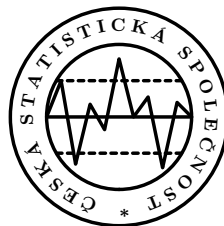


Informační Bulletin



České statistické společnosti

číslo 4, ročník 18, 24. 12. 2007

Prof. Ing. JAROSLAV JÍLEK, CSc.

*** 1931 – † 2007**

S lítostí oznamujeme, že ve středu 7. listopadu zemřel ve věku 76 let jeden ze zakládajících členů České statistické společnosti prof. Ing. Jaroslav Jílek, CSc.

7. listopadu 2007 opustil nečekaně řady statistiků na Fakultě informatiky a statistiky VŠE jeden z nejstarších a nejvýznamnějších z nich – prof. Jaroslav Jílek.

Narodil se 8. února 1931 v Olomouci, kde absolvoval obchodní akademii. Poté studoval v letech 1949-1954 na Vysoké škole ekonomické v Bratislavě obor ekonomika průmyslu, a na zdejší katedře statistiky a matematiky začal hned po promoci svou pedagogickou dráhu. Obhájil zde i svou kandidátskou disertační práci (1963). V té době už pracoval na katedře statistiky VŠE v Praze (od r. 1960) a tomuto pracovišti zůstal věrný po celých 47 let. Po odchodu doc. Kašpara na Statistický úřad OSN v New Yorku (1966) se stal hlavním představitelem oboru ekonomické statistiky na VŠE. Samostatnou katedru získal tento obor při reorganizaci VŠE v r. 1990.

Prof. Jílek se věnoval zejména problematice měření produkce, zvláště systémovému začlenění meziodvětvových bilancí. Jeho publikační činnost je velmi rozsáhlá, obsahuje více než dvě stovky prací z různých oblastí hospodářské statistiky – včetně čtyř vysokoškolských učebnic, několika desítek učebních textů, řady výzkumných prací i řešení grantových projektů. Habilitační práci obhájil již v roce 1968, profesorem mohl být jmenován až roku

1990 po pádu komunistického režimu. Na VŠE působil mj. jako předseda celoškolské akreditační komise, člen vědecké rady VŠE i Fakulty informatiky a statistiky, opakovaně byl zvolen členem senátu této fakulty.

Prof. Jílek už od 60. let spolupracoval se státní statistickou službou – nejprve především v problematice meziodvětvových (Input-Output) bilancí, poté v oblasti statistického výzkumu. V letech 1968-1970 byl zástupcem ředitele Výzkumného ústavu statistiky a účetnictví a v letech 1990-1992 statutárním místopředsedou Federálního statistického úřadu. Po jeho zrušení v souvislosti se zánikem federace působil prof. Jílek jako expert v sekretariátu předsedy Českého statistického úřadu, často jako renomovaný odborník zastupoval českou statistiku v zahraničí.

Dlouhá léta pracoval prof. Jílek i v redakčních řadách odborných časopisů – *Statistika* (v letech 1991-2002 jako předseda redakční rady), *Politická ekonomie* a *Statistics in Transition*.

V profesorovi Jílkovi ztratili čeští statistici jednoho z nejvýznamnějších a nejzasloužilejších odborníků, jeho přátel z VŠE oblíbeného kolegu a nejmladší adeпти statistiky obětavého učitele a ochotného rádce.

Čest jeho památce.

Výbor České statistické společnosti

Poznámka redakce

Řada členů České statistické společnosti může potvrdit, že profesor Jílek byl aktivním diskutujícím na setkání odborníků a zástupců českých vysokých škol s profesorem Peterem Hacklem v Praze ještě 11. října 2007. Je to až k nevíře!

STAKAN 2007 – DOJMY A REFLEXE

Jaromír Běláček

Na seminář *STAKAN 2007* konaný ve dnech 25.–27. 5. 2007 mě pozval můj dlouholetý kamarád, přítel (a t. č. vedoucí) Honza K. Existenci semináře s výše uvedeným názvem jsem zaregistroval již dříve na seznamu akcí pořádaných Slovenskou štatistickou a demografickou spoločnosťou (zkr. SŠDS – viz www.sdds.sk), kterou již po delší dobu považuji – podobně jako Českou štatistickou spoločnosť (zkr. ČStS – viz také www.statapol.cz) dnes – za reprezentativní a důstojnou odbornou platformu demografie a statistiky (nejen pro Slovensko). Že název semináře znamená něco jako *STATistika & KANtoři*, jsem mohl Honzovi do té doby pouze věřit. Právě tak by totiž samotný název semináře mohl někomu podsouvat *SKLENICI, ne nutně neprázdnou...* Obzvláště když se seminář měl konat někde v lesích na moravsko-slovenském pomezí (přesněji v rekreačním zařízení Jestřábí-Rusava nedaleko Bystřice pod Hostýnem).

Odborný seminář byl *naplněn* především překvapivě širokým spektrem profesionálů z oblasti teorie, výuky a aplikací statistiky, povětšinou z českých a slovenských vysokých škol. Pro ilustraci uvádím na semináři aktivně působící reprezentace za ČStS (prof. RNDr. Jaromír Antoch, CSc., z MFF UK Praha, bývalý předseda ČStS; doc. RNDr. Gejza Dohnal, CSc., z FS ČVUT, stávající předseda ČStS), za SŠDS (Doc. Ing. Jozef Chajdiak, CSc., – vědecký tajemník – a další přítomní kolegové ze SR, vesměs z Univerzity Komenského v Bratislavě), z pozvaných přednášejících za VŠE Praha (prof. Ing. Petr Hebák, CSc.; Doc. Ing. Hana Řezanková, CSc., a další) a z regionu organizátorů semináře (prof. Ing. Zdeněk Molnár, CSc., a Ing. Pavel Stříž, Ph.D., z Univerzity Tomáše Bati ve Zlíně). Výše uvedenému odpovídá rovněž spektrum problematik referovaných účastníky v rámci cca 25 přednesených příspěvků. Protože ze semináře byl pořizován videozáznam a většina příspěvků byla publikována ve slovenském časopise *Forum Statisticum Slovaca*, omezím se dále jen na to, co mě osobně nejvíc oslovilo a zaujalo.

Výuka statistiky na VŠ je v současné době *syčená* již zcela samozřejmou dostupností statistického software s širokou podporou internetových aplikací. Mělo by tedy být rovněž samozřejmostí, že se studenti VŠ naučí používat všechny základní statistické metody potřebné pro jejich budoucí praxi. Jak podtrhl v rámci syntetické úvodní přednášky „Výuka statistiky dříve, dnes a zítra“ prof. P. Hebák, je toto samozřejmě záležitostí nejen zájmu samotné profesionální obce statistiků, ale rovněž zodpovědností vysoce postavených manažerů vědy, výzkumu a vzdělávání, i kompetentních vládních představitelů, aby se s mnohem větší měrou *opájeli* objektivními statistickými daty,

neboť se nacházíme v situaci, kdy samotná matematika je (cit.) „... potlačována anebo dokonce vytlačována z výuky na středních školách ...“. Rozvoj statistiky samotné by měl vycházet ze stejných teoretických základů, na kterých byla postavena již v 19. a 20. století (perspektivy v syntéze bayesovských a „klasických parametrických“ postupů?).

