INFORMAČNÍ BULLETIN

České statistické společnosti

Ročník 30, číslo 3, září 2019

PRŮKOPNÍCI STATISTIKY VE VĚDÁCH O ČLOVĚKU V 19. STOLETÍ

THE PIONEERS OF STATISTICS IN THE SCIENCE OF MAN IN THE 19TH CENTURY

Jan Kalina[1], Lubomír Soukup[2]

Adresa: [1]Ústav informatiky AV ČR, Pod Vodárenskou věží 2, 182 07 Praha 8, [1]Ústav teorie informace a automatizace AV ČR, Pod Vodárenskou věží 4, 182 08 Praha 8

E-mail: `kalina@cs.cas.cz`

Abstrakt: V rámci věd zabývajících se člověkem byly statistické metody poprvé systematicky využívány v průběhu 19. století. Tento článek je věnován třem všestranně nadaným osobnostem, které považujeme za průkopníky statistiky ve vědách o člověku. Konkrétně jsou jimi Adolphe Quetelet, Florence Nightingaleová a Francis Galton. Článek se zamýšlí nejen nad jejich přínostem pro statistiku v širším kontextu věd o člověku, a také o hlubších cílech a filosofickém pozadí jejich statistického uvažování.

Klíčová slova: statistické uvažování, normální rozdělení, kvantily, statistický důkaz, filosofie statistiky.

Abstract: Statistical methods started to appear systematically within the science of man in the course of the 19th century. This paper is devoted to three universally gifted personalities, which can be considered as pioneers of statistics in the science of man. Particularly, these are Adolphe Quetelet, Florence Nightingale, and Francis Galton. The paper reflects not only their contribution to statistics within a broader context of the science of man, but also their deeper aims together with a philosophical background of their statistical reasoning.

Keywords: statistical thinking, normal distribution, quantiles, statistical evidence, philosophy of statistics.

1. Úvod

Statistické uvažování založené na počtu pravděpodobnosti se zpočátku vyvíjelo ve třech různých oblastech: hazardní hry, pojišťovnictví, fyzikální měření. Tyto tři oblasti se již velmi brzy začaly vzájemně ovlivňovat. Na přelomu 17. a 18. století zkoumal francouzský matematik Abraham de Moivre (1667–1754) matematické základy hazardních her a své poznatky sepsal formou

učebnice teorie pravděpodobnosti [18]. Zabýval se též aplikací teorie pravděpodobnosti při analýze tzv. úmrtnostních tabulek, které sloužily jako cenný zdroj dat pro stanovování výše splátek životního pojištění; úmrtnostními tabulkami se tehdy zabývalo více myslitelů, např. i Edmund Halley (1656–1742). Již tehdy vnímal Moivre statistické odhadování jako inverzní pravděpodobnostní problém spočívající v odhadování pravděpodobnosti určitého jevu na základě reálně pozorovaných dat. Roku 1733 odvodil vzorec pro hustotu normálního rozdělení pravděpodobnosti jako limitní případ binomického rozdělení.

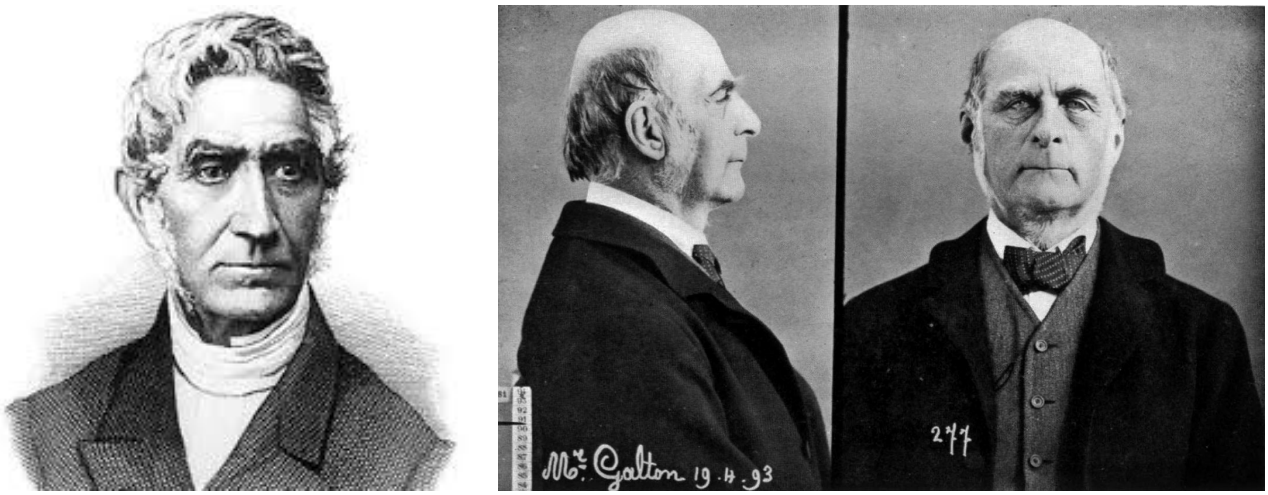
Carl Friedrich Gauss (1777–1855) byl jedním z prvních, kdo začal důsledně používat teorii pravděpodobnosti při zpracování měření z různých oblastí, zejména z astronomie, geodézie a fyziky. Na začátku 19. století rozvíjel teorii chyb měření a položil tak základy metody maximální věrohodnosti, jež se stala pilířem rodící se matematické statistiky. Pomocí metody maximální věrohodnosti odvodil originálním způsobem normální rozdělení pravděpodobnosti a teoreticky zdůvodnil metodu nejmenších čtverců. Astronomie byla cenným inspiračním zdrojem pro zpracování dat zatížených chybami a chybami měření se zabývalo mnoho astronomů té doby, např. Pierre Simon de Laplace (1749–1827), který výrazně přispěl k teorii pravděpodobnosti.

Vědy o člověku jsou pak čtvrtou oblastí, do níž statistické uvažování systematicky proniklo, a to v širokém záběru od biologické antropologie až po společenské vědy. Některé z těchto věd se přitom věnují člověku jako jednotlivci (biologická antropologie, daktyloskopie), jiné naopak pouze celé populaci lidí (sociologie) a na jednotlivce nehledí. Až od 19. století se systematicky využívá statistické uvažování jako vědecká metoda v rámci věd o člověku. K rozvoji statistiky v těchto oblastech nejvíce přispěli Adolphe Quetelet a Francis Galton. Zatímco Galton je dnes hodnocen kontroverzně, považujeme za nutné v tomto článku upozornit i na přínos Florence Nightingaleové, a to nejen k rozvoji statistiky, ale i na morální rozměr jejích aplikací statistiky pro celou společnost.

