny χ^2 o n stupních volnosti je roven hodnotě $2n$.*

Otázka 12.3 *Protože hustota rozdělení χ^2 není symetrickou funkcí vzhledem k žádné přímce, některé kritické hodnoty mohou být záporné.*

Otázka 12.4 *Testy dobré shody lze používat jen u diskrétních náhodných veličin.*

Otázka 12.5 *Při oboustranném χ^2 testu jsou kritické hodnoty χ^2_m, χ^2_v symetrické vzhledem k průměru, tj. $\chi^2_m = E\chi^2 - d$, $\chi^2_v = E\chi^2 + d$ pro jistou hodnotu d.*

Otázka 12.6 *Pro rostoucí n se hustota rozdělení $t^2(n)$ blíží hustotě rozdělení $\chi^2(n)$.*

Otázka 12.7 *Při testech v kontingenční tabulce je počet tříd četnosti vždy sudý.*

Tabulka 12.15: Přehled param. i neparam. testů

data	**účel testu**	**parametrický test**	**neparametrický test**
jeden vzorek	zjistit, zda střední hodnota populace, ze které byl vzorek vybrán, se liší od jisté hodnoty μ_0	U-test či t-test	znaménkový test
dva vzorky jednou	zjistit, zda střední hodnoty populací, ze kterých byly vzorky vybrány, se rovnají	U-test či t-test pro nezávislé skupiny	Mannův–Whitneyův test
jeden vzorek dvakrát	zjistit, zda střední hodnoty populací, ze kterých byly vzorky vybrány, se rovnají	U-test či t-test typu „jeden vzorek dvakrát“	znaménkový test nebo Wilcoxonův test
více než dva vzorky jednou	zjistit, zda populace, ze kterých byly vzorky vybrány, mají tutéž střední hodnotu	jednorozměrná ANOVA (F-test)	Kruskalův–Wallisův test
jeden vzorek více než dvakrát dvakrát	zjistit, zda populace, ze kterých byly vzorky vybrány, mají tutéž střední hodnotu	jednorozměrná ANOVA opakovaného měření	Friedmanův test
množina položek, z nichž každá má dva parametry	zjistit, zda tyto parametry či proměnné jsou korelovány	Pearsonův test korelace	Spearmanův test korelace
jeden vzorek	zjistit, zda populace, ze které byl vzorek vybrán, má jisté teoretické rozdělení		test χ^2 typu 12.2.2 nebo 12.2.3 nebo Kolmogorovův–Smirnovův test

Otázka 12.8 *Test dobré shody je jednostranným (konkrétně pravostranným) testem.*

Otázka 12.9 *Mannův–Whitneyův test místo jednotlivých hodnot měření zpracovává jen pořadí těchto měření v souboru uspořádaném podle velikosti hodnot.*

Otázka 12.10 *Neparametrické testy mají obecně větší sílu (= větší schopnost správně zamítnout H_0, když neplatí).*

Odpovědi na otázky viz 13.6.

13 Odpovědi na otázky a výsledky příkladů ke cvičení

13.1 Výsledky cvičení ke kapitole 1

Odpovědi na otázky

1.1 – A, 1.2-A, 1.3-A, 1.4 – N (ale lze, třetí axiom psti se často použije pro konečně mnoho disjunktních náhodných jevů), 1.5 – N ($P(A)$ musíme odečíst od jedničky), 1.6 – N (nemusí – stačí, když je v $\mathcal{A}$ tolik náhodných jevů, že jsou splněny dané tři axiomy jevového pole – $\Omega \in \mathcal{A}$ a $\mathcal{A}$ je uzavřená na rozdíly dvou libovolných jevů a sjednocení nekonečně mnoha po dvou disjunktních náhodných jevů), 1.7 – N (vzorec platí jen tehdy, když jevy A, B jsou disjunktní).

Výsledky příkladů

Příklady na procvičení jsou zvlášť ve skriptech pro cvičení.