Certifikovanou možností spolupráce mezi statistickými orgány a vysokými školami by bylo možno *zapít* např. ve vztahu k nedávno podepsané Dohodě mezi ČSÚ a UTB ve Zlíně (Hrbáček, Stříž). Metody statistické analýzy, Aplikovaná statistika, Počítačové zpracování dat nebo Ekonometrie jsou předměty vyučované podle „starých i nových“ osnov (Klímek, Stříž). V těchto souvislostech byli účastníci semináře seznámeni rovněž s rozšířením možností, které skýtá v oblasti statistiky nejnovější nástavba programové aplikace Microsoft Excel (přítomnými expertně oceněno jako přínosné). Od března 2007 bylo na FaME UTB vytvořeno centrum, které si klade za cíl „propojení aktivit na poli aplikovaného ekonomického výzkumu“ (Molnár). Těmto závěrům konvenovalo rovněž vystoupení J. Chajdiaka, kde byla podtržena úloha statistiky: „Ukázat svět, jaký je ... s cílem: vytvářet podporu, resp. doporučení procesu objektivizace manažerského rozhodování.“

Řada příspěvků prezentovaných na semináři měla poměrně sofistikovaný a specializovaný *statistický obsah*. Toto se týkalo například přednášky prof. Wimmera (Matematický ústav SAV) „Matematické modelování v jazykovede“, kde bylo představeno rozsáhlé spektrum aplikací speciálních pravděpodobnostních rozložení, profi-přednášky M. Urbaníkové „Využitie štatistiky v poistnej matematike“, referencí H. Řezankové (VŠE Praha) týkající se „Výuky jednorozměrné a dvourozměrné analýzy kategoriálních dat“, oživení „Metody latinských čtverců“ v podání prof. D. Jaruškové (ČVUT) anebo přednášky doc. Tvrdíka (Ostravská univerzita) „Grafy a tabulky ve statistice (aneb na co ve výuce většinou není čas)“ všímající si aspektů ne zcela postačujících a adekvátních způsobů prezentace statistických dat a výsledků analýz.

Až překvapivě *osamoceně* (z hlediska informatiky) na mne zapůsobilo na této platformě vystoupení doc. Žvávka (VŠE Praha) „Statistické výpočetní prostředí 2007“, ale i M. Dolejšové (UTB Zlín – její vystoupení pod názvem „Zařazení geografických informačních systémů do výuky předmětu Informatika ve veřejné správě“), které sice odlišným, ale zcela transparentním způsobem reflektovaly internetové směry a potřeby vizuálního rozvoje statistiky (ve světě, a tedy i v ČR?)... Že by jen *do prázdna* měl vyznít návrh doc. Žvávka, že by i profesionálové z ČR mohli být více aktivní při rozvoji a tvorbě některých speciálních stránek typu www.wikipedia.org v oblasti statistiky? (Pravděpodobně jsme to doposud považovali za záležitost více-

méně individuálního přístupu k uvedené problematice.) V tomto smyslu zapůsobilo nepochybně hodnověrně vystoupení M. Vojtkové a I. Stankovičové (UK Bratislava) „Viackriteriálne hodnotenie krajín EU na základe vybraných ukazateľov Lisabonskej stratégie“ vedoucí k typologii zemí EU podle objektivních rozvojových a sociálně-ekonomických ukazatelů.

Aplikace v medicíně byly na *STAKANU 2007* pokryty vystoupeními M. Malého (Státní zdravotní ústav) „K otázkám výuky a statistických konzultací“ zhodnocující jak praktické, tak i teoretické zkušenosti z biomedicíny a J. Běláčka (1. LF UK Praha + VFN) usilujícího o demonstraci obdobné problematiky v rámci příspěvku s názvem „Analýza laterality u dětí ve věku 9-11 let (aplikace statistických metod)“. Na *zdraví* se ve své aplikační části orientovalo rovněž vystoupení M. Žambochové (FSE ÚJEP v Ústí nad Labem) „Jak na rozhodovací stromy“ (zde byla shrnuta dlouhodobě nedostatečně průhledná metodika analýzy dat prostřednictvím regresních stromů) a přednáška I. Saxla (MFF UK Praha) „Alfred Binet a počátky testování inteligence“, která překročila rámec statistiky až do oblasti *historie psychiatrie*.

Kdybych byl býval věděl v jaké atmosféře na nás letošní *STAKAN čeká*, tak bych určitě na tuto akci přijel o den dříve. . . Z hektického běhu každodenních pracovních povinností byli všichni účastníci doslova od prvního vstupu *napojeni* na poklidnou a konstruktivně rehabilitující diskusní atmosféru, což bylo samozřejmě zásluhou všech oficiálních i neoficiálních pořadatelů akce. Týkalo se to nejen dopravy na místo, ubytování, stravování a vydařeného počasí, ale i organizace kulturního programu: vystoupení Taneční skupiny Kirké předvádějící břišní tance, pátečního táboráku, sobotního odpoledního výletu do okolí Rusavy. Do této kategorie lze přiřadit i vizuálně velmi atraktivní přednášku „Velkorozměrné struktury Vesmíru – zpracování dat o polohách galaxií“ (J. Vašátko – Zlínská astronomická společnost), která byla pro zájemce rozšířena o praktické pozorování objektů noční oblohy. Hvězdářským dalekohledem zvětšená planeta Jupiter (*k vidění dvakrát* se svými čtyřmi největšími měsíci se změnou jejich vzájemného postavení), na mě rozhodně zapůsobila pozitivněji, než by dokázala sklenice slivovice (*viděná nadvakrát* v obdobném postavení). . .

Na závěr se omlouvám všem přednášejícím (a zejména těm), jejichž vystoupení na semináři *STAKAN 2007* bylo dostatečně smysluplné a originální a v mém rádobychaotickém přehledu o nich nebyla zmínka. Účelem tohoto shrnutí nebyla *úplnost* nebo *úsilí o plnou* objektivitu (a už vůbec ne *degustace či hodnocení!*). Výuka statistiky v ČR má nepochybně silný a odborně erudovaný potenciál. V tomto smyslu byl letošní *STAKAN vyzvednut* určitě hodně vysoko. Těšme se na další. . .

Adresa: Ústav biofyziky a informatiky, I. lékařská fakulta, Univerzita Karlova

GRAFY A TABULKY VE STATISTICE (ANEBO NA CO VE VÝUCE OBVYKLE NENÍ ČAS)

Josef Tvrdík

Katedra informatiky, Přírodovědecká fakulta Ostravské university

Abstrakt: V článku jsou uvedeny některé jednoduché zásady a doporučení pro vhodnou prezentaci statistických výsledků, zejména tabulek a grafů. Tyto zásady a doporučení vycházejí z literatury a ze zkušeností z aplikací statistiky v různých oborech. Některé chyby v prezentaci výsledků jsou podrobně diskutovány a je také doporučeno vhodnější řešení.

Klíčová slova: Aplikace statistiky, prezentace výsledků, tabulky, statistické grafy.

Abstract:

Graphs and Tables in Statistics

Josef Tvrdík

Department of Computer Science, University of Ostrava, CZ

This paper deals with some simple rules for proper presentation of statistical results, mainly for graphs and tables. These rules are based on the recommendations published in literature and on the experience from application of statistical data analysis in various fields of research. Some mistakes in presentation are discussed in detail and better form of presentation is recommended.

Key words:

Applied statistics, presentation of results, tables, statistical graphs.

