Informační Bulletin

České Statistické Společnosti   číslo 3, říjen 1995, ročník 6.

Christian Huygens a vznik teorie pravděpodobnosti

Karel Mačák

V letošním roce dne 8. července uplynulo 300 let od smrti Christiana Huygense. O významu jeho díla *„De ratiociniis in ludo aleæ“* pro vznik teorie pravděpodobnosti už bylo v tomto bulletinu pojednáno v článku [1]; u příležitosti zmíněného třístého výročí by snad mohlo být zajímavé seznámit se s tímto Huygensovým dílem podrobněji. Protože nezbytná historická fakta a souvislosti byly uvedeny již ve zmíněném článku [1], můžeme zde přikročit přímo k výkladu Huygensova díla; vyjdeme přitom z textu obsaženého v prvním souborném vydání Huygensových spisů, které pod názvem *„Opera varia“* vyšlo v Leydenu v r.1724 [1] a které je dostupné v pražské Národní knihovně.

V těchto Huygensových spisech je práce *„De ratiociniis in ludo aleæ“* obsažena ve druhém dílu na str.725 - 744. První dvě strany obsahují Huygensův dopis F.Schootenovi; o tomto dopisu už byla řeč v [1] a zde ho ponecháme stranou. Vlastní práce začíná stručnou, nijak nenadepsanou úvodní částí, pak následuje hlavní text pojednání členěný do čtrnácti témat (nazývaných *Propositio*), která lze podle obsahu rozdělit do tří skupin, a na závěr je uvedeno pět neřešených problémů, z nichž u tří jsou uvedeny výsledky. Tohoto členění se přidržíme i zde; uvedeme jednak (poměrně) volný překlad čtyř *Propositií*, která lze z dnešního hlediska považovat za definice (*Propositio I. - III.*) nebo větu (*Propositio IX.*), jednak zadání všech příkladů obsažených v textu (ať už jsou uvedeny jako samostatné *Propositio* nebo jen na doplnění jiného textu), a to vše doplníme stručným komentářem.

[1]Dle předmluvy byl editorem tohoto souborného vydání známý nizozemský matematik a astronom G.J. 's Gravesande (1688 - 1742).

Propositio I. - III.

Propositio I. *Očekávám-li částku* a *nebo částku* b*, které obě mohu získat stejně snadno, pak hodnota mého očekávání je*

$$\frac{a+b}{2}.$$

Propositio II. *Očekávám-li částky* a, b *nebo* c*, z nichž každou mohu získat stejně snadno, pak hodnota mého očekávání je*

$$\frac{a+b+c}{3}.$$

Propositio III. *Je-li počet případů, v nichž obdržím částku* a*, roven* p*, a počet případů, v nichž obdržím částku* b*, roven* q*, a jestliže předpokládám, že všechny případy se mohou vyskytnout stejně snadno, pak mé očekávání bude mít hodnotu*

$$\frac{pa+qb}{p+q}.$$

Komentář k této části není nutný; vlastně je zde poprvé v historii matematiky zavedena střední hodnota diskrétní náhodné veličiny. Tento pojem se ale u Huygense nevyskytuje; všechny jeho úlohy se vztahují ke hrám o nějakou částku (sázku) a příslušný pojem se proto nazývá buď *expectatio* [2] nebo *sors* [3]. Všechny úlohy v Huygensově spisu jsou řešeny pouze pomocí těchto tří definic.

Propositio IV. - IX.

V této části spisu je řešena úloha o rozdělení sázky, která byla v počátcích teorie pravděpodobnosti jednou ze „základních“ úloh (viz [1]). V dnešní terminologii by bylo možno tuto úlohu (v nejjednodušším případu pro dva hráče) formulovat takto:

Dva hráči hrají sérii her o částku (sázku) C, přičemž tuto částku získá ten hráč, který jako první vyhraje k her. Pravděpodobnost výhry každé jednotlivé hry je pro oba hráče stejná (oba hráči jsou „stejně dobří“). Série her je předčasně ukončena ve chvíli, kdy jednomu hráči chybí do výhry m her, druhému hráči chybí do výhry n her. Jaké je spravedlivé rozdělení částky C mezi hráče?

[2] *Expectatio* nebo *exspectatio, onis, f.* = očekávání.

[3] *Sors, sortis, f.* = los, věštba.

Jak už bylo řečeno v [1], úlohu jako první správně rozřešili (pro některá konkrétní m a n) Pascal s Fermatem. Základem jejich řešení byla přesná formulace poněkud vágního pojmu „spravedlivé rozdělení sázky“: řečeno dnešní terminologií, sázka má být rozdělena mezi oba hráče ve stejném poměru, v jakém jsou v okamžiku přerušení série her jejich pravděpodobnosti výhry celé částky, kdyby se celá série her dohrávala.

Lze ukázat, že poměr, ve kterém má být sázka rozdělena, nezávisí ani na sázce C, ani na počtu her k. Huygens nejprve řeší následující případy pro dva hráče:

a) m = 1, n = 2 (výsledek 3 : 1);

b) m = 1, n = 3 (výsledek 7 : 1);

c) m = 1, n = 4 (výsledek 15 : 1);

d) m = 2, n = 3 (výsledek 11 : 5);

e) m = 2, n = 4 (výsledek 13 : 3).

Úlohy a), b), d) se vyskytují už v (prvním zachovaném) dopisu Pascala Fermatovi z 29.VII.1654 [4]. Huygensovu metodu řešení těchto úloh lze stručně demonstrovat na úloze a). Bude-li se v sérii her pokračovat, pak v první hře při pokračování jsou možné dva „stejně snadné“ (viz *Propositio I*) výsledky: buď vyhraje jeden hráč celou sázku, nebo vznikne situace, ve které jsou na tom oba hráči stejně (oběma chybí do celkového vítězství po jedné hře), což znamená, že při dělení sázky by každý dostal polovinu. Jestliže se tedy nebude v sérii her pokračovat a sázka bude rozdělena, pak hráči, kterému při přerušení chybí k celkovému vítězství už jen jedna hra, přísluší dle *Propositio I*

$$\frac{C + \frac{C}{2}}{2}$$

a druhému přísluší zbytek.

Úlohy b) – e) lze řešit postupně zcela analogicky vždy s využitím předešlé úlohy.

Pak přechází Huygens k řešení úlohy o rozdělení sázky pro tři „stejně dobré“ hráče; metoda řešení je zcela analogická metodě použité pro dva hráče. Nejprve řeší případ, kdy dvěma hráčům chybí po jedné hře a třetímu chybí dvě hry (výsledkem je dělení v poměru 4 : 4 : 1); pak formuluje (i když dle našeho názoru ne právě nejjasněji) svoji metodu zcela obecně pro řešení úlohy o rozdělení sázky pro libovolný počet hráčů:

PROPOSITIO IX. *Abychom mohli vypočítat podíl každého hráče při libovolně mnoha hráčích, z nichž některému chybí více a jinému méně her, je*

[4]Pascalovo a Fermatovo řešení úlohy d) nalézající se v (druhém zachovaném) dopisu Pascala Fermatovi z 23.VIII.1654 je uvedeno v [1].

třeba zjistit, co náleží hráči, jehož podíl má být zjištěn, když on sám nebo nějaký jiný hráč vyhraje následující hru. Sečtou-li se takto získané části dohromady a dělí-li se tento součet počtem hráčů, obdrží se hledaný podíl dotyčného hráče.

Jedná se vlastně o rekurentní postup, který je ilustrován na příkladu hry tří hráčů, z nichž jednomu chybí jedna hra a dvěma chybí po dvou hrách [5] (výsledkem je dělení v poměru 17 : 5 : 5). Celé *Propositio* je uzavřeno tabulkou, v níž jsou uvedeny poměry pro rozdělení sázky v sedmnácti situacích, které mohou nastat ve hře tří hráčů.

Propositio X. - XIV.

Další část Huygensovy práce obsahuje několik řešených úloh různého zaměření. Nejprve zde jsou dvě úlohy, které se v podstatě vyskytují už v Pascalově korespondenci s Fermatem; řečeno dnešní terminologií, v *Propositio X* se počítají pravděpodobnosti padnutí alespoň jedné šestky při jednom až šesti hodech jednou kostkou, v *Propositio XI* se počítají pravděpodobosti alespoň jednoho padnutí dvou šestek současně při jednom, dvou a čtyřech hodech dvěma kostkami. *Propositio XI* pokračuje úvahou (ne výpočtem) o 8 a 24 hodech dvěma kostkami a končí tvrzením (v dnešní terminologii), že pro 24 hodů je zkoumaná pravděpodobnost pořád ještě menší než $1/2$ a pro 25 hodů je už větší než $1/2$, což je známý problém rytíře de Méré z prvního dopisu Pascala Fermatovi (viz též [1]).