13.2 Výsledky cvičení ke kapitole 2

Odpovědi na otázky

2.1-N, 2.2-A, 2.3-N (může a nemusí, ale striktně tro zakázáno není; stené pravděpodobnosti jsou speciálním případem diskrétní pravděpodobnosti), 2.4-A (pro $a = b$ je $\int_a^a f(x)dx = 0$ pro jakékoli a; odtud také plyne, že v axiomu 3 pro hustotu psti mohou být v intervalu ostré i neostré závorky a výsledek je stále stejný), 2.5-N (například ve 2.6 má hustota $f(x)$ bod nespojitosti v nule), 2.6-A, 2.7-A, 2.8-N (například pro $f(x) = 3$ na intervalu $(0; \frac{1}{3})$ a $f(x) = 0$ pro ostatní x vidíme, že funkční hodnoty jsou větší než 1, a přesto se jedná o hustotu psti – nezápornou funkci, z níž je integrál od minus nekonečna do plus nekonečna roven jedné).

Výsledky příkladů

Příklady na procvičení jsou zvlášť ve skriptech pro cvičení.

13.3 Výsledky cvičení ke kapitole 3

Odpovědi na otázky

3.1-A, 3.2-N (ale může – nule nemůže být rovno v tomto případě jenom $P(B)$, protože to při výpočtu dosazujeme do jmenovatele), 3.3-A (odpověď je sice ANO, jedná se o de Morganovo pravidlo pro čtyři množiny:

$$\overline{A_1 \cup A_2 \cup A_3 \cup A_4} = \overline{A_1} \cap \overline{A_2} \cap \overline{A_3} \cap \overline{A_4},$$

ale počítat tuto pst přímo není vůbec jednoduché, protože náhodný jev $\overline{A_i}$ se vyjadřuje pouze tím, jaká volba nesmí být na i-té pozici; zde právě vypočteme pst sjednocení $A_1 \cup A_2 \cup A_3 \cup A_4$ a odečteme od jedničky), 3.4-A, 3.5-A (například tehdy, když všechny dílčí jevy jsou stochasticky nezávislé; protipříkladem je příklad 3.4, kde při výpočtu psti průniku můžeme použít jen pravou stranu rovnosti, nikoli levou), 3.6-A, 3.7-N (v jednom případě tento vzorec nemá smysl, a tedy neplatí, a sice když $P(A) = 0$), 3.8-A.

Výsledky příkladů
Příklady na procvičení jsou zvlášť ve skriptech pro cvičení.

13.4 Výsledky cvičení ke kapitole 4

Odpovědi na otázky
4.1-A, 4.2-N (nabývat může navíc ještě hodnotu 0), 4.3-A (veličina Y udává, s jakou pstí naměříme v N opakováních experimentu relativní četnost úspěchu; její hodnoty nejsou tedy celočíselné, ale psti, kterých nabývá, jsou stále Bernoulliovy psti), 4.4-N (při neprázdném průniku $H_i \cap H_j$ bychom museli vzorec pozměnit a nějakou pst odečíst, protože pst možných výsledků v průniku jevů bychom počítali dvakrát), 4.5-A, 4.6-N (v situaci zmetkovitosti balení bonboniér například $\sum_1^3 P(A|H_i) = 0{,}11$. Tyto psti jsou vázány jevem A a jejich součet roven jedné být nemusí), 4.7-A (výrazy na obou stranách počítají pst průniku jevů $A \cap H_i$ a průnik je operace komutativní, tj. použitím obecné věty o průniku jevů dostaneme obě možnosti), 4.8-A, 4.9-A (odpověď je sice ANO, ale k praktickému výpočtu vzorec není, protože $P(A)$ v čitateli zlomku na pravé straně rovnosti lze zkrátit se jmenovatelem zlomku a dostaneme rovnost $P(H_i|A) = P(H_i|A)$).