Úvod

Zkušenosti s pracemi studentů, články a zprávami výzkumníků různých oborů i s některými publikacemi ukazují, že prezentace výsledků statistických analýz je často zbytečně nepřehledná a obtížně čitelná. Mnohdy jako by autoři odmítali užívat zdravý rozum. Při prezentaci výsledků vůbec nemyslí na čtenáře a na to, aby jejich výsledky mohl někdo pokud možno snadno a správně vnímat a jejich badatelské úsilí také docenit. Právě zdravý rozum a ohled na čtenáře jsou důležité zásady, které by měly pomoci k přehledné prezentaci

výsledků, ale zřejmě takto vágní a různými lidmi různě vnímané doporučení nestačí. Přesnější formulace univerzálních zásad přehledné prezentace výsledků však není patrně možná.

Řadu dílčích doporučení, která mohou být užitečná, uvádí van Belle v knize [5], v níž je zásadám správné prezentace statistických výsledků věnována kapitola 7, nazvaná „Words, Tables, and Graphs“. Jako ilustrace toho, že na formě prezentace výsledků záleží, je na začátku této kapitoly uveden následující příklad tří způsobů prezentace téhož jednoduchého výsledku:

- The blood type in the population of the United States is approximately 40%, 11%, 4% and 45% A, B, AB, and O, respectively.
- The blood type in the population of the United States is approximately 40% A, 11% B, 4% AB and 45% O.

The blood type in the population of the United States is approximately,

O	45%
A	40%
B	11%
AB	4%

Rozdíly ve snadnosti či obtížnosti vnímání tohoto jednoduchého výsledku nepotřebují žádné další vysvětlování a snad jsou dostatečným argumentem pro to, že na způsobu prezentace výsledků záleží a že bychom se nad tím měli důkladně zamýšlet. V následujících třech odstavcích jsou uvedeny příklady a doporučení pro prezentaci tabulek a grafů a také pro interpretaci výsledků podle [5].

Prezentace tabulek

Některé chyby ukazuje tabulka 1, ve které jsou uvedeny počty pracovníků v různých zdravotnických profesích v USA roku 1980, názvy kategorií jsou ponechány v angličtině. Tabulka je nedokonalá nejméně ve dvou ohledech:

- Číselné údaje jsou téměř jistě zatíženy různou nepřesností. Zatímco u lékařů, sester, dentistů a optiků to jsou hodnoty získané z příslušných registrů, u některých jiných kategorií jako řečových, fyzických a pracovních terapeutů nebo pedikérů (podiatrists) jde jen o odhad v tisících, údaje v tabulce však vyvolávají dojem, že všechna čísla jsou přesná,
- van Belle jako chybu uvádí i to, že řádky tabulky jsou seřazeny podle abecedního pořadí názvů profesí, ne podle číselných hodnot. Možná se nám tato výhrada zdá neoprávněná, jsme asi zkaženi návyky jak z místních publikací, tak i většinou statistického softwaru, kde je četnostní

Tabulka 1: Počet aktivních zdravotníků v USA v roce 1980 (ze zprávy National Center for Health Statistics, 2000)

Occupation	1980
Chiropractors	25 600
Dentists	121 240
Nutritionists/Dieticians	32 000
Nurses, registered	1 272 900
Occupational Therapists	25 000
Optometrists	22 330
Pharmacists	142 780
Physical Therapists	50 000
Physicians	427 122
Podiatrists	7 000
Speech Therapists	50 000

Tabulka 2: Údaje z tabulky 1 seřazené podle počtu, zaokrouhleno na tisíce.

Occupation in 1000's	1980
Nurses, registered	1 273
Physicians	427
Pharmacists	143
Dentists	121
Physical Therapists	50
Speech Therapists	50
Nutritionists/Dieticians	32
Chiropractors	26
Occupational Therapists	25
Optometrists	22
Podiatrists	7

tabulka seřazena podle názvů kategorií nebo jejich číselných kódů. Ale argument, že pořadí řádků by nemělo záviset na tom, v jakém jazyku publikujeme, nelze jen tak vyvrátit.

Podle [5] by tabulka měla mít formu uvedenou v tabulce 2, tj. číselné údaje zaokrouhlené na tisíce a řádky seřazeny sestupně podle číselných hodnot.

Van Belle dále doporučuje užívat rozumný počet významných číslic a zejména v tabulkách brát ohled na tzv. „efektivní číslice“. To jsou ty číslice, jejichž hodnoty nejsou konstantní, ale mění se. Např. šestimístná čísla 354 691, 357 234, 356 991 mají jen čtyři efektivní číslice. Pokud bychom je chtěli prezentovat přijatelněji, pak bychom měli odečíst od těchto hodnot 350 000 a uvádět tento výsledný rozdíl. V tabulkách ovšem mají být pokud možno nejvýše dvě efektivní číslice, neboť tři a více efektivních číslic člověk obtížně vnímá.

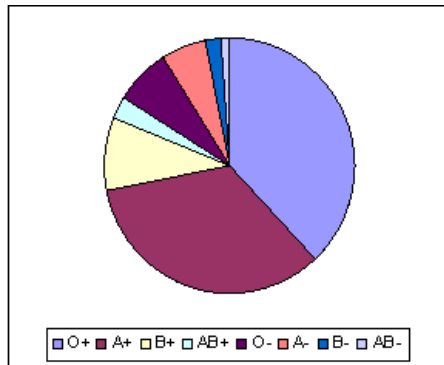
Grafy nebo tabulky?

Všeobecně hlášaná zásada, že grafy místo číselných údajů jsou lepší, není vždy správná. Někdy je tabulka vhodnější než graf, zejména když zvolený typ grafu neodpovídá struktuře dat. Jedním z doporučení [5] je **neužívat výsečové grafy**. Van Belle odkazuje citát: „*Jediná věc je horší než výsečový graf – několik nebo dokonce mnoho výsečových grafů*“.

Výsečové (koláčové) grafy ignorují strukturu dat, čtenář si musí propojovat legendu s výsečemi. Další van Bellův argument proti výsečovým grafům působí na první pohled úsměvně – při tisku výsečových grafů se spotřebuje moc inkoustu. Ale pokud se nad tím zamyslíme, je oprávněný. Porovnáme-li spotřebu inkoustu na bodový graf závislosti hodnot dvou veličin, kdy při malé spotřebě inkoust získáme náhled na tuto závislost se spotřebou na výsečové grafy, kdy při velké spotřebě nezískáme nic (viz příklad, obr. 1), pak závažnost argumentu musíme uznat.

Z výsečového grafu na obr. 1 se opravdu mnoho nedozvíme, struktura grafu neodpovídá struktuře dat, propojování legendy a výsečí je zbytečně namáhavé a spotřeba inkoustu velká. Tabulka 3 prezentuje stejný výsledek daleko přehledněji a srozumitelněji.