Další dvě úlohy nejsou zajímavé ani z hlediska historického, ani metodického: v *Propositio XII* je počítána (v dnešní terminologii) pravděpodobnost padnutí dvou šestek současně v prvním hodu při házení třemi kostkami (s obecnou úvahou pro více kostek), v *Propositio XIII* je řešena následující úloha:

Hráči A a B spolu hrají o nějakou sázku tak, že jeden z nich jednou hodí dvěma kostkami; padne-li sedm bodů, vyhrává A, padne-li deset bodů, vyhrává B, při každém jiném počtu bodů bude sázka rozdělena mezi oba hráče rovným dílem. Jaký je poměr středních hodnot výher obou hráčů?

(Výsledek je 13 : 11.)

Následující *Propositio XIV* lze charakterizovat jako úvod k úlohám, které jsou neřešené umístěné v dodatku; je zajímavé tím, že k řešení úlohy Huygens sestavuje soustavu dvou rovnic o dvou neznámých. Zadání je následující:

Hráči A a B spolu hrají o nějakou sázku tak, že střídavě hází dvěma kostkami a hráč A hází jako první; A vyhraje, hodí-li jako první šest bodů,

[5] Tento příklad se rovněž vyskytuje v Pascalově korespondenci s Fermatem.

B vyhraje, hodí-li jako první sedm bodů. Jaký je poměr pravděpodobností výher obou hráčů? (Výsledek je 30 :31.)

Tuto úlohu Huygens sdělil dopisem Robervalovi [6] a od něho se úloha dostala k Fermatovi a Pascalovi; ti Huygensovi na oplátku (prostřednictvím Carcavyho [7]) poslali jiné úlohy, které se pak objevily v dodatku k Huygensovu spisu (podrobněji o této korespondenci viz [3]).

Huygensův postup řešení lze stručně zapsat takto:

Označme x očekávanou výhru hráče B před hodem hráče A, y očekávanou výhru hráče B před vlastním hodem. Protože 6 bodů může padnout pěti způsoby a 7 bodů může padnout šesti způsoby, dle *Propositio III* platí

$$x = \frac{5.0 + 31y}{36}, \qquad y = \frac{6C + 30x}{36},$$

a řešením této soustavy dostaneme $x = \frac{31}{61}C$, z čehož plyne hledaný poměr.

Problemata

V dodatku ke své práci uvedl Huygens pět neřešených úloh, z nichž u tří uvedl výsledky. Uvedeme zde znění všech pěti úloh s malými komentáři.

Problema I. *A a B hrají se dvěma kostkami tak, že A vyhraje, když hodí jako první šest bodů, a B vyhraje, když hodí jako první sedm bodů. A začíná hru jedním hodem, pak hází B dvakrát za sebou, pak má A dva hody a tak dále, dokud jeden z nich nevyhraje. Jaký je poměr pravděpdodobností výher obou hráčů?* (Odpověď: 10355 : 12276.)

Úlohu zadal Huygensovi Fermat prostřednictvím Carcavyho.

V r.1687 byla v Haagu vydána anonymní práce obsahující jednak pojednání o duze, jednak pojednání o teorii pravděpodobnosti. V tomto „pravděpodobnostním“ pojednání je uvedeno všech pět úloh z Huygensova dodatku a první z nich je řešena. V současné době se považuje za jisté, že autorem této práce byl známý nizozemský filosof Benedictus Spinoza (1632 - 1677), což svědčí o aktivním Spinozově přístupu k aktuálním matematickým problémům oné doby. Podrobnosti lze najít v článku [4], který je dostupný v knihovně Matematického ústavu AV čR.

[6] Giles Personnier de Roberval (1602 – 1675) byl profesorem matematiky na Collége Royal ([2], II, str.876).

[7] Pierre de Carcavy (? – 1684) byl parlamentním radou v Toulouse (1622 – 1636) a v Paříži (1636 – 1647), pak byl ve službách vévody z Liancourtu a od r.1663 byl konservátorem královské knihovny ([2], II, str.758).

Problema II. *Tři hráči A, B a C mají dvanáct kamenů, z nichž čtyři jsou bílé a osm je černých, a hrají spolu tak, že zvítězí ten z nich, který jako první naslepo vytáhne bílý kámen; jako první táhne A, pak B, poté C, pak zase A a tak dále. Jaký je vzájemný poměr pravděpodobností výher všech tří hráčů?*

U této úlohy Huygens neuvedl výsledek. Jakob Bernoulli ve svém spise *„Ars conjectandi"* upozornil na to, že zadání úlohy není jednoznačné, protože není jasné, zda každý hráč má svých dvanáct kamenů nebo zda všichni tři hráči tahají z jedné hromady a není ani jasné, zda se vytažený kámen vrací nebo nevrací zpět; pro všechny tyto varianty Bernoulli podal řešení. Z Huygensovy korespondence s van Huddem [8] uskutečněné v r.1665 (viz [3]) plyne, že Huygens měl na mysli variantu s vracením; pak nezáleží na tom, kolik hromádek kamenů je a úlohu lze poměrně snadno řešit „huygensovsky" (výsledek je 9 : 6 : 4). Van Hudde řešil úlohu bez vracení při společné hroámadce kamenů (pak je výsledek 77 : 53 : 35). úlohu bez vracení při individuálních hromádkách kamenů řešil Bernoulli a našel výsledek 6476548 : 4231370 : 2768457; výpočet je dost komplikovaný.

Problema III. *A vyhraje nad B, když ze čtyřiceti hracích karet, z nichž vždy deset má stejnou barvu, vytáhne čtyři karty různých barev, jinak vyhrává B. Jaký je poměr pravděpodobností výher obou hráčů?*

(Odpověď: 1000 : 8139.)

Úlohu opět zadal Huygensovi Fermat prostřednictvím Carcavyho; z dnešního hlediska se jedná o jednoduchou kombinatorickou úlohu.

Problema IV. *Hráči A a B mají opět dvanáct kamenů, čtyři bílé a osm černých, a hráč A vyhraje nad B, když naslepo vytáhne sedm kamenů, mezi nimiž se budou nalézat tři bílé; jinak vyhrává B. Jaký je poměr pravděpodobností výher obou hráčů?*

Huygens u této úlohy neuvedl odpověď, ale řešil ji v již zmíněné korespondenci s van Huddem (výsledek je 35 : 64); van Hudde řešil i variantu této úlohy, při které *A* vyhraje, vytáhne-li nejméně tři bílé; pak je výsledek 14 : 19.

Problema V. *A a B mají po dvanácti mincích a hrají spolu třemi kostkami tak, že padne-li jedenáct bodů, pak A dá B jednu minci, padne-li*

[8] Johan van Waweren Hudde (některé prameny uvádějí jméno ve tvaru Hudden) (1628 - 1704) byl 30 let starostou Amsterodamu; zabýval se matematikou (viz [2], II, str.801, 919), výpočty rent (viz [2], III, str.48) a dalšími problémy. Jeho rukopisy nebyly nikdy vydány, ale dle Leibnize obsahovaly mnoho cenných výsledků (dle [5]).

ale čtrnáct bodů, obdrží A od B jednu minci. Hru vyhrává ten hráč, který jako první získá všechny mince. Jaký je poměr pravděpodobností výher obou hráčů? (Odpověď: 244 140 625 : 282 429 536 481.)

Úlohu zadal Huygensovi Pascal prostřednictvím Carcavyho.

Závěr

Huygensův spis *„De ratiociniis in ludo aleæ"* byl prvním tištěným pojednáním o úlohách teorie pravděpodobnosti; jak už bylo řečeno v [1], vyšel v r.1657 jako příloha ke spisu Huygensova učitele F. van Schootena *„Exercitationum mathematicarum libri quinque"*. Výrazně ovlivnil počáteční fázi formování teorie pravděpodobnosti; Jacob Bernoulli ve svém spise *„Ars conjectandi"* (který vyšel r.1713, ale fakticky vznikl mezi lety 1679 - 1685 (viz [2], III, str.339)) věnuje zhruba čtvrtinu svého spisu novému otištění a podrobnému komentování této Huygensovy práce. Po dobu přibližně půl století (až do vydání již zmíněného *„Ars conjectandi"* a prací Montmortových a Moivreových [9]) byl Huygensův spis základní prací v oblasti teorie pravděpodobnosti. Přes všechna uvedená fakta byla tato poměrně ranná Huygensova práce zastíněna jeho pozdějšími díly a dnes stojí poněkud stranou pozornosti; lze proto považovat za vhodné připomenout si ji trochu podrobněji u příležitosti třístého výročí Huygensova úmrtí.