Výsledky příkladů
Příklady na procvičení jsou zvlášť ve skriptech pro cvičení.

13.5 Výsledky cvičení ke kapitole 11

Odpovědi na otázky
11.1 – A, 11.2 – N, 11.3 – N, 11.4 – A, 11.5 – A, 11.6 – N, 11.7 – N, 11.8 – A, 11.9 – N, 11.10 – N (průměr se bere nikoli aritmetický, ale vážený), 11.11 – A (pokud se oba rozptyly liší o více než čtyřnásobek, je vhodnější místo paramatrického testu použít test neparametrický).

Výsledky příkladů
Příklady na procvičení jsou zvlášť ve skriptech pro cvičení.

13.6 Výsledky cvičení ke kapitole 12

Odpovědi na otázky
12.1 – N (χ^2 je definováno jako součet veličin U^2), 12.2 – A, 12.3 – N (ze vztahu $\chi^2(1) = U^2$ je vidět, že veličina s tímto rozdělením může nabývat pouze kladných hodnot (resp. záporných hodnot nabývá s nulovou pravděpodobností)), 12.4 – N (příslušné četnosti lze počítat i u spojitých veličin, více viz příklad **??**), 12.5 – N (ne, protože hustota není symetrickou funkcí vzhledem k přímce $x = E\chi^2(n) = n$), 12.6 – N (pro rostoucí n se $t^2(n)$ blíží rozdělení $\chi^2(1)$ ... stupeň volnosti je pouze jeden), 12.7 – N (např. v příkladu 12.3 stačí, abychom uvažovali tři politické strany místo dvou, a počet tříd by byl $J \times K = 3 \cdot 3 = 9$), 12.8 – A, 12.9 – A, 12.10 – N (právě naopak, parametrické testy

mají větší sílu, protože neparametrické testy užívají většinou spíše jen pořadí měřených hodnot než přímo tyto hodnoty).

Výsledky příkladů

Příklady na procvičení jsou zvlášť ve skriptech pro cvičení.

Seznam literatury:

(Budíková, Králová, Maroš 2009) Průvodce základními statistickými metodami, Grada Publishing 2009.

(Fajmon, Hlavičková, Novák, 2014) Matematika 3. Učební text FEKT VUT, v rámci jehož druhé části je probírána pravděpodobnost a statistika. Dostupný online.

(Fajmon, Hodková, 2021) MA08 Teorie pravděpodobnosti – cvičení. Elektronický text v předmětu Ma0008 na Pedagogické fakultě MU Brno. Slouží zejména pro cvičení.

(Fajmon, Nezval, 2021) MA0008 Teorie pravděpodobnosti – s využitím programu Excel. Elektronický text v předmětu Ma0008 na Pedagogické fakultě MU Brno. Slouží zejména pro vypracování zápočtových příkladů.

(Loftus, J., Loftusová, E., 1988) Essence of Statistics, Second Edition, Alfred Knopf, New York 1988. Na základě této knihy jsem kdysi vybudoval svou první výuku pravděpodobnosti a statistiky, dále jsou odtud vzaty tabulky distribuční funkce U, χ^2 a kritické hodnoty rozdělení t a Mannova-Whitneyho testu.

(Robová, Hála, Calda, 2013) Komplexní čísla, kombinatorika, pravděpodobnost, statistika. Edice Matematika pro SŠ, Prometheus 2013. Výborná učebnice představující čtyři obory uvedené v názvu, až na kombinace s opakováním, které by mohly být vysvětleny i srozumitelněji (viz výuka či konzultace).

(Rumsey 2006) Deborah Rumsey: Probability for Dummies. Wiley Publishing 2006.

(Rumsey 2016) Deborah Rumsey: Statistics for Dummies. 2nd Edition, John Wiley and Sons 2016.

(Schmuller 2016) Joseph Schmuller: Statistical Analysis with Excel for dummies. 4th Edition, John Wiley and Sons 2016.