V diskusi na konferenci Stakan připomněl Jaromír Běláček, že u číselných údajů v procentech by znak % měl být uveden za každou číselnou hodnotou. Domnívám se, že tato zásada nemá univerzální platnost a autor by měl vždy zvážit, kdy znak % srozumitelnosti pomáhá a kdy je nadbytečný. Ale procenta, ať se znakem % či bez, jsou čitelnější než relativní četnosti jako desetinná čísla z intervalu [0, 1]. U takto uváděných relativních četností při-



Obrázek 1: Relativní četnosti (v %) krevních skupin a Rh faktoru v populaci USA

Tabulka 3: Relativní četnosti (v %) krevních skupin a Rh faktoru v populaci USA

Blood Type	Rh+	Rh-	Total
O	38	7	45
A	34	6	40
B	9	2	11
AB	3	1	4
Total	84	16	100

Tabulka 4: Relativní četnosti krevních skupin a Rh faktoru v populaci USA

Blood Type	Rh+	Rh-	Total
O	0,38	0,07	0,45
A	0,34	0,06	0,40
B	0,09	0,02	0,11
AB	0,03	0,01	0,04
Total	0,84	0,16	1,00

bude neefektivní číslice 0, neefektivní je i znak desetinné tečky nebo čárky a vytratí se grafická informace o počtu efektivních číslic ve sloupcích a řádcích, čímž se zhorší přehlednost a srozumitelnost, jak se můžeme přesvědčit porovnáním tabulek 3 a 4.

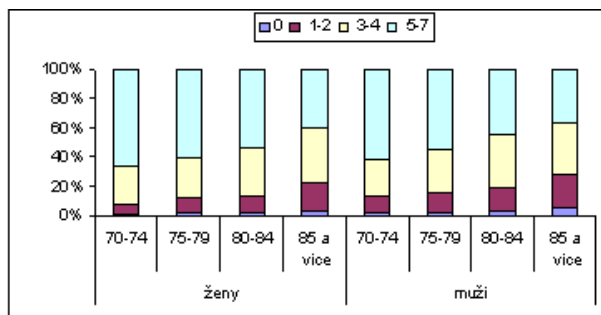
Další zásada podle [5] je **neužívat skládané sloupcové grafy**. Skládané (kumulované, stackbar) sloupcové grafy jsou hůře čitelné než jednoduché sloupcové grafy a často lze najít efektivnější možnost, jak nahlédnout do struktury dat. Van Belle to ilustruje na následujícím příkladu.

Souhrnná zdrojová data z průzkumu počtu aktivit provozovaných seniory v průběhu dvou týdnů jsou uvedena v tabulce 5. Ve zprávě Státního centra pro zdravotní statistiku byly tyto údaje prezentovány formou skládaného sloupcového grafu (obr. 2), což ke vnímání jejich obsahu nijak nepřispělo, spíše naopak. Prezentace by měla usnadňovat odpovědi na následující jednoduché a přirozené otázky: Mají více aktivit muži nebo ženy? Jak se mění počet aktivit s věkem? Liší se tyto změny u mužů a žen? To ovšem spojovaný sloupcový graf rozhodně neusnadňuje.

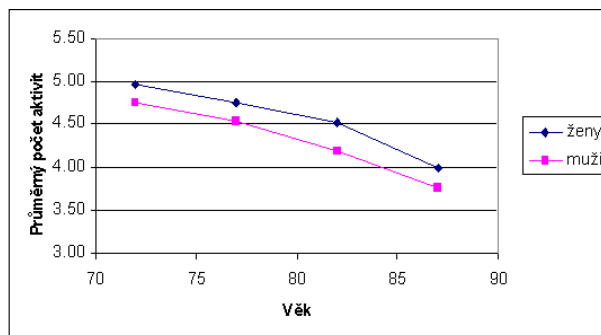
Tabulka 5: Počet aktivit seniorů v průběhu dvou týdnů – četnosti v %

	Počet aktivit	70-74	75-79	80-84	85 a více
Ženy	0	1	1,3	2,1	3,1
	1-2	6,8	10,5	11,9	19,2
	3-4	26,8	27,5	32,5	38,3
	5-7	65,4	60,7	53,5	39,4
Muži	0	1,9	1,7	2,9	5,3
	1-2	10,5	13,3	15,9	23
	3-4	26,3	30,3	36,7	35,9
	5-7	61,2	54,7	44,5	35,9

Přitom docela jednoduchý přepočet a grafické zobrazení průměrných hodnot aktivit pro muže a ženy podle věkových kategorií (obr. 3) vypovídá, že ženy jsou o trochu aktivnější, počet aktivit s věkem klesá a rychlost tohoto poklesu je u obou pohlaví zhruba stejná.



Obrázek 2: Počet aktivit v průběhu dvou týdnů – četnosti v % (Kramarov et al., zpráva National Center for Health Statistics, 1999).



Obrázek 3: Průměrný počet aktivit podle věku a pohlaví

Opatrně při interpretaci výsledků statistické analýzy

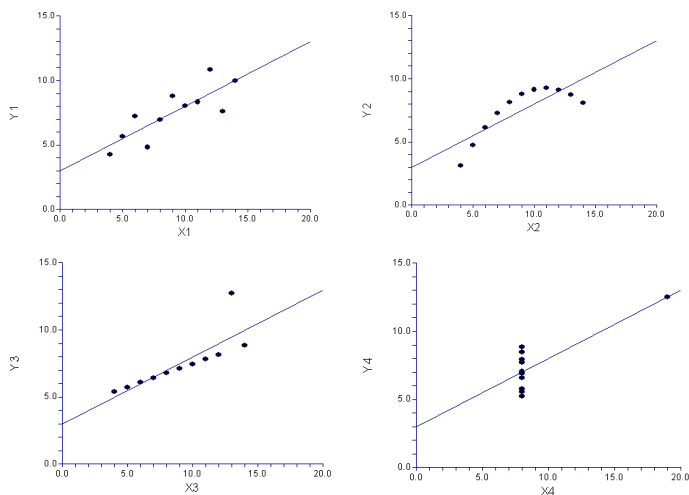
To je další doporučení, kterému se van Belle věnuje dosti podrobně. Neměli bychom přeceňovat (zdánlivě) podrobné numerické výstupy a při interpretaci se držet dobré zásady „Důvěřuj, ale prověřuj“. Ošidnost interpretace „věrohodného“ numerického výstupu je ilustrována známým příkladem (Anscombe [1]), který je užít i jako ilustrace na titulní straně knihy [5]. Číselné výsledky analýzy závislosti dvou veličin z tohoto příkladu jsou uvedeny dále.

Na obr. 4 jsou grafy čtyř závislostí veličiny Y na x . Patří výsledky z tabulek 6 a 7 k některé závislosti z tohoto obrázku? Jen k té první? Možná vás překvapí správná odpověď, že ke všem čtyřem. Tento příklad je dosti populární, data a různé komentáře lze najít i na webu, např. [2, 3]. Zde ke správné interpretaci výsledků pomůže jednoduchý graf. Ale ne vždy lze nesprávnou interpretaci podrobných numerických výsledků odhalit tak snadno, viz např. nějaký složitější zobecněný lineární model ap.

Jak vyjít vstříc čtenáři?

Na to nelze dát všeobecné doporučení, ale některé možnosti naznačí následující příklad. Je to výsek výsledků z článku [4], kde jsou porovnávány stochastické algoritmy pro odhady parametrů nelineárních regresních modelů. Jednou ze sledovaných veličin charakterizující časovou náročnost algoritmů je počet vyhodnocení účelové funkce, tj. součtu residuálních čtverců, potřebných k dosažení podmínky konvergence. V tabulce 8 jsou výsledky porovnání tří algoritmů ve formě, která se v prezentacích popisných statistik užívá často zcela automaticky, jako by to bylo dáno normou. Pro každý algoritmus a pro každou z osmi testovaných úloh je uveden vždy průměrný počet vyhodnocení funkce (n_e) a směrodatná odchylka (s) počtu vyhodnocení.

