

(10)

Zákon malých čísel

Ze studie výskytu rakoviny ledvin provážené ve všech 3 141 okresech Spojených států vyplynul zajímavý model. Okresy, ve kterých je četnost případů rakoviny ledvin nejnižší, jsou převážně venkovské řídké osídlené okresy v tradičně republikánských státech v oblasti amerického středozápadu, jihu a západu. Co o tom soudíte?

Během několika posledních veřejn byla vaše mysl velmi aktivní a jednalo se převážně o činnost Systému 2. Cíleně jste hledali v paměti a formulovali hypotézy. Vynakládali jste při tom úsilí, což by dosvědčily rozšířené zornice a zrychlený srdeční tep. Ale ani Systém 1 nezahálal: práce Systému 2 závisí na faktech a návrzích získaných z asociativní paměti. Myslénku, že by republikánští politici poskytovali ochranu proti rakovině ledvin, jste asi odmítli. Ve finále úvah jste se zřejmě zaměřili na skutečnost, že okresy s nízkým výskytem rakoviny jsou převážně venkovské. Známí statistikové Howard Wainer a Harris Zwetling, od kterých tento příklad mám, výsledky komentovali takto: „Je snadné a taky lákavé vyvodit, že nízký výskyt rakoviny v těchto okresech přímo souvisí s čistým životním prostředím a zdravým životním stylem – žádné znečištění vzduchu, žádné znečištění vody, přístup k čerstvým potravinám bez aditiv.“ Dává to do-
konalý smysl.

Nyní se podívejme na okresy, ve kterých byla četnost případů rakoviny nejvyšší. Tyto neduživé okresy jsou převážně venkovské řídké osídlené okresy v tradičně republikánských státech v oblasti amerického středozápadu, jihu a západu. Wainer a Zwetling ironicky komentují: „Je snadné

vyvodit, že vysoká míra rakoviny v těchto okresech bude přímo souviset s chudobou a nezdравými aspekty venkovského životního stylu – ne zde přístup ke kvalitní zdravotní péči, strava vysoká na obsah tuku, příliš mnoho alkoholu a příliš mnoho tabáku. “Něco taky samozřejmě nehrají Venkovský životní styl nemůže vysvětlovat jak velmi vysoký, tak velmi nízký výskyt rakoviny ledvin.

Klíčovým faktorem není, zda okresy byly venkovské nebo převážně republikánské. Jde o to, že venkovské okresy mají malý počet obyvatel. A hlavní ponaučení, které si máme z tohoto příkladu vzít, se netýká epidemiologie, ale složitého vztahu mezi naší myslí a statistikou. Systém 1 je velmi zběhlý v jedné formě uvažování: automaticky a bez úsilí identifikují *příčinná spojení* mezi událostmi, někdy dokonce i tehdy, když je spojení falešné. Když se mluví o okresech s vysokým výskytem rakoviny, okamžitě předpokládáte, že tyto okresy se od ostatních z nějakého důvodu liší a že musí existovat příčina, která tento rozdíl vysvětluje. Jak však uvidíme Systém 1 neumí dobře pracovat s „čistě statistickými“ fakty, která změní pravděpodobnost výsledků, ale přitom *nejsou* jejich přímou příčinou.

Náhodná událost ze své podstaty (je náhodná) nevede sama o sobě k vysvětlení, ale *soubor* náhodných událostí se chová podle vysoce pravdělného modelu. Představte si velkou nádobu plnou kuliček. Polovina kuliček je červených, polovina bílých. A teď si představte velmi trpělivého člověka (nebo robota), který vždy poslepu vytáhne z nádoby čtyři kuličky, zapíše počet červených a počet bílých v daném vzorku, pak hodí kuličky zpátky do nádoby, a takto pokračuje znovu a znovu. Pokud byste výsledky shrnuli, zjistíte, že výsledek „2 červené, 2 bílé“ nastal (téměř přesně) častěji než „4 červené“ nebo „4 bílé“. Tento vzťah je proěřeným matematickým faktem. Výsledek opakovaného výběru náhodného vzorku z nádoby můžete předvídat se stejnou jistotou, jako můžete předvídat, co se stane, když udeříte do vejce kladivem. Nemůžete předvídat každý detail toho, jak se skořápka rozletí na kusy, ale můžete si být jisti základní představou. Je tu však rozdíl: uspokojivý pocit kauzality (souvislosti mezi příčinou a následkem), který zažíváte při představě kladiva tříštícího vejce, zcela chybí, když pomyslíte na výběr náhodného vzorku.