Reference

[1] Coufal, J.: *Alea iacta est aneb půl tisíciletí od vytištění úlohy rytíře de Méré.* Informační bulletin České statistické společnosti 5 (1994), č.1 a 2.

[2] Cantor, M.: *Vorlesungen über Geschichte der Mathematik.* Band II, III. Teubner, Leipzig 1900, 1901.

[3] Huygens, Ch.: *Œuvres complètes.* T.XIV. Haag, 1920. [10]

[4] Dutka, J.: *Spinoza and the theory of probability.* Scripta mathematica 19 (1953) č.1, str.24 – 33.

[5] *The New Encyclopædia Britannica.* 1991.

---

[9] Pierre Remond de Montmort (1678 - 1719): *Essai d'analyse sur les jeux de hazards.* První vydání 1708.

Abraham de Moivre (1667 - 1754): *De mensura sortis, seu, de probabilitate eventuum in ludis a casu fortuido pendentibus.* První vydání ve „Philosophical Transactions" Nr.329, 1711.

[10] Celé toto vydání Huygensových sebraných spisů vycházelo v letech 1888 – 1950; bohužel není dostupné v žádné knihovně v naší republice.

Svatý Augustin (∗ 354 – †430) se o matematice vyjádřil se značnou zdrženlivostí:
Dobrý křesťan se má stříci matematiků a všech těch, kteří dělávají prázdné předpovědi, zvláště však tehdy, když se tyto předpovědi splní. Je totiž nebezpečí, že matematici ve spolku s ďáblem matou rozum a zaplétají lidstvo do spárů pekelných.
(De genesi ad literam. 2, XVII, 37)

První počítač v historii?

Jan Coufal, VŠE

Archeologické nálezy dosvědčují úctyhodné stáří dětských hraček. Stavebnice bychom čekali nejspíše v antickém Řecku. Už proto, že Řekové byli vynikající architekti, navíc i proto, že jejich architektura, při vší úctě k její kráse a dokonalosti, užívala konstrukčních principů ne nepodobných těm, jakými jsme se pyšnili v dětství, když jsme ze svých stavebnic budovali skvělé paláce.

V jedné stavebnici Řekové nesporné prvenství mají. A to v přímo největší stavebnici na světě. Udělali si stavebnici z celé planetární soustavy, z celého tehdy známého vesmíru. Na první pohled jde o podivný špás a krutý omyl. Jistěže si staří Řekové nehráli se skutečnými planetami. Stačilo jim si pohrávat s touto vznešenou stavebnicí alespoň v duchu, příp. nestačila-li představivost, kreslili do písku. I takové myšlené pohrávání se stavebnicí může mnoho ukázat. Pracuje se přitom podle zásady *“co by se stalo, kdyby . . . ”* Právě proto, že byli vynikajícími geometry, jako stavitelé tohoto druhu vynikli pýthagorejci[1]. Zřejmě právě takovou hrou se stavebnicí a zkoušením, čím je možno otáčet, přišli na to, že naše Země, o níž už

[1]Byli to příslušníci školy, která svůj původ odvozovala od Pýthagora z ostrova Samos. Pýthagorás (řecky se jeho jméno psalo $\Pi\upsilon\vartheta\alpha\gamma o\varrho\eta\varsigma$) se zřejmě narodil mezi léty 580–572 před naším letopočtem na ostrově Samos. O jeho životě se uchovaly jenom zlomky zpráv. V roce 497 před Kristem umírá v Metapontu v jižní Itálii. Pýthagorás chodil po světě a hledal, pozoroval a zkoumal. Porovnával posloupnosti čísel popisující měsíční fáze s posloupností mořského přílivu; dozvídal se o mystice egyptských číselných údajů použitých při stavbách pyramid; studoval závislost výšky tónu na délce chvějící se struny (všímal si i toho, jak příslušné číselné poměry prostřednictvím hudby uvádějí tanečníky do extáze). Najednou uzřel společný jmenovatel všech těchto jevů – zjevilo se mu **číslo**. Je přítomné v pohybu planet, v majestátu pyramid i ve víru tance. Je všudypřítomné a neomylné. Na rozdíl od konkrétních jevů tohoto světa je číslo dokonalé a absolutní. Jeho zákonitosti neznají výjimky, které jsou běžné ve věcech poznávaných smysly. Na tom nemůže nic změnit ani Zeus ani Moira. Číslo a jeho zákonitost jsou tím opěrným sloupem, který nezávisí na našem světě. Číslo je více než náš svět, je to kosmos, svět dokonalosti.

věděli, že musí být kulatá, se podle všeho otáčí kolem své osy jednou za den. Tím vysvětlovali zdánlivé otáčení hvězdné oblohy. Pýthagorejci věřili v mystické vlastnosti čísel, desítka měla vyjadřovat dokonalost. Proto předpokládali, že kolem centrálního ohně musí obíhat právě deset těles, aby byl celý vesmír dokonalý.

Zkušenosti s pomyslnou stavebnicí vedly k pokusům napodobovat stavebnicovým skládáním zdánlivě nepravidelné pohyby planet. Musíme předpokládat, že alespoň nějaké jednoduché mechanické modely stály v pozadí komplikovaných planetárních teorií, které ve 4. století před Kristem vytvořili astronomové Eudoxos a Kallippos v Athénách. I věhlasný filosof Platón v téže době dosvědčuje, že o planetárních teoriích *"mluvit bez názoru a obrazů by byla marná práce"*. A že Řekové v mechanickém modelování vesmírných pohybů dosvědčuje nález fragmentů mechanického strojku, získaný v roce 1901.

Přenesme se ve stroji času na úsvit našeho 20. století. V dubnu r. 1900 zahnala bouře řeckého lovce mořských hub *Eliase Stadiatise* při zpáteční plavbě od tuniského pobřeží k téměř neobydlenému ostrovu *Antikýthera*[2]. Když se tam pokusil potopit za houbami, spatřil pod vodou vrak lodě naplněný množstvím starých bronzových a mramorových soch. Vynořil se nad hladinu i s bronzovou rukou v nadživotní velikosti, kterou si vzal z vraku sebou, protože se právem obával, že by jeho zprávě nikdo neuvěřil[3]. Lze předpokládat, že se loď potopila někdy mezi lety 80 až 50 před Kristem při plavbě z Athén do Malé Asie.

Vrak lodi byl vyzvednut v roce 1901. Mezi bronzovýni a mramorovými sochami se při třídění kořisti našly i beztvaré úlomky, které byly významnější než všechny sochy dohromady. Šlo snad o nejpozoruhodnější ukázku práce antických mechaniků. Rekonstrukce zlomků trvala dva roky. Po ošetření a bedlivém restaurování byla odkryta bronzová destička s kruhy, nápisy a ozubenými kolečky. Brzy se ukázalo, že nápisy naznačují jakousi spojitost s astronomií. Teprve když byly součástky očištěny, vyloupla se prazvláštní konstrukce, skutečně důmyslný astronomický přístroj s pohyblivými ukazateli, složitými stupnicemi a popsanými kovovými destičkami. Podle rekonstrukce měl strojek více než dvacet koleček, jakýsi druh diferenciálního rozvodu a korunové kolo. Na jedné straně byla umístěna hřídel. Když se otáčela, uváděla rozličnými rychlostmi všechny stupnice do pohybu. Číselníky chránily bronzové štítky s dlouhými nápisy. Můžeme tváří v tvář tomuto strojku pochybovat o tom, že ve starověku nechyběli prvotřídní mechanici?

[2]Ostrov Antikýthera leží mezi Krétským a Jónským mořem, od severozápadu ostrova Kréty jej dělí úzký Antikýtherský průliv, mezi Antikýtherou a jihem poloostrova Pelopónnessos leží ostrov Kýthera a Kýtherský průliv.

[3]Duben roku 1900 je proto považován za počátek podmořské archeologie.