Usilujeme zde o celkový pohled na vývoj statistického uvažování ve vědách o člověku v 19. století a zamýšlíme se nad motivací tvůrců statistické analýzy a jejich filosofickými východisky. Faktografické životopisy těchto osobností lze najít ve specializovaných monografiích jako např. [7, 2, 28], přičemž ale hodnocení jejich přínosu pro budování matematické statistiky v kontextu vývoje jiných vědních oborů bývá opomíjeno. Proto se v tomto článku věnujeme především zasazení jejich díla do širšího rámce vědecké a společenské atmosféry překotného období na přelomu 19. a 20. století.

2. Adolphe Quetelet (1796–1874)

Adolphe Quetelet (též Quételet; obrázek 1) byl belgický všestranný vědec, který působil jako profesor astronomie a matematiky na Královské vojenské akademii v Bruselu. Dobře ovládal fyziku, která v 19. století přinášela fascinující objevy, i matematiku. Snažil se využívat počet pravděpodobnosti při analýze hromadných dat ve fyzice a ovšem také v oborech nefyzikálního charakteru, zejména v sociologii, demografii, kriminalistice apod. Založil tak nový obor, tehdy nazývaný sociální fyzika nebo sociální mechanika, a stal se jeho hlavním představitelem a tudíž též předchůdcem dnešní sociologie. Uvádí se, že přenášel fyzikální metody do věd o člověku[1], ovšem oněmi fyzikálními metodami se myslí pravděpodobnost se statistikou, které byly výrazně spojovány právě s analýzou fyzikálních měření.



Obrázek 1: Vlevo: Adolphe Quetelet. Vpravo: snímek Francise Galtona pořízený v nejstarší kriminalisticko-technické laboratoři na světě, a to v Bertillonově laboratoři v Paříži v roce 1893.

Quetelet s oblibou vyhodnocoval celostátní ukazatele různých společenských jevů, nejraději pokud šlo o údaje o kriminalitě nebo počtu sebevražd, nebo také sestavoval tabulky počtů duševních nemocí pro různé státy ([16], str. 3). Jako prezident Belgického statistického úřadu a zároveň talentovaný organizátor aktivně přispěl k rozvoji oficiální státní statistiky. Zavedl zejména pravidelná sčítání lidu v desetiletých intervalech, počínaje rokem 1846, a to za pomoci moderních dotazníků ([19], str. 533). Zároveň byl hlavním ini-

[1] Obdobně vznikla i psychofyzika a za zmínku stojí její zakladatel Gustav Theodor Fechner (1801–1887), otec experimentální psychologie ([34], str. 54), který mj. na místo normálního rozdělení zavedl šikmé rozdělení s hustotou vzniklou spojením částí dvou Gaussových křivek s odlišnými rozptyly.

cátorem a organizátorem mezinárodní spolupráce statistiků. Také v dalších státech se organizovala (úřední) statistická služba podle Queteletova bruselského vzoru a Queteletovy myšlenky našly ohlas i v rakouské monarchii a již v 50. – 60. letech 19. století i v Praze ([38], str. 128).

Zároveň se Quetelet s oblibou věnoval svým vlastním měřením, zejména antropometrickým měřením různých vnějších rozměrů belgických mužů. Měl dobrou intuici na to, jaké veličiny měřit a jak data sbírat. Při analýzách pak používal především metody popisné statistiky, zároveň usiloval o hledání příčinných vztahů. Při svých měřeních zjistil, že výška je v mužské populaci přibližně normálně rozdělená. Přibližnou normalitu zjistil i při analýze měření obvodu hrudníku 5738 skotských vojáků či při měření jiných veličin. Quetelet se tak stal historicky prvním, kdo použil normální rozdělení na výsledky měření člověka, ale vůbec i na studium jiných jevů než chyb měření. Normální rozdělení, kterému se tehdy ještě říkalo zákon chyb či (Gaussova) křivka chyb, chápal z astronomie jako univerzální princip, který podle něj platí i při studiu skupin lidí ([16], str. 2). Přitom uvažoval homogenní populaci a jedince pouze porovnával s průměrem, aniž by modeloval individuální variabilitu.

Queteletova antropometrická měření vypovídala o rozměrech lidí z náhodného výběru z určité skupiny lidí. Průměrné hodnoty naměřených veličin tak poskytovaly určitou představu o vzhledu typického představitele dané populace. To vedlo Queteleta, aby ve svém hlavním díle [29] zavedl pojem průměrný člověk (*l'homme moyen*). Ten má ve všech ohledech ideální vlastnosti, ale zůstává jen fiktivní bytostí. V tom se liší od Gausse, který považoval hodnotu měřené fyzikální veličiny za konstantu, která objektivně existuje a bylo by možno ji zjistit měřením jen v ideálním případě, kdy by nepůsobily měřické chyby. Variabilita výsledků fyzikálních měření je navíc mnohem menší než u antropometrických měření. Quetelet si proto uvědomoval, že z toho důvodu potřebuje provádět mnohem větší počet měření.

Normální rozdělení (třeba výšky mužů) obhajuje Quetelet hlouběji než jen svou empirickou zkušeností. V první řadě jej zdůvodňuje filosoficky a teologicky jako zákon daný samotným tvůrcem vesmíru a klade jej svým významem na úroveň gravitačního zákona. Odsud pak usuzuje, že také při analýze reálných dat dává smysl uvažovat normalitu a počítat z nich průměr, který odpovídá jejich těžišti. Jednotliví lidé se pak vztahují k průměrnému člověku tak, jako vesmírná tělesa obíhají kolem svého těžiště [15, 3].

Při měření obezity navrhl Quetelet roku 1835 index tělesné hmotnosti dnes známý jako BMI (*body mass index*), který byl nazýván i jako Queteletův index (QI). Ten považoval za vhodnou míru obezity. Quetelet chápal, že každá veličina má určitou disperzi (*dispersion*) kolem průměru; hodnoty mezi přijatelnými limity variability pak považoval za normální, ostatní za patolo-

gické (abnormální). Snažil se tyto limity (či normy) zformulovat (třeba právě pro BMI), i když jen subjektivně. Dodnes má horní limit BMI svou medicínskou interpretaci a hodnota BMI překračující tuto mezní hodnotu slouží jako rizikový faktor pro určitá onemocnění.

V humanitních oborech dnes není Queteletův odkaz plně doceněn ([16], str. 1). Jeho statistické analýzy zločinu nevedly ke konkrétním společenským opatřením a překonána je už i jeho myšlenka, že každý má svou určitou (změřitelnou) dispozici ke zločinu i k sebevraždě. Zato je dnes Quetelet uznáván jako Otec statistiky, a to i pro svůj vliv na Nightingaleovou a Galtona ([16], str. 1). Přinesl nový pohled na význam normálního rozdělení a také si uvědomoval, že dokázat něco opakovanými experimenty může poskytnout statistický důkaz (anglicky *evidence*), ale ne rigorózní důkaz (anglicky *proof*).