Pro příklad s rakovinou je relevantní ještě další, související statistický fakt. Představte si, že u nádoby s kuličkami máme dvě velmi trpělivé osoby, které se při vytahování vzorků z nádoby střídají. Jack vytahuje vždy 4 kuličky, Jill vytáhne vždy 7. Oba zaznamenávají pouze případy, kdy vytáhnou homogenní vzorek, tedy samé červené nebo samé bílé. Kdyby pokračovali

ve své činnosti dost dlouho, Jack bude zaznamenávat takové extrémní výsledky častěji než Jill – konkrétně osmkrát častěji (očekávané procento je 12,5 % u Jacka a 1,56 % u Jill). Opět – žádné kladivo, žádná kauzalita, ale matematický fakt: vzorek 4 kuliček poskytuje extrémní výsledky častěji než vzorek 7 kuliček.

Nyní si představte populaci Spojených států jako kuličky v obří nádobě. Některé kuličky nesou označení RL jako rakovina ledvin. Vytahujete vzorky kuliček a „osazujete“ jimi jednotlivé okresy. Venkovské vzorky jsou menší než ostatní vzorky. Stejně jako ve hře Jacka a Jill se budou extrémní výsledky (velmi vysoký a velmi nízký výskyt rakoviny) nacházet s větší pravděpodobností v řídké osídlených okresech. To je k našemu přiblíhu vše.

Začali jsme u faktu, který volá po příčině: četnost případů rakoviny ledvin se mezi okresy USA výrazně liší a rozdílly jsou systematické. Vysvětlení, které nabízím, je statistické: extrémní výsledky (jak vysoký výskyt, tak nízký) lze s větší pravděpodobností zjistit v malých vzorcích než ve velkých vzorcích. Toto vysvětlení není kauzální. Malý počet obyvatel v okrese ani nezpůsobuje rakovinu, ani jí nezabraňuje. Pouze umožňuje, aby výskyt rakoviny byl mnohem vyšší (anebo mnohem nižší), než je tomu u větší populace. Ve skutečnosti tady není co vysvětlovat. V okresech s menší populací není výskyt rakoviny opravdově nižší nebo vyšší než normální, pouze se to tak u určitém roce *jeví* v důsledku práce s náhodnými vzorky. Kdybychom analyzovali další rok opakovali, budeme pozorovat stejný obecný model extrémních výsledků u malých vzorků, ale okresy, kde měla minulý rok rakovina vysoký výskyt, nebudou vykazovat nutně stejný výsledek. Pokud je to takto, rozdílly mezi hustě osídlenými a venkovskými okresy vlastně nemají váhu faktu: jsou tím, co vědci nazývají artefakty – tedy pozorováními, která jsou produkována výhradně některými aspekty výzkumné metody, v tomto případě rozdílly ve velikosti vzorků.

Právě popsaný případ vás možná překvapil, ale ve skutečnosti to pro vás nemohl být až takový objev. Dávno přece víte, že výsledky velkých vzorků si zaslouží více důvěry než výsledky malých vzorků, a i lidé zcela neznalí statistiky někdy slyšeli o zákonu velkých čísel. Ale „vědět“ neznamená vždy „rozumět“ a možná zjistíte, že následující výroky platí i pro vás:

- Vlastnost „řídké osídlené“ vás okamžitě neupoutala jako relevantní fakt, když jste četli informace o studii výskytu rakoviny.

- Alepoň lehce vás překvapil poměrně velký rozdíl ve výsledcích mezi vzorkem 4 kulíček a vzorkem 7 kulíček.
- I nyní musíte vyvinout jisté mentální úsilí, abyste vnímali, že následující dvě věty sdělují úplně stejnou věc:
 - Velké vzorky jsou přesnější než malé vzorky.
 - Malé vzorky poskytují extrémní výsledky častěji než velké vzorky.

První věta zní pravdivě, ale dokud vám nedá intuitivní smysl druhá věta, vlastně jste té první zcela nerozuměli.

Závěr: ano, věděl jste, že výsledky velkých vzorků jsou přesnější, ale asi jste si neuvědomili, že jste tomu nerozuměli úplně dobře. Nejsite v tom ale sami. Hned první studie, kterou jsme s Amosem Tverským prováděli, ukázala, že i sofistikovaní vědeckí pracovníci mají špatnou intuici a chabě rozumějí účinkům výběru náhodných vzorků.