Na počátku šedesátých let zkoumal fragmenty americký profesor Derek J. de Solla Price pomocí záření gama a rentgenem. To umožnilo zaznamenat skrytý počet zubů na jednotlivých kolech. Při pokusu o rekonstrukci na základě těchto nových informací se ukázalo, že šlo o astronomický stroj ne nepodobný astrolábu, který byl zároveň velmi komplexním mechanickým kalendářem. Zařízení je ostatně tak složité, že nemohlo jít o prototyp. Přístroje tohoto typu kdysi ukazovaly zdánlivý pohyb Slunce po zvířetníku, což umožňovalo po celý rok určovat jeho východy a západy i polohy jasných hvězd. Měl nejméně tři číselníky a byl původně zabudován do dřevěné skříňky vysoké asi 32 cm. Dva menší číselníky udávaly doby východu a západu Měsíce a pohyby pěti tehdy známých planet. Ukazoval zřejmě i další astronomické údaje, plnil tedy analogické funkce jako orloj.

Tento přístroj lze také vysvětlovat jako exemplář zvláštního počítacího stroje, umožňujícího vypočítat polohy Měsíce, Slunce a nejspíše i planet a jasných hvězd. Není asi tak důležité, že na stroji je naznačen rok výroby, který odpovídá asi roku 87 před Kristem. Asi by bylo zajímavější zjistit, kdo sestrojil první model tohoto miniaturního planetária.

Třebaže zatím zůstává úloha některých součástí přístroje tajemstvím, je možno bezpečně říci, že stupnice počítaly s tradičním egyptským kalendářem o 365 dnech. K vyrovnání šestihodinových rozdílů mezi kalendářem a roční cestou Slunce se posouvala stupnice. Lze říci, že šlo o astroláb s převodovým ústrojím. Badatelé sestavili podle torza model přístroje, který jim umožnil nejen objasnit funkci mechanismu, ale i určit, kdy byl přístroj sestaven. V letech 80 až 50 před naším letopočtem, krátce po sestrojení, klesl aparát s lodí ke dnu.

Pro úplnost poznamenejme, že fragmenty strojku jsou dnes uloženy v Řeckém národním archeologickém muzeu v Athénách.

Co lze říci závěrem? Každopádně je možné tvrdit, že pýthagorejci jsou vynálezci modelování v nauce, které uplatnili hned při tom nejvyšším úkolu, který je vůbec myslitelný – při řešení stavby celého vesmíru. Vynalezli metodu, která je neustále nejen živá a dráždivá, ale také velmi účinná, i když přísný filosof Aristotelés bystře postihl nebezpečí, které tu hrozí, a zavázal stavitele geometrických modelů povinností zkoumat, zda to, co může dobře fungovat jako pomyslná stavebnice, je přitom fyzikálně možné. Dlužno říci, že jeho příliš ostrý výpad tehdy zbytečně zmrazil vývoj, který byl slibný.

Literatura

C. W. Ceram : *Oživená minulost (Dějiny archeologie v obrazech).* Orbis Praha, 1974.

Ivan Cigánek: *Divy světa.* Práce Praha, 1978.

Jan Coufal, Jindřich Klůfa: *Matematika I. (pro Vysokou školu ekonomickou).* VŠE Praha, 1994.

Erich von Däniken: *Vzpomínky na budoucnost (nerozluštěné hádanky minulosti).* Orbis Praha, 1971.

Zdeněk Horský, Zdeněk Mikulášek, Zdeněk Pokorný: *Sto astronomických omylů uvedených na pravou míru.* Nakladatelství Svoboda, 1988.

Znovu k české terminologii: pohled z jiné strany

Zbyněk Šidák

K napsání tohoto článku mě podnítily nedávné články J. Žváčka [1], J. Anděla [2] a J. Machka [3] v IB. Konstatuji také, že s některými místy v těchto článcích plně souhlasím, s jinými však zásadně nesouhlasím a bouřil se v nich můj jazykový cit. Nepodléhám ovšem iluzi, že by se po mém článku změnila navyklá terminologie některých lidí a pracovišť, chtěl jsem pouze ozřejmit, jak já mám určitá svá stanoviska zdůvodněná (viz [2]).

1. Na základě své dlouholeté práce v komisích pro obhajoby a při jiné posuzovatelské činnosti mohu jen potvrdit, že některé kandidátské práce a jiné elaboráty a občas i jejich posudky (někdy i odborně dost uznávaných posuzovatelů) jsou tristní co do češtinářských poklesků (chybné užívání čárek, slova jako „standartní“ nebo „vyjímka“, atd., viz [1]).

2. Co se týče termínů *rozdělení* či *rozložení*, já osobně jsem proti termínu *rozdělení*, poněvadž příliš navozuje představu dělení něčeho na kousky, např. řezáním, představu dělení vysloveně diskrétního (takto třeba *rozdělujeme* koláč nebo jablka dětem, ...). Jde-li však o spojitou náhodnou

veličinu, pak termín *rozdělení* je podle mého názoru velmi nevhodný, protože spojitou hmotu či pravděpodobnost nemohu přece *rozdělovat* na kousíčky s nulovou pravděpodobností a pak takové kousíčky někam rozmísťovat. Tedy shrnuto: termín *rozdělení* se mi zdá vhodný pouze pro diskrétní náhodné veličiny, kdežto pro spojité náhodné veličiny je vhodný jedině *rozložení* - takže univerzální termín pro oba případy je pak rovněž *rozložení*.

Není sporu o tom, že řada statistických termínů byla převzata z fyziky (*momenty, difuze, spektrální hustota*, atd.). V českých fyzikálních učebnicích (namátkou uvádím [4], str. 89, 90, 370, 552,…,[5], str. 16,160, 163,…,[6], str. 22, 23, 27-34, 135, 151,…) se zásadně hovoří o *rozložení* hmoty, napětí, sil, elektrického náboje, atd. To by naznačovalo, aby čeští statistici se rovněž drželi tohoto běžného fyzikálního termínu, když přece představa o jeho faktické náplni je vlastně stejná ve statistice jako ve fyzice.

Tento termín rovněž koresponduje se zeměpisem (totéž viz [3]), kde říkáme, že „Čechy se rozkládají (samozřejmě spojitě) mezi Krkonošemi, Šumavou, atd.“ Ironizující poznámka J. Anděla [2], že mu to připomíná „hnilobný rozklad“, mi připadá dost přehnaná. Ve fyzice je to termín běžný a co by vůbec tedy měli říkat chudáci zeměpisci?

Historicky má pravdu J. Machek [3], že ve škole prof. L. Truksy se říkalo *rozložení*, ve škole prof. J. Janko *rozdělení*. (Proč vlastně prof. Janko se v tomto slově odchýlil od jinak běžně používaných termínů původu fyzikálního?) Bohužel vím, že zde vedu donquijotský boj: žáků prof. Truksy valem ubývá, zatímco žáci prof. Janko (přímí i nepřímí) zaplavili českou statistiku (z MFF UK i z VŠE). Prof. J. Hájek jako žák prof. Janko říkal samozřejmě *rozdělení*, to pak přinesl na MFF UK a naučil to zde první generaci svých studentů, ti pak to učili další a další generace, takže tento podle mne nevhodný jazykový úzus se od té doby stále reprodukuje. (Abych byl dobře pochopen: prof. J. Hájka jsem si nesmírně vážil, jeho přínos vědě je skvělý, jeho jiné nově ražené termíny jako *pořadové a pořádkové statistiky, vychýlení a nevychýlené odhady*, atd. považuji za výborné, ale v otázce *rozdělení* či *rozložení* jsme se nikdy neshodli.)

3. Ve stanovisku k termínům *vícerozměrná* či *mnohorozměrná analýza* naprosto souhlasím s J. Andělem [2]. Nevím, odkud J. Žváček [1] vzal tvrzení, že „sémanticky je nám nesporně nejbližší němčina“ a „že u nás historicky převládá termín *vícerozměrná analýza*“, a pokládám je za nepravdivá. Druhé tvrzení možná vzniklo zkreslenou optikou pracovníka VŠE; nahlédnutím do řady knih jsem zjistil, že termín *vícerozměrná analýza* se používá na VŠE, kdežto na MFF UK se říká *mnohorozměrná*. (Podobně pro *vícerozměrná* a *mnohorozměrná rozdělení*.)

Také nevím, proč *mnohorozměrná analýza* je pro J. Žváčka rusofilský termín (se zřetelným pejorativním nádechem). Myslím, že stejně dobře bych jej mohl pokládat za anglofilský, protože jak doložil J. Anděl [2], anglické *multi-* se převážně překládá *mnoho-* (navíc to pochází z latinského *multum* = *mnoho*). Souhlasím s J. Žváčkem, že „duchu doby dnes odpovídá výraz co nejangličtější“. Žádné anglické *morevariate* nebo *moredimensional* však neexistuje, takže *vícerozměrná analýza* je vlastně naopak germanizmus.