3. Florence Nightingaleová (1820–1910)

Florence Nightingaleová (obrázek 2)[2], ač pocházela ze zámožné rodiny, se rozhodla být zdravotní sestrou a pečovatelkou. Žila v Anglii, působila zejména v Londýně. Vynikala v matematice a analytickém uvažování. Už když jí bylo devět, na zahradě svých rodičů sbírala data o sklizni ovoce a zeleniny do četnostních tabulek a vyhodnocovala je. Když ve dvaceti letech toužila po studiu matematiky, musela si to na svých rodičích tvrdě vybojovat, a uvádí se (ovšem jen bez důkazu [27]), že ji pak soukromě učil James Joseph Sylvester (1814–1897), jehož Sylvestrův zákon pro kvadratické formy známe z lineární algebry. V oblasti metodologie považovala za svého mentora Queteleta, který pro ni byl zakladatelem nejdůležitější vědy na světě ([6], str. 73), tj. statistiky.

Pracovala v nemocnicích, kde sbírala různá data (třeba statistiky lékařské péče nebo úmrtnosti) a snažila se najít a vysvětlit vztahy mezi naměřenými veličinami. Přitom potřebovala také hledat inovativní postupy pro správný sběr dat. Také vyhodnocovala veřejně dostupná data z oblasti veřejného zdraví. Všechny tyto analýzy ale prováděla s velmi praktickou motivací. Snažila se pomáhat v první řadě pacientům a s ohledem na ně usilovala o změny v celém britském zdravotnictví, prosazovala i změny státní politiky a sociální reformy. Uvědomila si význam zdravotnických dat a statistické analýzy používala při psaní různých zpráv pro úředníky či členy vlády. Při analýze dlouhých tabulek se statistickými údaji, které považovala za dokonale občerstvující, využívala zejména popisnou statistiku, ale také normální rozdělení, počty a součty či vhodné vizualizace. Konkrétně využívala a popularizovala kruhový (též koláčový) diagram nebo vymyslela polární diagram (též nazý-

[2] `http://www.kingscollections.org/exhibitions/specialcollections/nightingale-and-hospital-design/florence-nightingale-and-hospital-design`

vaný polární oblastní diagram). Díky důrazu na srozumitelnost se jí díky statistice opravdu podařilo prosadit reformu britských nemocnic, třeba z hlediska prevence infekcí (zavedla častější větrání) nebo zlepšení čistoty a stravy.

Nightingaleová je dnes opěvována jako hrdinka pro své sociální cítění a láskyplnou pomoc pacientům a zraněným (např. osobně organizovala pomoc v krymské válce), kterou považovala za svou křesťanskou povinnost. Vedla vyčerpávající boj za své ideály nejen humanity, ale také statistiky, protože podle ní umožňuje studium statistiky poznávat Boží záměry ([21], str. 222) a představuje tak určitou formu bohoslužby ([22], str. 8). Stala se zakladatelkou moderního ošetřovatelství jakožto vědního oboru, ovlivnila vývoj epidemiologie a byla i předchůdkyní dnešních ideálů medicíny založené na důkazech (*evidence-based medicine*) ([1], str. 404). Jejich cílem je vědecký přístup k medicíně, který umožní zlepšení zdravotní péče na základě medicínských dat a statistických důkazů (*evidence*) získaných jejich analýzou.

Obrázek 2: Obraz „Florence Nightingaleová s umírajícím vojákem ve Scutari“ namaloval Alexis Soyer roku 1857, který na vlastní oči viděl, jak Nightingaleová pečuje o raněné v krymské válce (1853–1856).

4. Francis Galton (1822–1911)

Francis Galton byl všestranný vědec (obrázek 1)[3], profesor eugeniky v Londýně, který je uznáván jako jeden z geniálních a univerzálních vědců. V roce 1909 byl povýšen do rytířského stavu. Galtonovu rozsáhlou biografii napsal jeho hlavní žák a obdivovatel Karl Pearson (1857–1936) [28].

Galton byl se svým lékařským vzděláním výborně připraven studovat člověka z hlediska různých vědních oborů. Byl doslova posedlý měřením a kvantifikací a nejen ve své laboratoři, ale i ve svém volnu neustále měřil nebo počítal četnosti různých jevů ([26], str. 38). Charakterizujme nyní postupně ty vědy o člověku, kterým se nejvíc věnoval.

Antropologie. Galton byl vynikající antropolog, vedl i Královský antropologický ústav v Londýně. Založil Antropometrickou laboratoř v Londýně roku 1884, v níž se během 6 let podařilo naměřit tělesné rozměry zhruba 9000 osob. Zároveň se vyšetřovala jejich ostrost sluchu a zraku či reakční čas. Zde naměřená data porovnával Galton mezi různými povoláními nebo rasami a velmi usiloval o odhad inteligence jedince z fyzických rozměrů.

Daktyloskopie. Galton zpracoval kompletní sady otisků prstů získané u zhruba 8000 osob a stanovil fyziologické zákony papilárních linií na prstech. Z nich se snažil identifikovat jedince, přičemž potřeboval rozlišovat variabilitu mezi jedinci a vnitrotřídní variabilitu (pro opakovaná měření na téže osobě). Dále navrhl třídicí a registrační systémy pro databáze otisků prstů. Galton tak výrazně překonal snahu Alphonse Bertillona (1853–1914), který v Paříži usiloval o identifikaci jedinců (zločinců) na základě antropometrických měření (obrázek 1). Galtonovým kolegům se v Anglii také záhy podařilo prakticky realizovat takové systémy, které skutečně sloužily pro účely identifikace osob v kriminalistické praxi ([30], str. 161–164).

Psychologie. Galton založil nový vědní obor, jejíž náplní bylo měření psychických jevů ([36], str. 14). Pojmenoval jej názvem psychometrie (český též psychometrika). Sestavil testy duševní schopností, jejichž výsledky kombinoval s tělesnými měřeními, a studoval individuální rozdíly duševně nemocných, čímž překonal Queteletův koncept průměrného člověka. Galtonův přístup prezentovaný v knize [12], která představuje jeho hlavní dílo (nejen) z hlediska psychometrie, lze označit jako behaviorismus [15]. Například nezkoumal, co dotyčný jedinec prožívá ve své mysli, ale jeho výzkum začíná experimentem (měřením), pokračuje jeho analýzou a končí zobecněním na obecné zákonitosti, opět ve smyslu statistických důkazů.

[3] `https://fineartamerica.com/featured/1893-francis-galton-criminal-photograph-paul-d-stewart.html`

Analýza obrazu obličeje. Galton studoval průměrný obličej [31] získaný z fotografií obličeje různých osob a hledal rysy typického zločince. Jinou jeho snahou bylo určit z pouhé fotografie obličeje, zda má dotyčný jedinec sklony ke zločinu.