ZÁKON MALÝCH ČÍSEL

Když jsem počátkem 70. let začal spolupracovat s Amosem, zaměřili jsme se na hypotézu, že lidé, kteří nemají statistické vzdělání či praxi, mohou být dobrými „intuitivními statistiky“. Amos vpravěl studentům v mém semináři o výzkumu na Michiganské univerzitě, ze kterého vyplynuly poměrně opimistické závěry o intuitivní statistice. Já jsem měl na tuto hypotézu poměrně vyhraněný osobní názor: nedávno předtím jsem totiž zjistil, že nejsem dobrý intuitivní statistik, a nevěřil jsem, že bych byl horší než ostatní.

Pro vědeckého psychologa není variančnost při výběru vzorků ničím zvláštním; znamená to nepřijemnosti a nákladné překážky, které proměňují každý výzkumný projekt v hazardní hru. Uvažujte, že byste chtěli potvrdit hypotézu, že slovní zásoba průměrné šestileté dívky je větší než slovní zásoba průměrného šestiletého chlapce. Tato hypotéza v běžné populaci platí, průměrná slovní zásoba dívek je v tomto věku opravdu větší než u chlapců. Jednotiví chlapci a dívky se však hodně liší a může se vám při výběru stát, že zrovna vyberete vzorek, ve kterém bude rozdíl nepřešvédčivý, anebo kde dokonce chlapci budou vykazovat větší slovní zásobu než dívky. Pokud jste výzkumným pracovníkem, takové závěry vás přijdou pěkně drahé, protože jste vynaložili spoustu prostředků i svůj čas a přitom se vám nepodařilo prokázat, že je vaše hypotéza pravdivá. Pokud byste chtěli snížit toto riziko, pak jediný způsob, jak to udělat, představuje

dostatečně velký vzorek. Výzkumníci, kteří pracují s příliš malým vzorkem, se vydávají na milost a nemilost štěstí při výběru vzorku.

Riziko chyby lze pro jakoukoliv velikost vzorku odhadnout pomocí celkem jednoduché metody. Psychologové však obvykle při rozhodování o velikosti vzorku nepoužívají sofistikované výpočty. Spolehnají na svůj úsudek, který bývá často zkreslený. Krátce před onou diskusí s Amosem na semináři jsem četl článek, ve kterém se demonstroval omyl, jehož se vědci pracovníci dopouštěli (a stále dopouštějí). Autor doložil, že psychologové běžně vybírají vzorek tak malý, že se vystavují 50% riziku nemohou potvrdit svoji vlastní pravdivou hypotézu! Žádný vědec při zdravém rozumu by takové riziko neakceptoval. Podle autora bylo důvodem to, že rozhodnutí psychologů ohledně velikosti vzorku odráží běžně rozšířenou intuitivní chybnou představu o rozsahu variančnosti při výběru náhodných vzorků.

Článek mě doslova šokoval, protože vysvětloval potíže, které jsem zažil při svých výzkumech i já. Podobně jako většina výzkumných psychologů jsem rutinně vybíral vzorky, které byly příliš malé, a často jsem docházel k výsledkům, které nedávaly smysl. Teď jsem věděl, proč: podivné výsledky byly ve skutečnosti artefakty mé výzkumné metody. U mě byl tento omyl obzvlášť trapný, protože jsem vyučoval statistiku, a věděl jsem tedy, jak vypočítat velikost vzorku, která by snížila riziko neúspěchu výzkumu na přijatelnou úroveň. Jenže jsem se o velikosti vzorku nikdy nerozhodoval na základě výpočtu. Podobně jako mí kolegové jsem při plánování experimentů důvěřoval tradici a vlastní intuici, statistickou stránkou věci jsem se nikdy vážně nezabýval. V době, kdy Amos přijel na můj seminář, jsem již došel k závěru, že má intuice je nedostatečná, a v průběhu semináře jsme se rychle shodli, že opimistické na Michiganské univerzitě neměli pravdu.