Jelikož jde o porovnání časové náročnosti algoritmů (a variability této časové náročnosti) na několika úlohách, můžeme porovnávání čtenáři usnadnit, viz tabulka 9. Především můžeme místo směrodatné odchylky užít koeficient variace vyjádřený v procentech průměrné hodnoty (sloupce vc). Tím oproti tabulce 8 podstatně snížíme počet efektivních čísel charakterizujících variabilitu. Dále je možné průměrnou časovou náročnost druhého a třetího algoritmu uvádět jako relativní změnu (v procentech) vzhledem k algoritmu 1. Potom budou záporné hodnoty znamenat zrychlení proti algoritmu 1, kladné hodnoty naopak větší časovou náročnost. Navíc bezprostředně vidíme, zda je tato změna časové náročnosti věcně významná. Pokud by nás zajímala i statistická významnost, lze ji z údajů v tabulce vyhodnotit. Tabulky 8 a 9 jsou



Obrázek 4: Závislosti Y na x k příkladu z Anscombe [1]

Tabulka 6: Popisné statistiky veličin

Proměnná	n	Průměr	Sm. odch.
x	11	9,00	3,32
Y	11	7,50	2,03

Tabulka 7: Výsledky lineární regrese – závislost Y na x

	$b(i)$	$s(b_i)$	T	p
Absolutní člen	3,00	1,125	2,67	0,026
x	0,50	0,118	4,24	0,002
Rovnice regresní přímky:	$Y = 3,00 + 0,50x$			
Koeficient determinace (R^2):	0,67			
Korelační koeficient:	0,82			
Směrodatná odchylka reziduí:	1,24			

Tabulka 8: Porovnání časové náročnosti tří algoritmů na osmi úlohách – průměry a směrodatné odchylky.

Úloha	Algoritmus1		Algoritmus2		Algoritmus3	
	ne	s	ne	s	ne	s
bennett5	41335	14054	47122	16493	36788	14347
boxbod	1308	92	1478	118	824	74
eckerle4	2629	105	2944	147	1709	103
mgh09	10422	834	11777	1060	8859	974
mgh10	20761	2076	24290	972	20969	1677
rat42	2942	177	3354	168	1912	115
rat43	4807	192	5432	217	2932	147
thurber	13915	417	15446	309	9741	292

co do informačního obsahu shodné, jednu z druhé lze snadno přepočítat, ale tabulka 9 je pro porovnávání algoritmů výrazně pohodlnější.

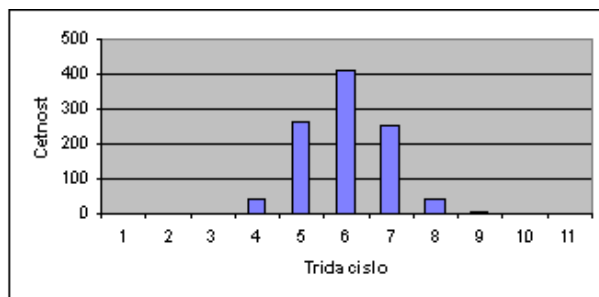
Tabulka 9: Porovnání časové náročnosti tří algoritmů na osmi úlohách – koeficient variace v procentech.

Úloha	Algoritmus1		Algoritmus2		Algoritmus3	
	ne	vc	ne	vc	ne	vc
bennett5	41335	34	14	35	-11	39
boxbod	1308	7	13	8	-37	9
eckerle4	2629	4	12	5	-35	6
mgh09	10422	8	13	9	-15	11
mgh10	20761	10	17	4	1	8
rat42	2942	6	14	5	-35	6
rat43	4807	4	13	4	-39	5
thurber	13915	3	11	2	-30	3

Příklady chyb v prezentaci výsledků

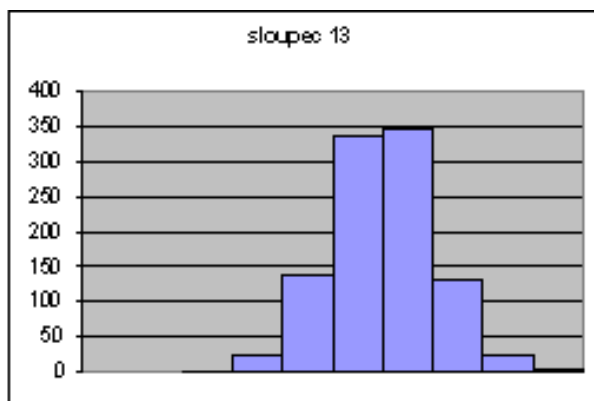
V tomto odstavci jsou komentovány chyby z korespondenčních úloh studentů v kombinované formě bakalářského studia aplikované informatiky v předmětu Analýza dat, který mají studenti ve třetím ročníku studia na Přírodovědecké

fakultě Ostravské univerzity. Ukázky jsou z letního semestru roku 2007, komentáře jsou psány kurzívou.



Obrázek 5: Histogram – častá chyba z naprosté nedbalosti

Histogram na obr. 5 je prezentován tak, jak ho nabízí Excel, zdravý rozum si vybral dovolenou, ohled na čtenáře žádný. Ponechány mezery mezi sloupci, nevhodně zvolené měřítko vodorovné osy (pět tříd s nulovou četností), nic nevyovídající popis vodorovné osy. Šedé pozadí, chybí čeština, bitmapa.



Obrázek 6: Histogram – další častá chyba způsobená nedbalostí

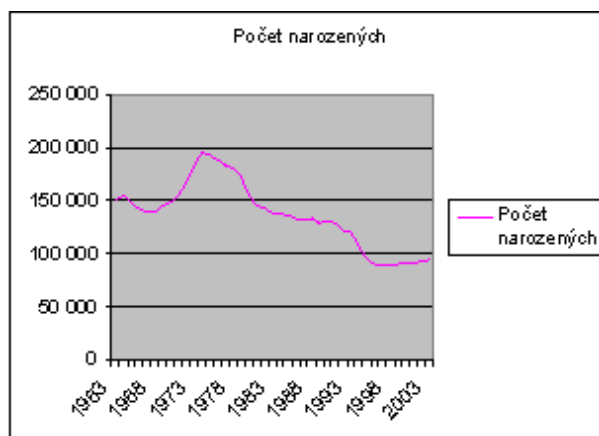
V histogramu na obr. 6 chybí popis os, zbytečný je nic neříkající nadpis histogramu, opět nevhodně zvolené měřítko vodorovné osy. Šedé pozadí, bitmapa.

$H_0: \mu = 6$, průměr $x = 5,959409417$, $s = 0,99046792$
hodnota testového kritéria: $-1,29593994$

Typická ukázka nesprávného a nepřehledného prezentování číselných výsledků s nadbytečným počtem platných číslic.

$b_1 = 0,90711042$, $b_0 = 17,0189542$
 $Se = \sum(Y_i - b_0 - b_1x_i)^2 = 423,839904$, $s^2 = Se/(n - 2) = 26,489994$

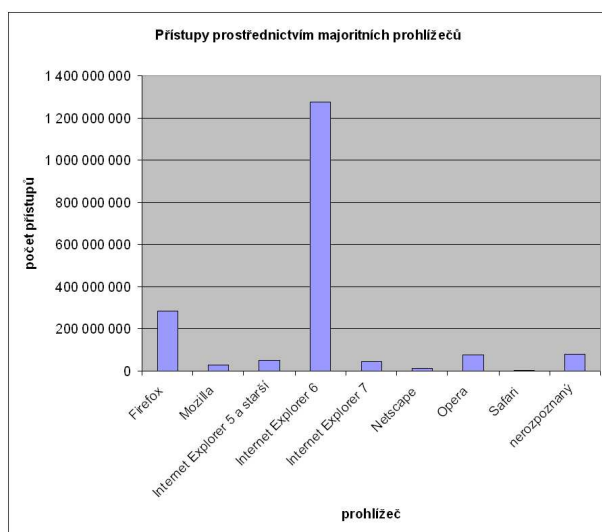
Podobné chyby jako v předchozí ukázce, tady navíc i neobratný a nepřesný zápis symbolů a vzorců. Bez sazby indexů.