Evoluce. Galton rozvíjel principy evoluce a přírodního výběru aplikované na lidskou populaci. S evolucí se setkal nejprve ve své rodině, byl bratrancem Charlese Darwina (1809–1882) a myšlenku evoluce organismů poprvé rozpracoval už jejich dědeček Erasmus Darwin (1731–1802). Galtona zajímala dědičnost a její principy, ale zejména konkrétně dědičnost u člověka (ač sám zůstal bezdětný). V již zmíněné knize [12], která i z hlediska dědičnosti představuje jeho hlavní dílo, se silně věnuje dědičnosti inteligence. Rád by přispěl k tomu, aby se rodili geniální jedinci, a věřil v možnost zvyšovat inteligenci lidstva jako celku. Lidstvo je podle něj evolučně tvárné doslova jako jíl. Nadto si i uvědomoval, že v evoluci hraje svou roli náhoda, a to v nahodilém vývoji organismů, které se teoreticky mohly vyvíjet nějak jinak.

Eugenika. Galton založil eugeniku s cílem vědecky realizovat myšlenky o šlechtění lidí s dobrými biologickými vlastnostmi ([5], kap. 2). Z eugeniky vycházely ostatní jeho vědecké snahy. Eugenici chtěli zachránit lidstvo před chudobou, narůstající velikostí populace a údajně rostoucím podílem takzvaně slabomyslných[4], jejichž rozmnožování chtěli potlačit ([24], str. 32). K tomu chtěli mít pod kontrolou boj o přežití (přirozený výběr), ke kterému (podle nich) dochází v lidské populaci. Dodejme ovšem, že v té době už vznikala vědecká psychiatrie v moderním duchu [37], která oproti eugenice nechtěla omezovat pacienty na lidských právech ani ponižovat jejich hodnotu. Eugenika měla politické ambice ([5], kap. 14–16), usilovala o moc pro „intelektuálně nadřazené“, a zároveň šlo o alternativu vůči všem náboženstvím[5]. Galton také založil Eugenickou vzdělávací společnost a stal se jejím prvním prezidentem ([20], str. 524).

Galton je také považován za vůbec prvního, kdo měřil inteligenci vědecky založeným testem [8]. Za pomoci statistiky hledal rasy s nižší inteligencí a šířil názory, které dnes hodnotíme jako evidentní rasismus nebo trapné urážky žen ([17], str. 119). Galton se sice snažil statistickými argumenty oslovit politiky (jako Nightingaleová), ale vliv si nezískal, alespoň ne pozitivní. Před druhou světovou válkou byl ve Velké Británii zrealizován jen jediný cíl eugeniků, a to byly nucené sterilizace ([24], str. 40). Po Galtonovi se stal hlavním eugenikem Leonard Darwin (1850–1943), syn Charlese Darwina, a eugenická ideologie se záhy rozšířila po celém světě ([24], kap. 3.4). Buď měla za cíl

[4] `https://is.muni.cz/el/1431/jaro2008/Bi6170/um/Eugenika.pdf`
[5] `http://law2.umkc.edu/faculty/projects/ftrials/conlaw/eugenics.html`

diskriminovat nežádoucí jedince, za které byli pokládáni bezdomovci, alkoholici, postižení či duševně nemocní, nebo přímo eliminovat. Eugenika ovšem ve svých počátcích nebyla považována za odsouzeníhodnou. Sami eugenici přitom považovali své cíle za chvályhodné a svou vědu za hluboce mravně založenou [5]. V Německu byla eugenika následně zneužita pro vědecké obhajování nacistického programu. Odtud usuzujeme, že uplatňování vědeckého výsledku v praxi vyžaduje i rozvahu o etické dimenzi a že když je nějaký názor vědecky podpořen, tak nemusí být akceptovatelný ani správný.

4.1. Galton jako statistik

Galton dokázal výrazně ovlivnit statistiku svými idejemi, přestože je neformuloval exaktně matematicky ([33], str. 126). Ve skutečnosti ani nebyl dostatečně matematicky vzdělán; studoval nejprve matematiku v Cambridgi od roku 1840, ale ve třetím ročníku studia zanechal. Se statistikou se začal seznamovat až po roce 1869 ([4], str. 512), kdy cítil potřebu statistiku aplikovat ve svém výzkumu člověka. Ani jako viceprezident Královské statistické společnosti (od roku 1875) nejevil o tuto svou funkci zájem a nesdílel své myšlenky s jinými statistiky. Galton publikoval články s originálními statistickými postupy už od 70. let, přičemž ale hlavní své myšlenky, které nyní popíšeme, publikoval v krátkém tříletém období mezi roky 1885 a 1888 ([4], str. 527).

Zavedl korelační analýzu, přičemž studiem korelací na místo kauzality zcela změnil evropskou statistickou tradici. Poprvé korelační analýzu vysvětlil na antropologické úloze při porovnávání výšky dětí a délky jejich kostí [10], ale potřeboval ji primárně pro studium dědičnosti ([35], str. 75). Když se v roce 1888 přestal aktivně věnovat vědě, přenechal dobudování korelační analýzy mladším a Pearson pak získal dnešní (Pearsonův) korelační koeficient jen drobnou modifikací Galtonovy původní definice [32].

Dále Galton jako první použil pojem regrese, přičemž regresí nejprve rozuměl regresi k průměru. Pro odhad regresních parametrů používal Gaussovu metodu nejmenších čtverců. Galton studoval reverzi dědičnosti, tj. kladl si třeba otázku, jak vysocí jsou synové vysokých otců. Při té příležitosti pozoroval (a jako vůbec první popsal) regresi k průměru (*regression to the mean*, též filiální regrese): syn je (v průměru) ve své výšce blíž průměru než jejich otec, ale zároveň vysocí synové mají otce, kteří jsou v průměru nižší než oni sami. Tento zdánlivý paradox je výborně zpopularizován článkem [33]. Galton nejprve vnímal regresi v průměru jako genetickou záležitost a regresi nazýval reverzí (*reversion*), než si uvědomil, že jde o statistický fenomén ([13], str. 231).

Pro Gaussovo rozdělení zavedl Galton název *normal curve of distributions* nebo prostě *normal curve*[6], kde pojem normální zřejmě vztahuje k normě jako ke standardu, modelu, či vzoru. Jako první objevil dvourozměrné normální rozdělení; při jeho popisu však potřeboval požádat o pomoc matematika Hamiltona Dicksona ([4], str. 523). Krom toho Galton zavedl pojmy percentily (nejprve jako *centiles* či *per-centiles*) a kvantily, a také jejich různé speciální případy (kvartil, decil). V práci [9] je používá na míry naměřené ve své Antropometrické laboratoři. Definicí kvantilu překonal Galton intuici A. Queteleta, o kterém jsme již zmínili, že subjektivně hledal vhodné limity pro normálně rozdělené veličiny. Dřívější pojem chyba (*error*) pak Galton nahradil deviací (*deviation*) od normality. V eugenickém duchu pak vnímal jedince vybočující normě za (patologicky) nenormální, tj. deviantní.