Začali jsme s Amosem zkoumat, jestli jsem jediným hlupákem, anebo jestli jsem členem velké skupiny hlupáků: rozhodli jsme se otestovat, zda se vědci pracovníci zběhli v matematické dopouštějí stejných omylů. Vytvořili jsme dotazník, ve kterém byly popsány realistické výzkumné situace, včetně replikací úspěšných experimentů. Respondenti měli pro jednotlivé případy zvolit velikost vzorku, ohodnotit riziko selhání, které jejich rozhodnutí ohledně velikosti vzorku obnáší, a poskytnout hypotetickému postgraduálnímu studentovi, plánujícímu daný výzkum, nějakou radu. Amos získal jako respondenty skupinu účastníků konference Společnosti pro matematickou psychologii, do které patřili i dva autoři učebnic statistiky. Výsledek našeho experimentu byl jednoznačný: nejsem jediným

hlupákem. Jeden každý omyl, kterého jsem se dopustil, se mnou sdílela valná většina respondentů. Bylo zřejmé, že ani odborníci na matematiku nevěnovali dostatečnou pozornost velikosti vzorků.

První společný článek, který jsme s Amosen napsali, jsme nazvali „Belief in the Law of Small Numbers“ (Věra v zákon malých čísel). S lehkou ironií jsme vysvětlovali, že „intuice ohledně výběru náhodných vzorků se zdá uspokojovat zákon malých čísel, který říká, že zákon velkých čísel se vztahuje rovněž na malá čísla“. Také jsme do článku zahrnuli striktně formulované doporučení, aby vědci pracovníci „pohlíželi na svoji statistickou intuici s velkou nedůvěrou a kdekoliv je to možné, nahradili uváření dojmů řádnými výpočty“.

ZKRESLENÍ: JISTOTA PŘEVÁŽÍ NAD POCHYBNOSTMI

V telefonickém průzkumu tři set seniorů 60 % z nich podpořilo prezidenta.

Kdybyste měli shrnout několika slovy sdělení této věty, jak by znělo? Asi určitě něco ve smyslu „starší lidé podporují prezidenta“. Tato slova vystihují jádro příběhu. Opomenuté detaily o průzkumu, tedy že byl proveden telefonicky a se vzorkem 300 osob, nejsou samy o sobě zajímavé, poskytnutí doplňkové informace, které přitáhnou málo pozornosti. Své shrnutí byste formulovali úplně stejně, i kdyby byl vzorek odlišný. Jistě, totálně absurdní čísla by upoutala vaši pozornost („V telefonickém průzkumu 6 voličů staršího věku [nebo 60 milionů voličů staršího věku]...“). Ale pokud nejste povoláním statistik, asi byste nereagovali příliš odlišně na vzorek 150 a na vzorek 3 000 osob. Toto v praxi znamená výrok, že „lidé nejsou adekvátně vnímaví k velikosti vzorku“.

Úvodní věta o výsledku průzkumu obsahuje dva druhy informací: příběh a zdroj příběhu. Přirozeně se zaměřujete na příběh a ne na otázku zdroje, tedy spolehlivosti výsledků. Pokud by byla spolehlivost evidentně nízká, příběh by se diskreditoval. Kdyby vám někdo řekl, že „skupina přívrženců prezidenta provedla nekorrektní předpojatý průzkum, aby ukázala, že starší lidé podporují prezidenta...“, samozřejmě výsledky takového průzkumu odmítnete; výsledky se nestanou součástí toho, čemu věříte. Předpojatý průzkum a jeho falešné výsledky by se namísto toho staly novým příběhem o politických lžích. V takto jednoznačných případech se můžete rozhodnout sdělení nevěřit. Ale rozlišujete dostatečně mezi „Čel

jsem v *New York Times*, že...“ a „Někdo u automatu na kávu říkal, že...“? Dokáže váš Systém 1 odlišit míru důvěry v nějaké tvrzení? Z principu WYSIATI vyplývá, že nikoliv.

Jak jsem již popisoval v předchozích kapitolách, Systém 1 nemá sklony k pochybnostem. Podlačuje dvojnáčnosti a spontánně konstruuje příběhy, které jsou tak logické, jak jen mohou být. Pokud není sdělení okamžitě negováno, asociace, které sdělení vyvolá, se budou šířit tak, jako by sdělení bylo pravdivé. Systém 2 dokáže pochybovat, protože umí současně udržet v paměti nekompatibilní možnosti. Jenže pracovat s pochybnostmi je podstatně náročnější než sklouznout k jistotě. Zákon malých čísel je manifestací obecného zkreslení, které upřednostňuje jistotu nad pochybnostmi. S tímto zkreslením se setkáme v řadě dalších kapitol.