I mně, tak jako J. Andělovi, slovo *vícerozměrný* sugeruje otázku „více než co“. *Mnohorozměrný* chápu jako matematický protiklad ke slovu *jednorozměrný*, takže pro mne (i v našich krajích) 2 je už skutečně *mnoho*.

4. J. Machek [3] se zmínil též o překladu anglického *power function*; to je historicky dost podobné jako v bodu 2. Ve škole kolem prof. Truksy, např. u Dr. Fischera, se říkalo *mohutnost*, ve škole prof. Janko *síla* nebo *silofunkce testu*. Opět zvítězila terminologie prof. Janko (proti těmto termínům však nemám žádné námitky).

5. K rozumně nepřeložitelným slovům *bootstrap* a *jackknife* (viz [1]) lze přidat řadu dalších, např *martingal*, *fuzzy* množiny, *distribution-free* tests. Poslední termín značí ovšem *testy nezávislé na distribuci*, ale jak to říci česky výstižně a krátce? Zpravidla se jim říká *neparametrické testy*, ale to není ono, významová náplň je odlišná. Zvláště uvážíme-li, že *neparametrické testy* se velmi často užívají vlastně pro testování parametrů! Nebo: je vhodné anglické *score* překládat podivným českým *skór*?

6. Nevím, proč se J. Machkovi [3] nelíbí slovo *nerovnice* a jaký jiný název by použil? Já proti tomuto slovu nic nemám a *nerovnice* samozřejmě není totéž jako *nerovnost*.

Snad je však už na čase po tolika vážných úvahách přidat také něco v lehčím tónu.

7. Naše zkušenosti v Matematickém ústavu Akademie při konzultacích s češtinářskými odborníky byly bohužel vesměs špatné. V padesátých letech byly studovány (zejména v teorii hromadné obsluhy) náhodné procesy (či vstupní proudy) *bez posledějstvija*, německy *nachwirkungsfreie* (viz [7], [8] aj.). V podstatě šlo o celočíselné neklesající procesy s nezávislými přírůstky. Z Ústavu pro jazyk český nám v dopisu z 25.3.1957 navrhli říkat jim *nespojité* nebo *isolované* nebo *samostatné* nebo *jedinečné* procesy, tedy jeden termín horší než druhý, všechny vzniklé naprostým nepochopením.

Jinou módou byly tenkrát *learning processes* (viz [9] aj.), matematicky modelující psychologické procesy při učení se něčemu. V témže dopise nám jazykový odborník doporučuje termín *učné* procesy.

V češtině vůbec je kamenem úrazu zvratné zájmeno *se*, viz např. *branching processes*, kterým se říkalo buď *větvící* nebo *větvící se* procesy, později též *rozvětvovací*. Nevím, co s těmito typy termínů.

8. Také jsem mnohokrát odpovídal na telefonické dotazy překladatelů (někdy spíše „překladatelů") odborné literatury, např. zdali ruské *geometrija na ploskosti* je česky *geometrie na ploskosti*?!

9. Vždycky mi vrtalo hlavou, proč v češtině používáme dost nesmyslné a nevhodné přídavné jméno *směrodatná* odchylka; anglické *standard* deviation je aspoň trochu vhodnější. Při přípravě tohoto článku jsem objevil publikaci [10], kde se (v r. 1948!) dokonce běžně používají termíny *směrodatná úchylka* a *úchylky od průměru*!

LITERATURA

[1] Žváček, J.: *Kšaft umírající matky - české statistické terminologie.* Informační bulletin České statistické společnosti **5** (prosinec 1994), 20-23.

[2] Anděl, J.: *K problémům české statistické terminologie.* Informační bulletin České statistické společnosti **6** (duben 1995), 4-7.

[3] Machek, J.: *Terminologické úvahy.* Informační bulletin České statistické společnosti, **6** (duben 1995), 8-11.

[4] Nachtikal, F.: *Technická fysika.* JČMF, 1946.

[5] Brdička, M.: *Mechanika kontinua.* NČSAV, 1959.

[6] Votruba, V., Muzikář, Č.: *Theorie elektromagnetického pole.* NČSAV, 1955.

[7] Chinčin, A. Ja.: *Potoki slučajnych sobytij bez posledějstvija.* Těorija věrojatnostěj **1** (1956), 3-18.

[8] Zítek, F.: *K těorii ordinarnych potokov.* Čechoslov. mat. ž. **8(83)** (1958), 448-459.

[9] Bush, R.R., Mosteller, F.: *Stochastic models for learning.* Wiley, 1955.

[10] *Užití korelačního počtu.* (Návrh čs. normy.) Čs. společnost normalizační, Praha 1948. (Autoři neuvedeni, ale předmluvu podepsali J. Kaucký, J. Novák, V. List.).

Výběrový rozptyl

Jiří Anděl

Mějme reálná čísla $x_1, \ldots, x_n$, kde $n \geq 2$. Označme $\bar{x}$ jejich aritmetický průměr. Dále položme

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2, \quad m_2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2, \quad C_2 = \frac{1}{n+1}\sum_{i=1}^{n}(x_i - \bar{x})^2.$$

Pokud lze $x_1, \ldots, x_n$ interpretovat jako výběr z nějakého rodělení s konečným rozptylem σ^2, pak s^2 je nestranným odhadem právě tohoto parametru σ^2. Na mnoha kalkulačkách je k disposici s^2 a m_2.

Název článku prof. Čermáka z Informačního bulletinu České statistické společnosti **6** (1995), č. 2, str. 1–3, však zní:

Patří do jmenovatele výběrového rozptylu n nebo $n - 1$?
Nikoliv, patří tam $n + 1$!

Z této formulace plyne, že prof. Čermák definuje výběrový rozptyl jako C_2 (písmeno C jsem si ostatně vypůjčil právě ze jména Čermák). V jeho článku je to zdůvodněno tím, že právě C_2 má nejmenší čtvercovou odchylku od σ^2, pokud $x_1, \ldots, x_n$ je výběr z $N(\mu, \sigma^2)$ s neznámými parametry μ a $\sigma^2 > 0$. Já se pokusím čtenáře přesvědčit o něčem jiném:

Do jmenovatele výběrového rozptylu patří n.

Napřed několik obecných poznámek. Na první pohled se zdá, že si v matematice můžeme objekty nazývat a označovat jak chceme. Ale jsou zde patrně určité meze. Je zcela běžné pracovat s posloupností kladných čísel ε_n, která konverguje k nule při $n \to \infty$. V knížce Steenrod N. E. a kol. (1973): How to Write Mathematics (vyd. Amer. Math. Soc.) se v příspěvku prof. Halmose na str. 27 dočteme: „Noční můrou matematika je posloupnost n_ε, která konverguje k nule, když ε roste nade všechny meze.“ To se týká označení. A pokud jde o názvosloví, i ve velmi abstraktních prostorech definujeme třeba pojem koule tak, aby měl obdobné vlastnosti jako obyčejná koule v trojrozměrném Euklidově prostoru.

Zkusme nyní domyslet, k čemu by vedlo, kdybychom se při definici výběrových momentů řídili kritériem nejmenší čtvercové odchylky (stručně MSE z anglického *mean square error*).

Nechť $x_1, \dots, x_n$ $(n \geq 2)$ je výběr z $N(\mu, \sigma^2)$ s neznámými parametry μ a $\sigma^2 > 0$. Je známo, že pro centrální momenty rozdělení $N(\mu, \sigma^2)$ platí

$$\mu_{2k-1} = 0, \qquad \mu_{2k} = \frac{\sigma^{2k}}{2^k}\frac{(2k)!}{k!} \qquad \text{pro} \qquad k = 1, 2, 3, \dots .$$

Zaveďme výběrové centrální C_p momenty

$$C_p = \gamma_p \sum_{i=1}^{n} (x_i - \bar{x})^p,$$

kde koeficienty γ_p jsou voleny tak, aby každý moment C_p měl minimální střední čtvercovou odchylku

$$MSE_p = E(C_p - \mu_p)^2.$$

Z článku prof. Čermáka víme, že $\gamma_2 = \frac{1}{n+1}$. Dále je zřejmé, že musí platit $\gamma_p = 0$ pro každé liché $p > 0$, protože pro tato p jsou nulové centrální momenty μ_p. Už to, že $C_{2k-1} = 0$, by mělo vzbudit naši ostražitost. Nebývá zvykem, aby výběrové charakteristiky byly přesně rovny odpovídajícím teoretickým ukazatelům.