Galtonovy biologické výsledky jsou z dnešního pohledu přísně hodnoceny ([17], str. 143). Zejména je mu vytýkán ideologický přístup nedostatečně podpořený vědeckou argumentací. Konkrétně se v literatuře uvádí, že přistupoval ke statistice primárně jako k nástroji pro podporu (potvrzení) svých eugenických představ, které zastával ještě před samotným prováděním měření ([4], str. 516; [20], str. 500; [26], str. 38). V literatuře bývá Galton označen jako statistik s primitivními metodami ([14], str. 120), zapálený eugenik bez nestranného vědeckého úsudku ([26], str. 52–54), vědec ovládaný politickou ideologií ([4], str. 509), či předpojatý kompilátor pochybných statistických údajů ([17], str. 115), jehož názory stejně zůstaly neměnné bez ohledu na statistické výsledky. Některé jeho závěry mohou působit z dnešního hlediska absurdně ([4], str. 525). Domníváme se, že to může být dáno i tím, že Galton nevhodně navrhoval své experimenty.

Galtonův přínos statistice byl každopádně výrazný, a to také díky tomu, že jeho myšlenky rozpracoval a zpopularizoval velmi vlivný Karl Pearson. Ten také založil a vedl anglickou biometrickou školu a Galton na sklonku svého života vystupoval jako její patron. Cílem biometrie bylo provádět měření vnějších fyziologických znaků na živých organismech včetně člověka. Nedílnou součástí biometrie bylo použití statistických metod pro zpracování takto získaných měření s cílem studovat variabilitu živých organismů. Zatímco pojem biometrie užíval již Galton, teprve Pearson dokázal pojem biometrie naplnit konkrétním obsahem, zejména důrazem na statistické metody, k jejichž rozvoji biometři nebývale přispěli [23]. Když byl Galton mrtev, Pearson vyznal, že smyslem jeho života je pracovat na Galtonově znovuobživení (tj. na propagování jeho myšlenek) ([28], předmluva k 1. dílu).

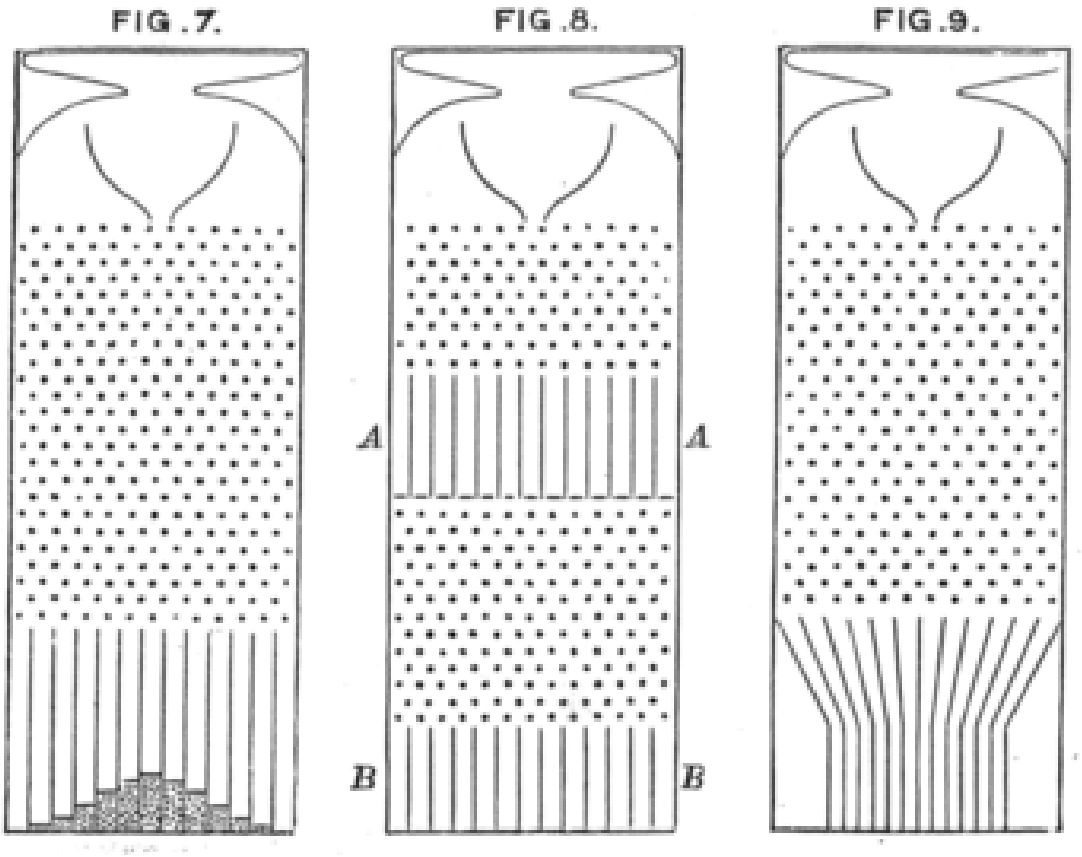
[6] `http://condor.depaul.edu/ntiourir/NormalOrigin.htm`

4.2. Galtonovo prkno

Galton navrhl tři verze mechanické pomůcky, kterou dnes nazýváme Galtonovým prknem (anglicky *quincunx*, *Galton board*, *bean machine*). Původní název pro tato dřevěná prkna (viz obrázek 3)[7], jejichž principy zde s využitím současného pravděpodobnostního formalismu popíšeme, byl *Instrument to illustrate the principle of the Law of error or Dispersion*. Zákon chyb byl přitom tehdejší název pro centrální limitní větu, která Galtona fascinovala, a jeho hlavní motivací vskutku bylo sestrojit nástroj pro její vizualizaci.

Jednoduché Galtonovo prkno (obrázek 3 vlevo) elegantně demonstruje, jak lze pomocí padajících kuliček znázornit pravděpodobnostní funkci binomického rozdělení. Při pádu se každá kulička odráží od hřebů zatlučených v prkně, přičemž odraz vlevo i vpravo má shodnou pravděpodobnost $\pi = \frac{1}{2}$. V přihrádkách ve spodní části prkna se hromadí kuličky. Počet řad hřebů nad přihrádkami označme pomocí n. Pokud by v ideální situaci mělo prkno také

[7]Bulmer M. (2003): *Francis Galton, Pioneer of Heredity and Biometry.* The Johns Hopkins University Press, Baltimore.



Obrázek 3: Galtonova kresba jednoduchého, dvojitého a konvergentního Galtonova prkna [11].

n přihrádek, pak podíl kuliček v i-té přihrádce přibližně odpovídá pravděpodobnostní funkci binomického rozdělení

$$p_i = \binom{n}{i} \pi^i (1-\pi)^{n-i} = \binom{n}{i} \left(\frac{1}{2}\right)^n ,$$

kde $i \in \{0, 1, 2, \ldots, n\}$. Galtonovo prkno na obrázku 3 vlevo však nemá přihrádky s nejnižšími a nejvyššími indexy, aby nemuselo být příliš široké. Pro toto prkno s $n = 23$ řadami hřebů tedy platí $i \in \{5, 6, 7, \ldots, 17\}$. Pro velký počet řad hřebů (tj. pro velké n) je pak tvar pravděpodobnostní funkce dané hodnotami p_i blízký Gaussově křivce, což názorně demonstruje centrální limitní větu.