Silný sklon k tomu, věřit, že malé vzorky přesně vystihují populaci, ze které byly vybrány, je také součástí obecnějšího tvrzení: máme sklon nadhodnocovat konzistentnost a koherenci čehokoliv, co vidíme. Přehnaná důvěra vědců v to, co můžeme zjistit z několika pozorování, těsně souvisí s haló-efektem, tedy pocitem, že známe a chápeme člověka, o kterém toho ve skutečnosti víme velmi málo. Systém 1 předbláhá fakta a sestavuje bohatý obraz na základě útržků informací. Náš stroj na rychlé závěry tedy funguje tak, jako by věřil v zákon malých čísel. Řečeno obecněji, vytváří reprezentace skutečnosti, které někdy nedávají příliš mnoho smyslu.

PŘÍČINA A NÁHODA

Asociační mechanismus stále hledá příčiny. Problém, který máme se statistickými jevy (reprezentativnosti statistických vzorků), spočívá v tom, že si vyžadují jiný přístup. Statistický pohled se nezabývá tím, jak došlo k tomu, že se daná událost přihodila, zajímá se o to, co se mohlo stát *namísto toho*. Nic konkrétního nepůsobilo, že je událost taková, jaká je – náhoda ji vybrala z mnoha dalších alternativ.

Vzhledem k naší slabosti pro kauzální myšlení se vystavujeme závazným omylům v hodnocení nahodilosti skutečně nahodilých událostí. Vezměte si například pohlaví šesti po sobě narozených dětí v nějaké porodnici. Pořadí holčiček a kluků je nepochybně náhodné; události jsou na sobě nezávislé a počet holek a kluků, kteří se v nemocnici během posledních několika hodin narodili, zjevně nemá na pohlaví dalšího narozeného dítěte žádný vliv. Nyní uvažujte tři možná pořadí:

KKKHHH
HHHHHH
KHKKHK

Jsou tato pořadí stejně pravděpodobná? Intuitivní odpověď – „samozřejmě ne!“ – je chybná. Protože jsou události nezávislé a protože výsledky K (kluk) a H (holka) mají přibližně stejnou pravděpodobnost, pak jakákoliv sekvence šesti výsledků nastane se stejnou pravděpodobností jako kterákoliv jiná. Dokonce i teď, když víte, že je tento závěr pravdivý, se tvzení vzpírá vaší intuici, protože pouze třetí sekvence vypadá jako náhodná. Podle očekávání lidé pořadí KHKHKH přisuzují mnohem větší pravděpodobnost než druhým dvěma. Jsme hledači modelů, věříme v koherentní svět, ve kterém se reprezentativnosti (jako je sekvence šesti dívek) nedějí náhodně, ale jsou výsledkem nějaké mechanické kauzality nebo něčeho záměru. Neočekáváme, že uvidíme reprezentativnost na základě náhodného procesu, a když zjistíme něco, co vypadá jako pravidlo, rychle odmítneme myšlenku, že ten proces je skutečně náhodný. Náhodné procesy produkují mnoho sekvencí, které lidi přesvědčí, že proces vlastně není vůbec náhodný. Cháparete již, proč využívání kauzality mohlo přinašet evoluční výhody. Je to součást všeobecné ostráživosti, kterou jsme zdědili po předcích. Automaticky dáváme pozor na náznaky, že se v prostředí kolem nás něco změnilo. Lev se může objevit na pláni kdykoliv, v náhodnou dobu, ale bude bezpečnější pozorovat okolí a reagovat na známky zvýšeného výskytu lvích smeček, i když tento výskyt ve skutečnosti podléhá výkyvným náhodného procesu.

Široké nepochopení náhodlosti má někdy významné dopady. V našem článku o reprezentativnosti jsme s Amosem citovali statistika Williama Fellerera, který ilustroval snadnost, s jakou lidé vidí modely tam, kde žádné neexistují. Během intenzivního bombardování Londýna během druhé světové války se všeobecně věřilo, že bombardování nemůže být náhodné, protože mapa zásahů odhalila nápadné mezery. Někteří lidé se domnívali, že se v těchto nepoškozených oblastech nacházejí němečtí špióni. Z pečlivé statistické analýzy vyplynulo, že rozdělení zásahů je typický náhodný proces – a typický také v tom, že evokuje silný dojem, že není náhodný. „Pro neutřouvané oko,“ poznamenává Feller, „se náhodlost jeví jako pravidelnost nebo tendence ke shluku.“