V dalším se omezíme jen na případ $n = 2$. Vedou nás k tomu dva důvody. Jednak rozdíly mezi s^2, m_2 a C_2 jsou tím větší, čím je n menší. A za druhé vzorce při $n = 2$ budou zvlášť jednoduché. Budeme přitom předpokládat, že $x_1 \neq x_2$.

Jelikož při $n = 2$ máme $\gamma_2 = 1/3$, po jednoduché úpravě dostaneme

$$C_2 = \frac{1}{6}(x_1 - x_2)^2.$$

Již víme, že $\gamma_3 = 0$, takže $C_3 = 0$. Ještě vypočteme γ_4. Nejdřív se snadno ověří, že

$$C_4 = \gamma_4 \frac{1}{8}(x_1 - x_2)^4.$$

Dál položíme $\xi = x_1 - x_2$. Z našich předpokladů vyplývá, že $\xi \sim N(0, 2\sigma^2)$. Obecné momenty veličiny ξ jsou tudíž totožné s jejími centrálními momenty. Proto

$$E\xi^4 = \frac{(2\sigma^2)^2}{2^2}\frac{4!}{2!}, \qquad E\xi^8 = \frac{(2\sigma^2)^4}{2^4}\frac{8!}{4!}.$$

Odtud dostaneme

$$MSE_4 = E(C_4 - \mu_4)^2 = E\left(\frac{\gamma_4 \xi^4}{8} - \mu_4\right)^2 = \sigma^8\left(\frac{105}{4}\gamma_4^2 - 9\gamma_4 + 9\right).$$

Minima je dosaženo pro $\gamma_4 = 6/35$, takže

$$C_4 = \frac{3}{140}(x_1 - x_2)^4.$$

Pro teoretické momenty μ_p libovolného rozdělení (nejen normálního) však platí celá řada různých nerovností. Jednou z nejdůležitějších je

$$\mu_2\mu_4 - \mu_3^2 \geq \mu_2^3.$$

Důkaz se najde třeba v knížce Anděl J. (1993): Statistické metody, Matfyzpress Praha, na str. 24. [1] Proč je zrovna zmíněná nerovnost tak důležitá, to by byl námět na jiný článek. Zde jen naznačím, že je to jakási obdoba toho, že rozptyl je vždy nezáporný.

Ptejme se nyní: Platí také nerovnost

$$C_2C_4 - C_3^2 \geq C_2^3?$$

Když sem dosadíme naše výsledky vypočtené pro $n = 2$, zjišťujeme, že by muselo platit $C_4 \geq C_2^2$, což je ekvivalentní s nerovností

$$\frac{3}{140} \geq \frac{1}{36}.$$

Tato poslední nerovnost však neplatí, neboť $0{,}021\,428\,5 \neq 0{,}027\,777\,8$. A tak neplatí ani příslušná nerovnost pro C_p momenty. Domnívám se, že to je už samo o sobě dostatečně silný důvod proti tomu, aby se C_p nazývaly výběrové momenty — a specielně proti tomu, aby se C_2 nazýval výběrový rozptyl.

Podle mého (ale nejen mého!) názoru při daných $x_1, \ldots, x_n$ by se měl definovat obecný výběrový moment m'_p a centrální výběrový moment m_p pomocí

$$m'_p = \frac{1}{n}\sum_{i=1}^{n} x_i^p, \qquad m_p = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^p.$$

Pak jsou totiž tyto momenty totožné s teoretickými momenty diskrétního rozdělení, které nabývá hodnot $x_1, \ldots, x_n$ a každou z nich s pravděpodobností $1/n$. Pokud by některá z čísel $x_1, \ldots, x_n$ byla totožná, brala by se jen jednou, ale pravděpodobnost by se pak rovnala počtu těchto totožných čísel dělenému n. Tím by se zaručilo, že pro m'_p a m_p budou platit všechny vztahy a nerovnosti jako pro teoretické momenty. Profesor Hájek nazýval postupy založené na takovýchto charakteristikách naivní statistikou (ale bez jakéhokoli pejorativního přídechu).

---

[1] Dovolte malou reklamu. Publikace je ke koupi jednak na MFF UK na Malostranském náměstí, jednak na VŠE Praha.

Chceme-li proto něčemu říkat výběrový rozptyl, měla by to být veličina m_2. Upřímně řečeno, tak trochu na to narazil i prof. Čermák hned v prvním odstavečku svého příspěvku. Upozorňuje tam na souvislosti s tzv. základním rozptylem při výběrech s konečných populací. Ale vždyť někdy tato konečná populace sama může být výběrem z populace ještě větší. Neměl by potom být základní rozptyl totožný s odpovídajícím výběrovým rozptylem, když i hranice mezi základním a výběrovým souborem může být jen relativní? To by mohl být další argument pro m_2.

Kterému odhadu stranit aneb variace na článek prof. Čermáka

Eva Jarošová

Uvažujme dva hypotetické soubory statistiků; jejich identifikačním znakem bude názor na vydatnost odhadu. Jedni spojují vydatnost pouze s nestranným odhadem, druzí jsou příznivci „vydatnosti za každou cenu“, tj. bez ohledu na případně zkreslení či vychýlení.

Odlišný úhel pohledu se projeví např. u odhadu rozptylu. Společným jmenovatelem prvních je $n-1$, zastánce druhé linie poznáte podle $n+1$ (viz článek prof. Čermáka). Jako polehčující okolnost je u vydatného odhadu rozptylu uváděna aspoň asymptotická nestrannost. Všimněme si ještě poměru vychýlení odhadu vůči jeho směrodatné odchylce, udávající relativní vychýlenost odhadu. Pro vychýlení vydatného odhadu rozptylu σ^2 platí

$$b(s_{n+1}^2; \sigma^2) = E(s_{n+1}^2) - \sigma^2 = -\frac{2}{n+1}\sigma^2$$

pro jeho rozptyl

$$D(s_{n+1}^2) = \frac{(n-1)^2}{n(n+1)^2}\mu_4 - \frac{(n-1)(n-3)}{n(n+1)^2}\sigma^4,$$

kde σ^2 je odhadovaný rozptyl, μ_4 je čtvrtý centrální moment a s_{n+1}^2 vydatný odhad rozptylu σ^2.

„S jistou mírou špičatosti“ (konkrétně $\mu_4 = 3\sigma^4$ u rozdělení $N(\mu, \sigma^2)$) lze psát

$$D(s^2_{n+1}) = \frac{2(n-1)}{(n+1)^2}\sigma^4$$

a

$$\left|\frac{b(s^2_{n+1}; \sigma^2)}{\sqrt{D(s^2_{n+1})}}\right| = \sqrt{\frac{2}{(n-1)}}$$

a je nutno přiznat, že pro větší n to s relativní vychýleností není tak zlé (obr.1).

Co se týče odhadu směrodatné odchylky, žádná ze stran si nedovolí odhad rozptylu jen tak odmocnit. Toto na první pohled líbivé řešení by nebylo ani nestranné v prvním případě, ani vydatné v případě druhém. Podívejme se, jak se s tímto problémem oba tábory vyrovnají. Omezme se přitom na výběry z normálního rozdělení.

Přístup „nestranníků“ je všeobecně známý. Drží se v podstatě svého a nestrannost zajistí přidáním c_n, tedy

$$\hat{\sigma}_{nestr} = c_n s.$$

kde s je odmocnina *jejich* nevychýleného odhadu s^2_{n-1}. A co zastánci vydatnosti? Jejich přístup se v literatuře příliš často neuvádí.

Aplikujme postup zmíněný v článku prof. Čermáka na směrodatnou odchylku, to znamená hledejme, pro kterou hodnotu m statistiky

$$s_m = \sqrt{\frac{1}{m}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

je střední čtvercová odchylka $E(s_m - \sigma)^2$ minimální. Lze dokázat, že pro přirozená $n > 1$ platí

$$\sqrt{m} = \frac{n-1}{\sqrt{2}} \frac{\Gamma\left(\frac{n-1}{2}\right)}{\Gamma\left(\frac{n}{2}\right)} = c_n\sqrt{n-1}.$$

Vidíme, že vydatný odhad směrodatné odchylky dostaneme odmocněním nestranného odhadu rozptylu a použitím stejného c_n, tentokrát ovšem ve jmenovateli, tj.

$$\hat{\sigma}_{vyd} = \frac{1}{c_n}s = s^\star.$$

Pro úplnost uveďme také u odhadu směrodatné odchylky výše zmíněný poměr. Platí

$$\frac{|b(s^\star; \sigma)|}{\sqrt{D(s^\star)}} = \sqrt{c_n^2 - 1}.$$

Nahradíme-li c_n přibližným vztahem

$$c_n \approx \frac{4n-4}{4n-5},$$

dostaneme

$$\frac{|b(s^\star;\sigma)|}{\sqrt{D(s^\star)}} \approx \sqrt{\frac{8n-9}{(4n-5)^2}}.$$

Relativní vychýlenost odhadu směrodatné odchylky konverguje rovněž k nule, přičemž pro každé přirozené $n > 1$ je menší než u odhadu rozptylu (obr.1).