Dvojité Galtonovo prkno (obrázek 3 uprostřed), které demonstruje složitější princip, je navrženo také překvapivě jednoduše. Prezentuje konvoluci dvou binomických rozdělení, a to velice elegantně jako

$$r_k = \sum_i p_i q_{k-i} \ ;$$

tento součet představuje vzorec pro úplnou pravděpodobnost, v němž q_{k-i} je podmíněnou pravděpodobností dopadu kuličky do k-té přihrádky v prostoru B za podmínky, že kulička padla do i-té přihrádky v prostoru A. Pravděpodobnost této podmínky je p_i. Výsledná pravděpodobnostní funkce $r_k = P(I + J = k)$ pak vyjadřuje diskrétní rozdělení náhodné veličiny $I + J$, kde náhodné veličiny I a J odpovídají binomickému rozdělení v horní, resp. spodní části desky. Prkno je tedy ukázkou mechanického počítače, který provádí stochastický výpočet konvoluce dvou pravděpodobnostních funkcí.

Konečně konvergentní Galtonovo prkno (obrázek 3 vpravo) názorně ukazuje zpřesnění původní náhodné veličiny při zmenšení velikosti odchylek od průměru. Zúžením šířky přihrádek koeficientem $s \in (0, 1)$ vznikne při $n \to \infty$ Gaussova křivka, jejíž směrodatná odchylka je s-násobkem směrodatné odchylky původní Gaussovy křivky přibližně zobrazené na obrázku 3 vlevo. Konvergentní Galtonovo prkno tedy představuje transformaci náhodné veličiny při změně měřítka.

Galtonova prkna později později posloužila jako základ některých mechanických hracích automatů. Za zhlédnutí stojí i videozáznamy Galtonova prkna na internetu (i když u některých videí se jedná o efektní padělek)[8]. Při popisu svých tří verzí prkna v práci [11] vyslovil Galton přiléhavé vyznání:

> Je obtížné pochopit, proč statistici často omezují svá šetření na průměry, a nelibují si v komplexnějších pohledech. Jejich duše

[8] `https://www.youtube.com/watch?v=S5WAswaJRjw`

> se zdají být tak otupělé k půvabu rozmanitosti jako onen [...], jehož pohled na Švýcarsko byl takový, že kdyby se jeho hory mohly naházet do jezer, odstranily by se dvě nepříjemnosti najednou. [...] (Statistika) je jediným nástrojem, jímž lze prosekat otvor do hrozivého houští potíží bránících v cestě těm, kdo se věnují Vědě o člověku.

Zřejmě zde má Galton na mysli, že právě normální křivka má oproti průměru schopnost získat z dat bohatší informaci.

Závěrem můžeme říci, že statistické uvažování v 19. století si vystačilo s elementárním pojetím pravděpodobnosti a ani nepotřebovalo definovat náhodnou veličinu. Společným znakem Queteleta, Nightingaleové a Galtona byl zájem o normalitu a její praktické využití ve vědách o člověku nebo i jejich vášeň pro sběr dat a jejich vyhodnocování. Pokud jde o odlišnosti uvedených osobností, Galton změnil evropskou statistickou tradici přechodem od studia kauzálních vztahů ke studiu korelací i faktickým položením základů mnohorozměrné statistiky, kterou následně matematicky rozvinul Karl Pearson. Také z hlediska společenské angažovanosti vnímáme výrazné odlišnosti mezi Nightingaleovou, která statistikou pomáhala nemocným a trpícím, a Galtonem, který statistikou obhajoval eugenické argumenty v neprospěch nemocných nebo chudých; navíc byla eugenika později zneužita pro obhajobu nacistického teroru.

Poděkování

Článek vznikl v rámci Centra Karla Čapka pro studium hodnot ve vědě a technice (CEVAST). Děkujeme dvěma recenzentům za pečlivé přečtení rukopisu a hodnotné náměty na zlepšení článku.

Literatura

[1] Aravind M., Chung K. C. (2015): *Evidence-based medicine and hospital reform: Tracing origins back to Florence Nightingale.* Plastic and Reconstructive Surgery **125**, 403 – 409.

[2] Bostridge M. (2009): *Florence Nightingale: The woman and her legend.* Penguin Books, London.

[3] Caponi S. (2013): *Quetelet, the average man and medical knowledge.* História, Ciências, Saúde–Manguinhos **20**, 830 – 847.

[4] Cowan R. S. (1972): *Francis Galton's statistical ideas: The influence of eugenics.* Isis **63**, 509 – 528.

[5] Darwin L. (1936): *Co je to eugenika?* Lékařské knihkupectví a nakladatelství, Praha.

[6] Diamond M., Stone M. (1981): *Nightingale on Quetelet.* Journal of the Royal Statistical Society A **144**, 66 – 79.

[7] Donnelly K. (2015): *Adolphe Quetelet, social physics and the average men of science, 1796–1874.* Routledge, London.

[8] Flanagan D. F., Kaufman A. S. (2004): *Essentials of WISC-IV assessment.* Hoboken, Wiley.

[9] Galton F. (1885): *Anthropometric per-centiles.* Nature **31**, 223 – 225.

[10] Galton F. (1888): *Co-relations and their measurement, chiefly from anthropometric data.* Proceedings of the Royal Society of London **45**, 135 – 145.

[11] Galton F. (1889): *Natural inheritance.* Macmillan, London.

[12] Galton F. (1892): *Hereditary genius. An inquiry into its laws and consequences.* 2. vydání. Macmillan, London.

[13] Gorroochurn P. (2016): *On Galton's change from "reversion" to "regression".* American Statistician **70**, 227 – 231.

[14] Hald A. (2006): *A history of parametric statistical inference from Bernoulli to Fisher, 1713 to 1935.* Springer, New York.

[15] Hankins F. H. (1908): *Adolphe Quetelet as statistician.* Columbia University, New York.

[16] Jahoda G. (2015): *Quetelet and the emergence of the behavioral sciences.* Springerplus **4**, Article 1.

[17] Johnson P. (2012): *Darwin. Portrét génia.* Barrister & Principal, Brno.

[18] Kalina J., Soukup L. (2018): *Doktrína šancí: 300. výročí první učebnice teorie pravděpodobnosti.* Informační bulletin České statistické společnosti **29**(1), 1 – 11.

[19] Louckx K., Vanderstraeten R. (2014): *State-istics and statistics: exclusion categories in the population census (Belgium, 1846–1930).* The Sociological Review **62**, 530 – 546.

[20] MacKenzie D. (1976): *Eugenics in Britain.* Social Studies of Science **6**, 499 – 532.

[21] Magnello M. E. (2006): *Victorian vital and mathematical statistics.* BSHM Bulletin: Journal of the British Society for the History of Mathematics **21**, 219 – 229.

[22] Maindonald J., Richardson A. M. (2004): *This passionate study: A dialogue with Florence Nightingale.* Journal of Statistics Education, **12**(1), 1 – 13.