Brzy jsem měl příležitost aplikovat Fellerovy poznatky v praxi. V roce 1973 vypukla Jomkipurská válka a můj jediný významný příspěvek

k válčícímu úsilí spočíval v tom, že jsem vysokým důstojníkům Izraelských leteckých sil poradil, aby zastavili jedno vyšetřování. Válka se v počátcích vyvíjela pro Izrael dost špatně vzhledem k neočekávaně dobrému výkonu egyptských balistických střel země-vzduch. Ztráty byly vysoké a zdálo se, že nejsou distribuované rovnoměrně. Dostal jsem informace o dvou izraelských letkách operujících ze stejné základny, z nichž jedna ztratila čtyři letouny, zatímco druhá neztratila žádný. Dotaz na mě byl vznesen v naději, že zjistím, proč si tato nešťastná letka vedla tak špatně. Neexistoval žádný objektivní důvod k tomu, domnívat se, že by jedna letka fungovala efektivněji než druhá, nebyly zjištěny žádné operační rozdíly, ale samozřejmě život pilotů se v mnoha náhodných ohledech odlišoval – například, jak si vzpomínám, v tom, jak často mezi jednotlivými misemi jezdili domů, nebo v tom, jak podávali jakési hlášení. Má rada zněla, že velení by mělo akceptovat, že různé výsledky jsou důsledkem slepé náhody a že by měli zastavit vysýlání jednotlivých pilotů. Argumentoval jsem tím, že nejpravděpodobnější odpověď je smůla či štěstí, že náhodné hledání nezejavné příčiny nemá naději a že v mezicase by piloti letky, která utrpěla ztráty, neměli být zatěžováni ještě dalším bremenem, kdy v rámci výslechů nabydou pocit, že na vině byli oni a jejich mrtví kamarádi.

O pár let později způsobil Amos a jeho studenti Tom Gilovich a Robert Vallone rozruch studii, kde zkoumali dezinterpretace náhodnosti v basketbale. „Skutečnost“, že hráči někdy mají šťastnou ruku, obecně uznávají jak hráči samotní, tak trenéři a fanoušci. Nemůžete tomuto úsudku odolat: hráč trefí tři nebo čtyři koše za sebou a vy dojdete k závěru, že ten hráč má šťastnou sérii, a tedy dočasně zvýšené šance na skórování. Tomuto úsudku se přizpůsobí hráči obou týmů – spoluhráči budou s větší pravděpodobností hráči se šťastnou rukou přehrávat, obránci z druhého týmu zase proti němu posílí obranu. Z analýzy tisíců sekvencí hodů na koš však vyplynul závěr, který vás zklame: nic takového jako šťastná ruka v profesionálním basketbale neexistuje, ani při střelení z pole, ani při testném střelení. Ovšem někteří hráči jsou přesnější než jiní, ale pořadí úspěšných a neúspěšných střel odpovídá všem testům náhodlosti. Šťastná ruka vzniká výhradně v očích diváků, kteří velmi rychle začnou vidět v náhodlosti řád a přičinnost. Šťastná ruka při různých hrách patří nejčastěji a nejrozšířenější kognitivní iluze.

Zajímavou součástí příběhu této studie je reakce veřejnosti. Výsledky studie se chytila média, protože šlo o dost provokativní zjištění, ovšem všeobecným postojem byla nedůvěra. Když Red Auerbach, oslavovaný

trénér Boston Celtics, slyšel o Gilovichovi a jeho studii, reagoval: „Kdo je ten chlap? No tak uďal nějakou studii. To mě fakt vůbec nezajímá.“ Tendence vidět modely v nahodilostech je skutečně neodolatelná – a určitě přitažlivější než nějaký chlap, který dělá studie.

Iluze modelu ovlivňuje naše životy v mnoha jiných situacích než na basketbalovém hřišti. Kolik „dobrých roků“ byste měli čekat, než vyhodnotíte svého investičního poradce jako neobvykle schopného? Kolik úspěšných akvizic je potřeba, aby správní rada došla k závěru, že generální ředitel má na tyto obchody mimořádný cit? Jednoduchá odpověď na tyto otázky zní, že pokud se budete řídit intuicí, budete častěji chybovat než nechybovat, protože budete mylně klasifikovat náhodnou událost jako systematickou. Příliš často odmítáme věřit názoru, že hodně z toho, co v životě vidíme, je nahodilé.