Obr.1. Relativní vychýlenost odhadu rozptylu a směrodatné odchylky normálního rozdělení

Jak velký může být rozdíl mezi průměrem a mediánem

Václav Čermák

V časopise *The American Statistician*, který si pro svou čtivost a žánrovou rozmanitost získal velkou oblibu i mezi našimi statistiky, byl v r. 1990 otištěn článek (viz [3]), v němž se dokazovalo, že rozdíl mezi průměrem a mediánem nemůže být (v absolutní hodnotě) větší než kolik činí hodnota směrodatné odchylky. Formálně :

$$|\mu - m| \leq \sigma. \tag{1}$$

Důkaz v tomto článku uvedený nepatřil právě k nejjednodušším, takže neudivovalo, že do redakce *The American Statistician* přišlo postupně několik krátkých noticek s náměty na jiné, zpravidla jednodušší důkazy. Zájemci je mohou nalézt otištěny v dalších číslech a ročnících uvedeného časopisu, a to vždy v oddílu „Letters to the Editor".

Ze zvědavosti jsme nejprve aplikovali nerovnost (1) na několik nejznámnějších rozdělení – diskrétních i spojitých – a také na četné reálné datové soubory. Ukázalo se, že rozdíl mezi průměrem a mediánem nejen že nebyl větší než *směrodatná* odchylka, ale nebyl větší než *průměrná* absolutní odchylka od *průměru* a dokonce že nebyl větší než průměrná absolutní odchylka od *mediánu.* Vnucovala se tak domněnka, že jde o obecně platný vztah, což jsme se pokusili dokázat. Náš postup přinášejí následující řádky.

Všeobecně jsou známy následující dvě nerovnosti:

(a) $\delta_1 \leq \sigma$, kde $\delta_1 = \int |x - \mu| dF$ je průměrná absolutní odchylka od průměru (od střední hodnoty).

Tato nerovnost vyplývá bezprostředně z Cauchyho–Schwarzovy–Buňakovského nerovnosti pro integrály – viz například knihu R. C. Rao „Lineární metody statistické indukce a jejich aplikace", str. 79. Připomeňme dále, že rovnost $\delta_1 = \sigma$ platí pouze v případech kausálního a dvoubodového symetrického rozdělení, jinak platí ostrá nerovnost.

(b) $\delta_2 \leq \delta_1$, kde $\delta_2 = \int |x - m| dF$ je průměrná absolutní odchylka od *mediánu* (viz například Cramérovu učebnici, oddíl 15.5)

Připommeňme opět, že rovnost platí pouze v případě symetrických rozdělení, jinak platí ostrá nerovnost.

Poněkud méně je známa nerovnost (viz [1], [2] nebo [4])

$$-1 \leq \xi^\star \leq 1, \tag{2}$$

kde $\xi^\star$ je Bonferroniova (modifikovaná Pearsonova) míra šikmosti

$$\xi^\star = \frac{\mu - m}{\delta_2}. \tag{3}$$

Poznámka. Krajních hodnot -1 a $+1$ nabývá $\xi^\star$ pouze v případech extrémně sešikmených rozdělení, například $\xi^\star = 1$ když $x_1 = x_2 = \cdots = x_{n-1} = a$, $x_n = b$, $a < b$.

Shrnutím nerovností (1), (a), (b) a (2) dostáváme celkový výsledek v podobě série nerovností

$$\boxed{|\mu - m| \leq \delta_2 \leq \delta_1 \leq \sigma} \tag{4}$$

Přímý důkaz tvrzení, že rozdíl mezi průměrem a mediánem není větší než průměrná absolutní odchylka od mediánu lze podat také tímto způsobem: nechť x značí centrovanou proměnnou ve smyslu $m = 0$, takže nerovnost $|\mu - m| \leq \delta_2$ lze psát

$$\left| \int x dF \right| \leq \int |x| dF, \tag{5}$$

což platí, jak je známo z teorie pravděpodobnosti, pro každé rozdělení F, pro něž existuje konečná střední hodnota. Některým čtenářům bude patrně známnější přepis nerovnosti (5) pro případ diskrétní proměnné x

$$|\frac{1}{n}\sum_i^n x_i| \leq \frac{1}{n}\sum_i^n |x_i| \tag{6}$$

$$|\sum_i^n x_i| \leq \sum_i^n |x_i|, \qquad \text{Q.E.D.} \tag{7}$$

Literatura

[1] Bonferroni, C. E. (1933): *Elementi di Statistica Generale.* Torino, Litografia E. Gili. (Dotisk 1941.)

[2] Frosini, B. V. (1987): *Lezioni di Statistica, Parte prima.* Milano, Vita e Pensiero.

[3] O'Cinneide, C. A. (1990): *The mean is within one standard deviation of any median.* The American Statistician, **44**, 4, 292 – 293.

[4] Čermák, V. (1990): *Srovnání některých měr nesouměrnosti rozdělení z hlediska jejich chování v rámci intervalu možných hodnot.* Statistická revue 10, SEVT Praha, str. 109 – 157.

[5] Čermák, V. (1993): *Diskrétní a spojitá rozdělení – vzorce, grafy, tabulky.* VŠE Praha.

Tabulky kvantilů rozdělení F

Josef Bukač

Rozšířené tabulky kvantilů rozdělení F byly vypočteny pro pravděpodobnosti 0.95, 0.975, 0.99, 0.995, 0.999. Jejich velikost je dána stupni volnosti v čitateli i jmenovateli od jedné do sta. Pro každou z pravděpodobností je to deset tisíc kvantilů.

Relativní přesnost sedm platných míst je zaručena, neboť výpočet byl proveden v intervalové aritmetice. Bylo pak testováno výpočtem distribuční funkce, zda opravdu výsledné intervaly obsahují kvantily a zjištěno, že chyba se nevloudila.

Dalším krokem bude nalezení aproximací kvantilů. Pak bude možné kvantily počítat přímo, to znamená bez iterací a cyklů pro libovolné stupně volnosti. Numericky vypočtené aproximace znamenají vždy ztrátu přesnosti. Lze očekávat, že dostaneme aproximace s přesností nejméně čtyři desetinná místa.

Nejtěžší je otázka, jakými funkcemi aproximovat. Není to předmět studia teorie aproximací, jedná se spíš o specifické vlastnosti aproximované funkce. Bohužel se o tom v literatuře najde velmi málo a nějaká pravidla jak postupovat asi vůbec ne.

Jako ukázku dávám aproximaci pro $P = 0.99$ a stupně volnosti pod a na diagonále. Koeficienty byly vypočteny z běžných tabulek. Ty obsahují velké mezery, takže vyšší přesnost se nedá zaručit.

```
function f990(ndf1,ndf2 : integer) : real;
var r,s,y,w : real;
begin
 { Místo kvantilů nad diagonálou jsou dosazeny nuly. }
 f990:=0.0;
 r:=ndf1; s:=ndf2;
 if ndf1>2 then
 begin  { Max. rel. chyba je 0.0006. }
  y:=1.0/sqrt(r);
  y:=((((0.15576*y-0.44886)*y-0.30314)*y+2.94126)*y+3.28995)*y*r+r;
  w:=((1281.84525/s+51.21173)/s+170.44791)*r/s/y;
  w:=w+((5.294417*r-25.947186)*r+5.208631*y*y-10.503048*y*r);
  y:=(((w-29.24918*y+65.6473018)/s+(2.0+y-r)*0.5)/s+1.0)*y/r;
 end;
 if ndf1=1 then
```

```
 begin
  y:=(((27.4612/s-9.45008)/s+17.92627)/s+11.35729)/s;
  y:=((y+8.86832)/s+4.91655)/s+2.57583;
  y:=y*y;
 end;
 if ndf1=2 then y:=0.5*s*(exp(-2.0*ln(1.0-0.99)/s)-1.0);
 f990:=y;
end; {f990}
```


Společenská rubrika

Dne 1. 6. 1996 se konala svatba dvou členů naší Společnosti, a to Josefa Arlta a Markéty Škuthanové, která změnila své příjmení na Arltová.