[23] Malý M. (2003): *Biometrika stoletá.* Informační bulletin České statistické společnosti **14**(1), 4 – 9.

[24] Matiásková A. (2012): *Bioetika III—Eugenika—historie a současnost.* Univerzita Karlova v Praze, Hradec Králové. Diplomová práce.

[25] Nesselroade J. R., Cattell R. B. (1988): *Handbook of multivariate experimental psychology.* 2. vydání. Plenum Press, London.

[26] Neužilová I. (2013): *Přínos Francise Galtona vědám o člověku.* Západočeská univerzita v Plzni, Plzeň. Bakalářská práce.

[27] Parshall K. H. (2006): *James Joseph Sylvester: Jewish mathematician in a Victorian world.* The Johns Hopkins University Press, Baltimore.

[28] Pearson K. (1914, 1924, 1930): *The life, letters and labours of Francis Galton I, II, III.* Cambridge University Press, Cambridge.

[29] Quetelet A. (1835): *Sur l'homme et le développement de ses facultés, ou Essai de physique sociale.* Bachelier, Paris.

[30] Rak R., Matyáš V., Říha Z. a kol. (2008): *Biometrie a identita člověka ve forenzních a komerčních aplikacích.* Grada, Praha.

[31] Rakover S. S., Cahlon B. (2001): *Face recognition: Cognitive and computational processes.* John Benjamins Publishing Company, Amsterdam.

[32] Rodgers J. L., Nicewander W. A. (1988) *Thirteen ways to look at the correlation coefficient.* American Statistician **42**, 59 – 66.

[33] Senn S. (2011): *Francis Galton and regression to the mean.* Significance **8**, 124 – 126.

[34] Sheynin O. (2004): *Fechner as a statistician.* British Journal of Mathematical and Statistical Psychology **57**, 53 – 72.

[35] Stigler S. M. (1989): *Francis Galton's account of the invention of correlation.* Statistical Science **4**, 73 – 79.

[36] Šajtarová N. (2016): *Francis Galton a jeho přínos k psychometrii.* Univerzita Palackého v Olomouci, Olomouc. Bakalářská práce.

[37] Šedivec V. (2008): *Přehled dějin psychiatrie.* Psychiatrické centrum Praha, Praha.

[38] Závodský P. (1992): *Vývoj statistické teorie na území Československa do roku 1948.* Federální statistiký úřad a Infostat, Praha.

POSTŘEHY ZE SETKÁNÍ PŘEDSTAVITELŮ NÁRODNÍCH STATISTICKÝCH SPOLEČNOSTÍ ZEMÍ V7 V LUBLANI

IMPRESSIONS FROM THE V7 STATISTICAL SOCIETIES MEETING IN LJUBLJANA

Ondřej Vencálek

E-mail: ondrej.vencalek@upol.cz

Dne 27. září 2019 se ve slovinské Lublani uskutečnilo setkání zástupců statistických společností zemí V7 – České republiky, Maďarska, Polska, Rakouska, Rumunska, Slovenska a Slovinska. Letos bylo zastoupeno všech sedm zemí.

Toto setkání konané každoročně hostí vždy jedna ze zemí skupiny V7. Jeho smyslem je načerpat inspiraci od kolegů ze zahraničí (zjistit co a jak dělají kolegové ze statistických společností v okolních zemích) a prostřednictvím setkávání navazovat nové osobní kontakty či upevňovat kontakty stávající.

Oficiální část programu sestávala z vystoupení zástupců jednotlivých statistických společností, v nichž představili činnost svých spolků. Účastníci zmiňovali zejména

- konference, které jejich společnosti pořádají (některé jsou mezinárodní),
- periodika, ktará vydávají,
- soutěže, které pořádají a ceny, které udělují. Jde např. o soutěže o nejlepší bakalářskou či diplomovou práci (Rakousko, Polsko, Slovensko), udělovány jsou však také ceny za popularizaci statistiky (cena Gerharta Bruckmanna v Rakousku) či za správné používání statistických metod (Slovinsko – cenu obdržela skupina žurnalistů Pod črto), polská statistická společnost dokonce organizuje statistickou olympiádu pro studenty středních škol,
- univerzitní vzdělávání v oblasti statistiky. Zástupci slovinské a rakouské společnosti zmínili existenci studijního programu EMOS (European Master in Official Statistics), který vzniká za účasti statistických úřadů těchto zemí. V ČR tento program garantuje VŠE Praha.

Zástupci jednotlivých společností rovněž „kvalifikovaně odhadli“ počet členů svých společností. Shodli se přitom, že počet členů bývá obvykle výrazně větší než počet aktivních členů (ti tvoří jen malou část – většinou pětinu nebo čtvrtinu celkového počtu oficiálně registrovaných členů). Hrubé odhady

počtu členů: ČR 200, Maďarsko 400, Polsko 700, Rakousko 300, Rumunsko 300, Slovensko 250, Slovinsko 80.

Náměty k zamyšlení:

- Četné soutěže a různá ocenění jistě přispívají k popularizaci statistiky. ČStS uděluje ocenění nejlepších prací studentů doktorského studia v rámci konference Robust. Ceny pro studenty na bakalářské či magisterské úrovni většinou organizují jednotlivá pracoviště samostatně (ať už na úrovni katedry, fakulty či univerzity). Možná by bylo dobré jednak na stránkách ČStS uvádět výsledky lokálních soutěží, jednak zvážit zavedení dalších ocenění a formu jejich předávání.
- Opakovaně byl zmiňován mezinárodní statistický den (20. 10.), ČStS zatím tuto popularizační akci „nevzala za svou“. Upřesněme, že tzv. World Statistics Day[1] je organizován od roku 2010 jednou za pět let – příští bude 20. 10. 2020, od roku 2016 se však každoročně slaví tzv. European Statistics Day[2].

Na závěr setkání byla krátce diskutována chystaná akreditace evropských statistiků. Ač má tato iniciativa mezinárodního sdružení FENStatS podporu členů společností zemí V7, o zavedení této akreditace v jednotlivých zemích se bude teprve na národní úrovni jednat.

[1] https://en.wikipedia.org/wiki/World_Statistics_Day
[2] https://ec.europa.eu/eurostat/web/ess/european-statistics-day

October 11-13: conference STAKAN
• STAKAN = STAtističti KANtoři (teachers of statistics)
• About 50 participants
• 5 sections ~ 5 topics ~ 5 invited lectures
Statistično
društvo
Slovenije
@StatSlovenija
primerljivost
podatki
šifranti
baze
infograf

KAKO SE

POZVÁNKA NA KONFERENCI ROBUST 2020
INVITATION TO THE ROBUST 2020 CONFERENCE

Redakce

ROBUST 2020
20+20=40

Vážené kolegyně, vážení kolegové,

dne 7. června 2020 se v Bardějově uskuteční slavnostní zahájení dvacáté první letní školy JČMF ROBUST 2020. Připomeňme, že série letních a zimních škol ROBUST je organizována skupinou pro výpočetní statistiku při **ČMS JČMF** obrok od roku 1980, takže se s řadou z Vás sejdeme po neplánovaných a neuvěřitelných **čtyřiceti letech**. Letos naši akci mimo jiné podpořila **ČStS**, **SŠDS** a **KPMS MFF UK**.