Uvedl jsem tuto kapitolu příkladem o výskytu rakoviny ledvin v rámci Spojených států. Tento příklad se dnes objevuje v knižkách určených pro učitele statistiky, ale já jsem se o něm dozvěděl z populárního článku dvou statistiků, které jsem zmínil již dříve (Howard Wainer a Harris Zwetling). Jejich článek se týkal velké investice, nějakých 1,7 miliardy dolarů, kterou provedla Gatesova nadace v návaznosti na pozoruhodné poznatky o charakteristických znacích neúspěšnějších škol. Hledáním tajemství úspěšného vzdělávání se zabývalo mnoho badatelů. Ve snaze zjistit, co odlišuje úspěšné školy od neúspěšných, identifikovali ty úspěšné a zkoumali jejich vlastnosti. Jeden z poznatků, ke kterým došli, bylo, že neúspěšnější školy jsou v průměru velmi malé. V průzkumu 1 662 škol v Pensylvánii, například, bylo 6 škol z nejlepší padesátky z kategorie malých škol, což je nadprůměrná reprezentace s faktorem 4. Tato data vedla Gatesovu nadaci k tomu, aby učinila velkou investici do tvorby malých škol, což se někdy dělo i tak, že velkou školu rozdělili do několika samostatných jednotek. Minimálně desítky dalších prominentních nadací typu Annenberg Foundation nebo Pew Charitable Trust se k této iniciativě připojila, a stejně tak výbor amerického ministerstva školství pro vzdělávání menších komunit.

Výše popsaný postup vám pravděpodobně dává intuitivně smysl. Je snadné zkonstruovat kauzální příběh, který vysvětluje, proč jsou malé školy schopné poskytovat výborné vzdělání, a tudíž produkovat studenty s vynikajícími výsledky – protože učitelé mohou dětem věnovat více osobní pozornosti a podpory, než by se jim dostalo na větší škole. Kauzální analýza se však zde bude mýjet účinkem, protože fakta jsou špatná. Kdyby statistkové, kteří pracovali pro Gatesovu nadaci, měli zjišťovat charakteristiky

neuhovných škol, zjistili by, že nehorší školy také vykazují tendenci být menší, než je průměr. Pravda je taková, že malé školy v průměru nejsou lepší; vykazují prostě vyšší variantnost. Naopak, uvádí Wainer a Zwetling, velké školy mají tendenci produkovat lepší výsledky, zejména ve vyšších ročnících, kde mají děti k dispozici v rámci osnov více studijních možností.

Díky nedávným pokrokům na poli kognitivní psychologie nyní můžeme jasně vidět to, co jsme s Amosem pouze „zřehně zahledli“: zákon malých čísel je součástí dvou obecnějších příběhů o fungování naší mysli.

- Nadměrná důvěra v malé vzorky je pouze jedním z příkladů obecnější iluze: věnujeme více pozornosti obsahu sdělení než informacím o jeho spolehlivosti: v důsledku toho si vyrváíme pohled na svět kolem nás, který je jednodušší a celistvější, než by z daných dat vyplývalo. Činit rychle závěry je bezpečnější ve světě našich představ, než je tomu v realitě.
- Statistika produkuje mnoho výsledků, které si zdánlivě žádají kauzální vysvětlení, ale které samotné k takovému vysvětlení nevedou. Mnoho skutečností ve světě kolem nás je způsobeno náhodou, včetně náhody při výběru vzorků. Kauzální vysvětlení náhodných událostí bude nevyhnutelně nespřávné.

HOVORY O ZÁKONU MALÝCH ČÍSEL

„Ano, studio mělo tři úspěšné filmy od té doby, co nastoupil nový generální ředitel. Ale je příliš brzy na to prohlašovat, že má šťastnou ruku.“

„Neuvěřím, že je ten nový obchodník génius, dokud to nezkonzultuji s nějakým statistikem a ten neodhadne, jestli jeho úspěšná série byla náhodná událost, nebo ne.“

„Vzorek pozorování je ale příliš malý na to, aby se z něj daly vyvodit nějaké závěry. Neříďme se zákonem malých čísel.“

„Mám v plánu držet výsledky tohoto experimentu v tajnosti, dokud nebudeme mít k dispozici dostatečně velký vzorek. Jinak bychom museli čelit kritice, že docházíme k předčasným závěrům.“