Dne 3. 6. 1996 se vdala další členka, a to Lenka Zychová, která přijala jméno Wijnhorstová.

Všem přejeme do dalšího života hodně štěstí a pohody.

Opravy a doplňky k adresáři členů ČStS z ledna 1995

Čas běží a my jsme tady s novými údaji a opravami, neboť i tiskařský šotek si zařádil.

K dnešnímu dni se naše členská základna rozrostla o osm členů:

Budíková Marie, RNDr.,
PřF MU, KAM (katedra aplikované matematiky),
Janáčkovo nám. 2a, 662 95 Brno, tel. 41321251
E-mail: budikova@@math.muni.cz

Bukač Josef,
LF UK, Šimkova 870, 500 38 Hradec Králové
E-mail: bukac@@lfhk.cuni.cz

Keprta Stanislav,
MFF UK, KPMS, Sokolovská 83, 186 00 Praha 8, tel.: 2191 3272
bydliště: č. 197, Kunvald, tel. 0446/8732

Křikavová Martina,
VŠK 17. listopadu,

Pátkova 5, B1509, 182 00 Praha 8, tel. 8556 152, l. 161
E-mail: xkrim04@@st.vse.cz, mkri127@@karlin.mff.cuni.cz

Mačák Karel, Doc. Ing. RNDr. CSc.,
Technická univerzita v Liberci, KMS
(katedra diskrétní matematiky a statistiky),
Hálkova 6, 461 17 Liberec, tel. 329
E-mail: karel.macak@@vslib.cz

Pavelka František, Doc. RNDr. CSc.,
VUT, fakulta technologií, KPE (katedra podnikové ekonomiky)
nám. TGM 275, 762 72 Zlín, tel. 842 274
bydliště: Příční 858, 763 61 Napajedla, tel. 942177

Radek Jan, RNDr. CSc.,
MZe ČR, informatika, Těšnov 17, 117 05 Praha 1, tel. 21812646
Bydliště: 294 41 Vinařice 109

Rytíř Vladimír, Ing.,
VUT, fakulta technologií, IME,
nám. TGM 275, 762 72 Zlín, tel. 842 277
bydliště: U splavu 3841, 760 01 Zlín

Dále si, prosím, opravte adresáře (adresář členů z ledna a E-mailový adresář z dubna 95), dle následujících údajů:

Ambrožová M.	– VZP ČR, Karlovo nám. 8, 128 00 Praha 2
Běláček J.	– platí pouze adresa domů
Ettlerová E.	– E-mail: ettle@@lfhk.cuni.cz
Hartmann M.	– E-mail: hartmann@@lfhk.cuni.cz
Holík M.	– PřF MU, katedra teoretické a fyzikální chemie, Kotlářská 2, 611 37 Brno, tel. 41129348
Kevická R.	– tel. 272335 – práce, tel. 496076 – domů
Klaschka J.	– tel. 66003174
Koróny S.	– platí pouze adresa domů
Krajčík Ján	– E-mail: krajcik@@vumiba.sk
Kunderová P.	– tel. 5414163
Matoušková M.	– ČSÚ, odbor 3210, Sokolovská 142, 186 04 Praha 8, tel. 66042955
Nový M.	– Konces, s.r.o., Šumavská 33, 612 54 Brno, tel. 41235341, 41321101
Procházka M.	– MŠMT ČR, Karmelitská 7, 118 12 Praha 1, tel.5193377
Rodová V.	– platí pouze adresa domů
Skalská H.	– Vysoká škola pedagogická, FŘIT,

	nám. Svobody 301, 501 91 Hradec Králové
Skibová J.	– tel. 61082523 E-mail: jesk@@medicon.cz
Sůva M.	– platí pouze adresa domů E-mail: suvova@@pef.vsz.cz
Semerák K.	– Česká pojišťovna, a.s., Na Perštýně 12, 113 04 Praha 1, tel. 24051210
Svoboda K.	E-mail: karel@@nb.vse.cz
Šikulová M.	– platí pouze adresa domů
Škuthanová M.	– mění se příjmení na Arltová E-mail: arltova@@nb.vse.cz
Zychová L.	– mění se příjmení na Wijnhorstová
Žváček. J.	– E-mail: zvacek@@zvacek.vse.cz

Změnilo se několik telefonních čísel:

- ústředna SZÚ má nové telefonní číslo 6708 1111; tato změna se týká Marka Malého, Bohumíra Procházky, Zdeňka Rotha, Evy Švandové a Ladislava Tomáška.

  Telefonní číslo na nynějšího předsedu naší společnosti Ing. Rotha je 67082380.
- telefonní ústředna MFF UK má nové telefonní číslo 21911111, takže si doplňte např.

Jiří Anděl	– 21913283,
Jaromír Antoch	– 21913275,
Josef Machek	– 21913274,
Karel Zvára	– 21913276.

Další opravy se týkají části „Seznam organizací". Vyškrtněte si, prosím organizace AV ČR – SEÚ,
ÚHP ČR,
VUT – FS.

Naopak připište si k MU – PřF jména Budíková Marie a Holík Miroslav. Masarykova univerzita Brno tak zvyšuje svůj počet z pěti na sedm a dostává se v tabulce „Zastoupení podle organizací" za MFF UK.

V naší republice dochází stále k velkým změnám. Patří k nim i vznik nových škol sloučením různých fakult, přejmenování škol apod. Protože jsme byli s adresářem v časové tísni, nestihli jsme do něj zapracovat změny názvů škol. Jistě jste o těchto změnách informováni, přesto některé nyní uvedeme:

VŠB Ostrava	– VŠB – TU (Technická univerzita) Ostrava
VŠCHT Pardubice	– Univerzita Pardubice
VŠST Liberec	– Technická univerzita v Liberci
VŠZ Č. Budějovice	– Jihočeská univerzita
VŠZ Praha	– Česká zemědělská univerzita

Oznámení

Česká statistická společnost ve spolupráci s Přírodovědeckou fakultou Ostravské university pořádá 7. února 1996 ve 12.30 na Přírodovědecké fakultě Ostravské univerzity

Ostravský den České statistické společnosti,

kterým chceme navázat na loňský úspěšný statistický den pořádaný v Olomouci a pokusit se založit tradici občasných setkávání statistiků na různých místech a přispět tak k šíření statistické osvěty v prostoru a čase.

Rádi bychom, aby přednášky byly změřeny především na oblast statistického software a aplikací statistiky, ale i další oblasti mohou být dotknuty. Předběžně přislíbili přednášku prof. Militký, prof. Kubáček, doc. Řezanková a doc. Ramík. Vyzýváme další kolegy k podobnému příslibu, i krátké sdělení je možné. Texty přednášek mohou být publikovány v recenzovaném sborníku Acta Oeconomica Pragensia, VŠE Praha.

Vzhledem k tomu, že předpokládaný program statistického dne skončí v pozdních odpoledních hodinách (nebo dokonce neformální diskusí v nočních hodinách v pohostinství), můžeme pro přespolní zájemce zamluvit nocleh).

Další informace budou uvedeny ve druhém oznámení.

Žádáme zájemce o Ostravský den České statistické společnosti, aby se přihlásili k účasti na adrese:

Josef Tvrdík, Katedra informatiky, PřF OU, Bráfova 7, 701 03 Ostrava,

tel. 62 22 808, l.231, e-mail: tvrdik@@osu.cz,

Česká statistická společnost společně s vedením Českého statistického úřadu Vás zvou na seminář

Česká státní statistická služba v transformaci,

který se uskuteční ve čtvrtek 30. listopadu 1995 od 13.30 hod. v místnosti 115 B v ČSÚ, Sokolovská 142, Praha – Karlín.

Úvodní referát přednese předseda ČSÚ p. ing. Edvard Outrata. Další sdělení přednesou vedoucí pracovníci ČSÚ. Poté bude následovat diskuse. Předpokládané ukončení semináře je asi v 16.30 hod.

Informační Bulletin České statistické společnosti vychází čtyřikrát do roka v českém vydání a jednou v roce v anglické verzi. Předseda společnosti: Ing. Zdeněk Roth, CSc, SZÚ Praha, MSP, Šrobárova 48, 100 42 Praha 10, E-mail: roth@szu.cz. ISSN 1210-8022

Redakce: Dr. Gejza Dohnal, Jeronýmova 7, 130 00 Praha 3, E-mail: dohnal@fsik.cvut.cz.