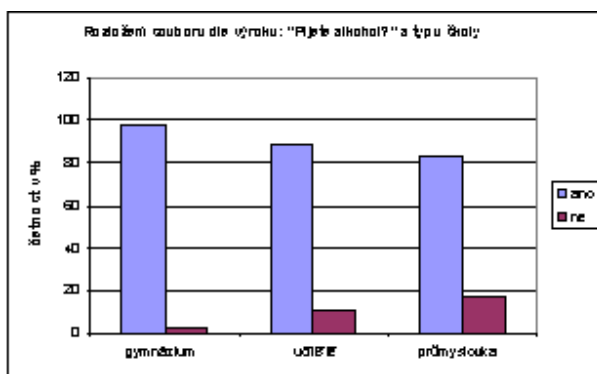
Obrázek 7: Časový průběh počtu narozených

Na obr. 7 chybí popis os grafu, nevhodné jednotky na svislé ose (tři neefektivní nuly, počet narozených měl být v tisících), legenda je nadbytečná a zbytečně zabírá značnou část kreslicí plochy, význam čáry nejasný (bylo užito nějaké vyhlazování?), časová řada by měla být nakreslena jako body, případně se spojnicemi.

Na obr. 8 jsou užity nevhodné jednotky na svislé ose sloupcového grafu (8 neefektivních číslic), vhodnější by bylo uvádět počet přístupů v milionech nebo lépe ve stovkách milionů. Zobrazení devíti značně odlišných četností formou sloupcového grafu není nejvhodnější způsob prezentace tohoto výsledku, tabulka by vypovídala o struktuře a obsahu dat lépe. Osa y by měla mít maximum v hodnotě 100, ne 120.

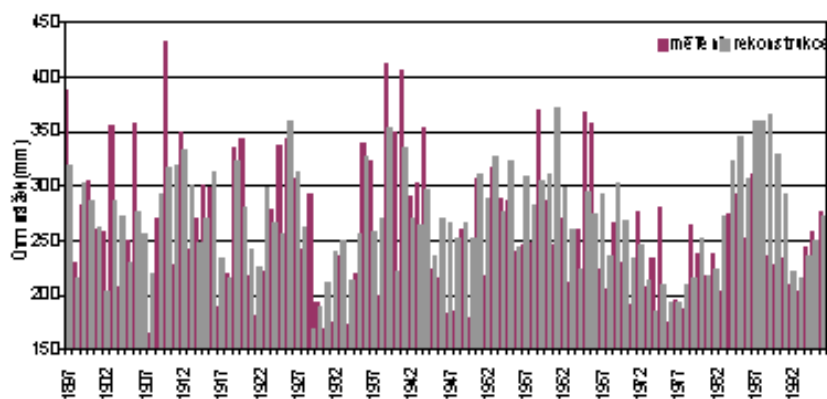


Obrázek 8: Nevhodný sloupcový graf



Obrázek 9: Další nesprávný sloupcový graf

Na první pohled (pomineme-li neobratnou formulaci nadpisu) sloupcový graf na obr. 9 vypadá uspokojivě. Ale jaký je význam druhých sloupečků? Jsou to doplňky do 100%, takže jsou nadbytečné stejně jako legenda. Tři zjištěné relativní četnosti stačilo uvést jako tabulku, zabralo by to méně místa a vypovídalo jasně. Nadpis a popisky os jsou díky bitmapě téměř nečitelné.



Obrázek 10: Nevhodně užitý typ grafu

Na obr. 10 je nevhodně zvolený typ grafu pro zobrazení dvou časových řad do jednoho obrázku, takže výsledek je nepoužitelný pro naprostou nečitelnost. Pro takové závislosti jsou vhodné bodové grafy, případně se spojnicemi bodů.

Závěr

Van Belle [5] požaduje, aby se v prezentaci výsledků statistických analýz věda spojovala s uměním. Možná je to požadavek až příliš náročný, ale rozhodně bychom měli dbát alespoň na dobrou řemeslnou úroveň, využívat základní prezentační dovednosti, při prezentaci výsledků statistických analýz užívat zdravý rozum, přihlížet k možnostem vnímání čtenáře, mít ke čtenáři respekt a snažit se o co největší přehlednost a srozumitelnost výsledků. Otázkou je, jak to učit. Existuje nějaká metodika takové výuky? Může být taková metodika účinná? Lze dobré „mravy“ prezentace statistických výsledků někoho naučit ve studijním kurzu? Není pro dobrou prezentaci statistických výsledků nejdůležitější poctivá snaha a praxí a omyly získané zkušenosti? Tyto otázky bohužel zůstávají dosud nezodpovězeny.

Literatura

- [1] Anscombe F. J. Graphs in statistical analysis. *The American Statistician* 27, 17-21, (1973).
- [2] Diggle P. J., www.maths.lancs.ac.uk/~diggle/gsse401/anscombe.dat, last modified October 30, 2001.
- [3] Scatterplots, Anscombe's datasets, <http://exploringdata.cqu.edu.au/anscomb1.htm>.
- [4] Tvrđík J., Křivý I., Mišík L., Adaptive population-based search: Application to estimation of nonlinear regression parameters, *Computational Statistics & Data Analysis*. 52(2), 713-724, 2007. Available online 9 November 2006, <http://www.sciencedirect.com/science/article/B6V8V-4M9HWXT-3/2/e4fa1077aa80c154b130396b3c486286>.
- [5] van Belle G.: *Statistical Rules of Thumb*, John Wiley & Sons, 2002

Poděkování: Tato práce byla zčásti podporována grantem 201/05/0284 Grantové agentury České republiky.

ROBUST AND NONPARAMETRIC STATISTICAL INFERENCE 1. – 6. ZÁŘÍ 2007, HEJNICE

Marek Omelka, Jan Kalina

Centrum Jaroslava Hájka pro teoretickou a aplikovanou statistiku pořádalo ve dnech 1. – 6. 9. 2007 workshop Robust and Nonparametric Statistical Inference. V příjemném prostředí Mezinárodního centra pro duchovní obnovu v Hejnicích se sešlo 63 vědců, aby se podělilo o poznatky v oblasti statistické inference zejména v oblasti robustních a neparametrických statistických metod. Mnozí z těchto účastníků z celé Evropy i USA patří mezi špičkové odborníky v oblasti metod matematické statistiky.

Organizačně byl workshop zajištěn pracovištěm Centra J. Hájka na Technické univerzitě v Liberci a na MFF UK v Praze. Celá akce navázala na

tradici předchozích workshopů Perspectives in Modern Statistical Inference I-III, které se konaly v letech 1998, 2002, 2005 ve spolupráci s dalším partnerským pracovištěm Centra J. Hájka, Institutem matematiky a statistiky Masarykovy univerzity v Brně. Z velmi inspirativních přednášek bylo tak jako v minulých letech patrné, že matematická statistika je vskutku živý obor, který reaguje na stále živější poptávku po kvalitním a zároveň rychlém zpracování dat. Více než kdykoli v minulosti se dnes pracuje s velkými datovými soubory s komplikovanou strukturou, což přináší stále nové a nové výzvy pro metodologickou i výpočetní oblast statistiky.