ROBUST 2020 bude věnován, tak jako vždy, současným trendům statistiky, pravděpodobnosti, finanční matematiky, optimalizace a analýzy dat. Vedle toho bude velká pozornost věnována problematice výuky našich oborů v době Internetu a existence volného přístupu k nástrojům jakými jsou Wikipedie, Google Search a Google Scholar, sytém R, apod., svádějících k mylné představě (nejenom studentů), že si řešení snadno a bezbolestně „vygůglíme“. Této problematice budou věnovány nejenom Vaše přednášky, ale i připravovaný panel.

Vedle zvaných přednášek se jako vždy počítá se sděleními účastníků a posterovou sekcí doktorandů a studentů. Nechte se překvapit, kdo přijal nabídku k přednesení zvaných přednášek. Intenzivní jednání probíhají.

- *Datum a místo konání*: 7. – 12. června 2020 (ne-pá), **Bardějov**; GPS $49°17'33.566''N, 21°16'16.726''E$.
- *Ubytování*: Ve dvou a tří lůžkových pokojích. Vzhledem k odlehlosti místa konání pro zájemce zajistíme za příplatek ubytování již od soboty 6. června 2020.
- *Stravování*: Celodenní (neděle večeře – pátek oběd).
- *Kulturní program*: Na středu plánujeme výlet po **dřevěných chrámech východoslovenských Karpat**, jež jsou zařazeny do světového dědictví UNESCO, a po stopách **karpatsko-dukelské operace**.
- *Technické vybavení*: K dispozici bude notebook, zpětný a datový projektor, tabule a wifi připojení.
- *Přihláška a registrace*: Registrace je otevřena na adrese `http://robust.nipax.cz`. Pro ty, kteří se zúčastnili předchozích Robustů, jsme zachovali již dříve vložené osobní údaje. Zkontrolujte je a v případě změn opravte. Při jakýchkoliv potížích s registrací se obraťte na kolegu Dohnala (`dohnal@nipax.cz`).
- *Abstrakt*: **Na svoji přednášku myslete již dnes**, a abstrakt v TEXu spolu s pdf souborem vložte přes Vaši registrační stránku na adrese `robust.nipax.cz` **nejpozději 31. března 2020**.
- *Stipendia*: Česká statistická společnost vypsala několik stipendií. Podrobnosti viz `www.statspol.cz`.
- *Konferenční poplatek*: 5 000 Kč. Pro studenty řádného studia a interní postgraduální studenty bez vedlejšího úvazku 4 800 Kč. Konferenční poplatek zahrnuje ubytování, stravu a náklady organizačního výboru.
- *Bankovní spojení*: Číslo účtu 2001215985/2010 u Fio banka, a.s. IBAN: CZ77 2010 0000 0020 0121 5985 BIC: FIOBCZPPXXX. Příjemce: Česká matematická společnost, sekce Jednoty Českých matematiků a fyziků, Žitná 609/25, 110 00 Praha 1. Variabilní symbol ve tvaru 20205xxx, kde xxx si zvolte sami, si nezapomeňte zapamatovat a vložit do Vaší registrační stránky!
- *Poznámky k placení*: Jako informaci pro příjemce platby uveďte Vaše jméno, případně instituci. Daňový doklad obdržíte během konference. Pokud potřebujete zálohovou fakturu dříve, například kvůli možnosti Vaši účast proplatit, obraťte se na paní Naxerovou (`naxerova@karlin.mff.cuni.cz`). Budete-li platit až na místě, dohodněte se s námi předem, abychom včas zajistili kapacitu.
- *Storno poplatky*: Při zaplaceném konferenčním poplatku a odřeknutí účasti:
  – Mezi 15. – 31. květnem 2020 je storno poplatek 25 %.
  – 1. června 2020 či později je storno poplatek 50 %.
  – Nedostaví-li se účastník, je storno poplatek 100 %.
  Pro stanovení data pro zrušení účasti je rozhodující datum informující organizátory o zrušení účasti.
- *Další oznámení*: Veškeré informace budou zveřejňovány na `www.karlin.mff.cuni.cz/~antoch`, a dle potřeb distribuovány buď e-mailem nebo klasickou poštou.
- *Adresa pro korespondenci*: ROBUST 2020, MFF UK, KPMS, Sokolovská 83, 186 75 Praha 8 – Karlín, tel. 221 913 287; e-mail: `antoch@karlin.mff.cuni.cz`.

Na setkání se těší: J. Antoch, G. Dohnal, D. Hlubinka, I. Stankovičová a – především – účastníci.

V Praze na Svatého Vendelína roku 2019

Obsah

Vědecké a odborné články

Jan Kalina, Lubomír Soukup
Průkopníci statistiky ve vědách o člověku v 19. století 1

Zprávy a informace

Ondřej Vencálek
Postřehy ze setkání představitelů národních statistických společností zemí V7 v Lublani 16

Redakce
Pozvánka na konferenci ROBUST 2020 20

Informační bulletin České statistické společnosti vychází čtyřikrát do roka v českém vydání. Příležitostně i mimořádné české a anglické číslo. Vydavatelem je Česká statistická společnost, IČ 00550795, adresa společnosti je Na padesátém 81, 100 82 Praha 10. Evidenční číslo registrace vedené Ministerstvem kultury ČR dle zákona č. 46/2000 Sb. je E 21214.

The Information Bulletin of the Czech Statistical Society is published quarterly.
The contributions in the journal are published in English, Czech and Slovak languages.

Předseda společnosti: Mgr. Ondřej Vencálek, Ph.D., Katedra matematické analýzy a aplikací matematiky, Přírodovědecká fakulta Univerzity Palackého, 17. listopadu 12, 771 46 Olomouc, e-mail: `ondrej.vencalek@upol.cz`.

Redakce: prof. RNDr. Gejza Dohnal, CSc. (šéfredaktor), prof. RNDr. Jaromír Antoch, CSc., doc. RNDr. Zdeněk Karpíšek, CSc., RNDr. Marek Malý, CSc., doc. RNDr. Jiří Michálek, CSc., prof. Ing. Jiří Militký, CSc., doc. Ing. Iveta Stankovičová, PhD., doc. Ing. Josef Tvrdík, CSc., Mgr. Ondřej Vencálek, Ph.D.

Redaktor časopisu: Mgr. Ondřej Vencálek, Ph.D., `ondrej.vencalek@upol.cz`.
Informace pro autory jsou na stránkách společnosti, `http://www.statspol.cz/`.

DOI: 10.5300/IB, http://dx.doi.org/10.5300/IB
ISSN 1210–8022 (Print), ISSN 1804–8617 (Online)

Toto číslo bylo vytištěno s laskavou podporou Českého statistického úřadu.