Mezi mnoha jinými zmiňme například příspěvek prof. Sena (USA) s aplikacemi v genetické oblasti, dále sugestivní příspěvek prof. Stuteho (SRN) inspirovaného zpracováním dat z komplexních marketingových průzkumů, či příspěvek prof. Pavlopoulose (Řecko) zabývající se statistickým modelováním dešťových srážek. Z čistě metodologických příspěvků jistě zaujala přednáška prof. Kagana (USA), která se dotýkala některých fundamentálních výsledků statistické teorie, dále neotřelý přístup prof. Daviese (SRN) k neparametrické regresi či nový přístup k odhadům založeným na pořadích prezentovaný prof. Pandaveinem (Belgie). Dopolední přednášky zvaných řečníků pak byly odpoledne doplněny kratšími příspěvky a nedělní večer byl věnován také posterům, u kterých se nezvykle živě diskutovalo. Přednášky zvaných řečníků budou publikovány ve zvláštním čísle časopisu Applications of Mathematics po připravovaném recenzním řízení.

Pořadatelům (a zde patří dík zejména liberecké části Centra Jaroslava Hájka) se podařilo vytvořit velmi přátelskou a neformální atmosféru, která podnítila řadu zajímavých diskusí. K upevnění a rozvinutí spolupráce zejména mezi Belgií a ČR přispěl také již tradiční čtvrtý ročník pražsko-bruselského semináře, kterému byl věnován první den workshopu. Velkým příslibem do budoucna je také účast více než 20 mladých vědců (do 30 let) z Belgie, ČR, Německa, Malt, Polska, Švédska a Turecka, kteří byli povzbuzeni a inspirováni setkáním s autory věhlasných monografií a článků.

Adresa: Marek Omelka, Jan Kalina,
Centrum Jaroslava Hájka pro teoretickou a aplikovanou statistiku

THIRD MEETING OF NATIONAL STATISTICAL SOCIETIES

Gejza Dohnal

V pořadí již třetí setkání předsedů národních statistických společností se uskutečnilo v letošním roce ve slovinském hlavním městě Ljubljani. Zástupce dalších pěti statistických společností sem pozval předseda Slovinské statistické společnosti Andrej Blejec. Do Ljubljany přijeli zástupci Maďarské statistické společnosti (předseda Lőrinc Soós a generální tajemnice Éva Laczká, oba z Maďarského statistického úřadu), Rakouské statistické společnosti (předseda Wilfried Grossmann, Universität Wien), Rumunské statistické společnosti (předseda Constantin Mitrut, Institute of Statistical Mathematics and Applied Mathematics a tajemník společnosti, Ioan Goreac), Slovenskej štatistickej a demografickej spoločnosti (předseda Peter Mach) a za naši společnost jsem se nakonec zúčastnil pouze já.

Jednání probíhalo v budově Slovinského statistického úřadu, kde nás přivítala jeho ředitelka, paní Irena Križman, informací o historii a organizaci úřadu. Poté informovali předsedové o činnosti jednotlivých národních společností a pořádaných akcích. Jménem naší společnosti jsem tlumočil návrh na uspořádání příštího setkání v roce 2008 v Praze. Po mém návrhu, aby toto setkání bylo spojeno s dvoudenní konferencí pro studenty statistiky, se zahájila široká diskuze. Návrh byl nakonec jednomyslně přijat. V příštím roce se tedy uskuteční kromě setkání předsedů národních statistických společností souběžně i konference, na které budou mít příležitost studenti magisterského a postgraduálního studia přednést své výsledky v angličtině před mezinárodním publikem. Vzhledem k tomu, že se tato akce bude konat v prvním týdnu měsíce září, tedy po prázdninách, bude třeba účastníky konference přihlásit ještě před koncem letního semestru.

Setkání pokračovalo odpolední prohlídkou města, v němž jsou velmi výrazné stopy slovinského architekta Josipa Plečnika, jehož dílo známe důvěrně i zde v Praze (rekonstrukce prvního a třetího nádvoří Pražského hradu, kostel Nejsvětějšího Srdce Páně v Praze na Vinohradech). V centru města, které má přibližně 250 tisíc obyvatel, jsou velmi příjemná místa okolo řeky Ljubljanica. Ljubljana je univerzitním městem, což je znát i v ulicích, kde jsme potkávali převážně mladé lidi. Poněkud rozpačitě vypadá dostavba Ljubljanského hradu, kde bylo částečně dostavěno původní gotické zdivo a doplněno betonovými a ocelovými prvky, které po čase vypadají dosti sešle (na rozdíl od původního kamene). Zato pohled na nedaleké zasněžené vrcholky Savinjských Alp působí velmi majestátně.

PRASTAN 2007

Gejza Dohnal

Ve dnech 4.–8. 6. 2007 se konal v Banské Bystrici další ročník konference Slovenskej štatistickej a demografickej spoločnosti PRASTAN'07. V letošním roce se na její organizaci nepodílela naše společnost, neboť my jsme organizovali týden před tím náš STAKAN'07 (viz informace ve 2. letošním čísle IB). Nicméně i tato konference, podobně jako STAKAN'07, byla hezkou ukázkou česko-slovenské vzájemnosti/vzájomnosti. Ač zahraniční, přesto naše!

Konference probíhala v přednáškovém sále právnické fakulty Univerzity Mateja Belu, která je umístěna ve stráni, s výhledem na okolní hory a město pod sebou. Ve stejné budově, pouze o pár poschodí výše, bylo zajištěno i příjmné ubytování (jen ta strava v závodní jídelně byla poněkud chudší).

Program byl rozdělen do několika sekcí:

- Numerika I a II (7 příspěvků)
- Statistika I a II (8 příspěvků)
- Statistické řízení jakosti I a II (7 příspěvků)
- Pedagogika (3 příspěvky)
- Kvantové logiky (4 příspěvky)
- Fuzzy množiny (3 příspěvky)

Přednášky probíhaly ve slovenštině a v češtině, což je na mezinárodní konferenci příjemné. Zazněla zde i jedna plenární přednáška doc. Janigy o „Normalizácii metod štatistického riadenia kvality“. Sborník příspěvků vyšel jako zvláštní číslo 2/2007 časopisu Forum Statisticum Slovaca. Řada příspěvků byla na velmi dobré úrovni a je jen dalším projevem ignorance našich „vědeckých“ úředníků, když tento časopis zařadili pro letošní rok mezi časopisy „s vědeckou hodnotou nula“!

Kulturní program zahrnoval – vedle návštěvy několika příjemných „zahradek“ na banskobystričském náměstí – i večerní návštěvu Štátnej opery v Banskej Bystrici, kde bylo představení Straussovy opery Netopýr. Ve čtvrtek (7. 6.) odpoledne jsme vyrazili „do hor“. Na programu byla návštěva Španieľ doliny, malebného místa uprostřed hor, na rozhraní Nízkyh Tatier a Velkej Fatry. Zde se už ve 13. století těžilo zlato a stříbro a kopce v okolí vesnice jsou provrtány tisícem chodeb, jako obrovská mraveniště. Těžba zde ustala až v roce 1888, kdy se zdejší ložiska zdála být zcela vyčerpána.



Výhodný nákup starších knih

Milí kolegové, chystá se likvidace dole uvedených knih. Myslím, že by měly doputovat těm, kteří by je upotřebili. Zvažte tento lukrativní nákup pro Vaše pracoviště. Srdečně zdraví Zdeněk Půlpán.

- K problematice hledání podstatného v humanitních vědách PŮLPÁN ZDENĚK, zásoby 118 kusů, rok vydání 2001 – nabízená cena 1,- bez DPH.
- K problematice měření v humanitních vědách PŮLPÁN ZDENĚK, zásoby 107 kusů, rok vydání 2000 – nabízená cena 1,- bez DPH.
- K problematice vágnosti v humanitních vědách PŮLPÁN ZDENĚK, zásoby 194 kusů, rok vydání 1997 – nabízená cena 48,- bez DPH.
- K problematice zpracování empirických šetření v humanitních vědách PŮLPÁN ZDENĚK, zásoby 156 kusů, rok vydání 2004 – nabízená cena 1,- bez DPH.

Marika Tomanová, vedoucí odbytu, Nakladatelství Academia, Vodičkova 40, 110 00 Praha 1, tel.: +420 221 411 478, +420 774 135 605, fax: + 420 224 941 982, ICQ 385-760-471.

tomanova@academia.cas.cz

NOVÁ KNIHA: VIACROZMERNÉ ŠTATISTICKÉ METÓDY S APLIKÁCIAMI

Iveta Stankovičová, Mária Vojtková

Milí priatelia,

dovoľujeme si Vás informovať, že vo vydavateľstve IURA EDITION vyšla naše nová kniha (členiek výboru SDŠS): *Viacrozmerné štatistické metódy s aplikáciami*.

Obsah knihy:

1. kapitola: Metódy viackriteriálneho hodnotenia
2. kapitola: Metóda hlavných komponentov
3. kapitola: Faktorová analýza
4. kapitola: Kanonická korelačná analýza
5. kapitola: Zhuková analýza
6. kapitola: Diskriminačná analýza
7. kapitola: Logistická regresia
8. kapitola: SAS Enterprise Guide

Kniha obsahuje aj DVD s použitými dátovými súborami v riešených príkladoch knihy a tiež výukový program SAS Enterprise Guide verzia 4. Vo výukovom programe SAS EG sa dajú spracovávať dátové súbory do rozsahu 1000 štatistických jednotiek. Inak program nemá žiadne iné obmedzenia v porovnaní so štandardným programom SAS Enterprise Guide verzia 4.

Knihu je možné objednať alebo priamo zakúpiť vo vydavateľstve:

IURA EDITION, spol. s r. o.

Oravská 17

821 09 Bratislava

Predajné hodiny v budove vydavateľstva:

Pondelok-Piatok 09-12 hod. a 13-15 hod.

<http://www.iura.sk/ekonlit.asp>

RECENZE SKRIPT PETRA KLÍMKA EKONOMETRIE: CVIČEBNICE

Recenzent: Pavel Stříž

Cvičebnice je po úvodním slově autora rozdělena do tří velkých částí:

- I. Ekonometrické modely.
- II. Ekonometrie – řešené příklady pro cvičení.
- III. Ekonometrie – neřešené příklady pro cvičení.

Ad I. Ekonometrické modely. Nalezneme zde šest komplexních příkladů:

1. Klasický a neoklasický model Phillipsovy křivky v České republice.
2. Vývoj HDP s jednou a dvěma zvažovanými proměnnými.
3. Zkoumání tendencí ve spotřebě s vlastním typem modelu, ukázkově řešeným ve statistickém programu Statistica.
4. Řešení použití koeficientů pružnosti v analýze spotřeby.
5. Modely časových řad pomocí dekompozice časové řady, klouzavých průměrů, exponenciálního vyrovnávání a významné Boxovy-Jenkinsonovy metodologie (ARIMA modelování), včetně predikce modelu, vše řešeno také pomocí programu Statistica (upoutávka na verzi 8 viz další strana).
6. Závěrečný komplexní příklad je náhled na vícerozměrný model, přesněji řečeno na shlukovou analýzu v marketingovém výzkumu trhu.

Ad II. Ekonometrie – řešené příklady pro cvičení. Příklady jsou postupně zopakováním a rozšířením předmětů *Metody statistické analýzy* a *Aplikované statistiky*, konkrétně oblastí regresní a korelační analýzy, analýzy časových řad, zjišťování trendu, sezónní, cyklické a náhodné složky u dekompozice časových řad, adaptivních modelů a korelace časových řad.

Další partie jsou zaměřeny na cenovou a bankovní statistiku a mezinárodní srovnávání zemí pomocí parity kupní síly. Parita kupní síly je řešena pomocí Fisherova indexu a postupu dle Eltetého, Kövese a Szulce. Závěr této kapitoly tvoří informace o data miningových technikách a příslušných softwarových výrobcích.

Ad III. Ekonometrie – neřešené příklady pro cvičení. Kapitola obsahuje několik neřešených příkladů na analýzu časových řad, a to se zaměřením na její elementární charakteristiky, analýzu trendu a sezónní složku.

Publikace operuje s reálnými daty, především je čerpáno z oficiálních statistik. Po těchto třech částech následuje do odborných bloků seříděná použitá a doporučená literatura a v přílohách umístěné statistické tabulky.

Cvičebnici lze uplatnit v předmětu *Ekonometrie* na vysokých školách.

„Upgrade, který se zaplatí velmi brzy.“

recenze v *Scientific Computing World* (06/2007)

STATISTICA™

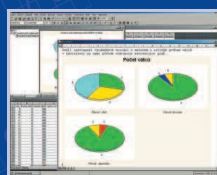
analýza dat business intelligence řízení kvality data mining



Verze 8 přináší novinky, které byste neměli přehlédnout:

Integrace Microsoft Office

Spolupráce softwaru *STATISTICA* s Microsoft Office je nyní tak snadná, jak jen může být. Dokumenty Microsoft Word mohou být použity přímo jako druh výstupního protokolu, podobně dokumenty Microsoft Excel mohou být otevřeny přímo prostřednictvím *STATISTICA* bez nutnosti importu. Grafy lze rovněž umístit do dokumentů Microsoft Wordu nebo Microsoft PowerPointu jednoduchým přetažením.

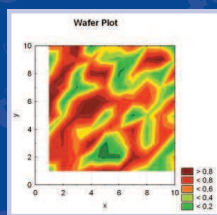


Uložení rozpracovaného projektu

Nyní můžete uložit *STATISTICA* projekt, který obsahuje zmrazený stav vaší právě rozpracované analýzy, včetně rozložení všech oken na obrazovce. Po otevření projektu může příslušná analýza pokračovat z místa, kde byla naposledy uložena. Je též nabídnuta možnost uložit nahrané makro asociované s danou analýzou.

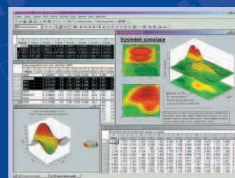
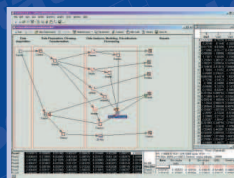
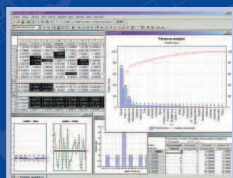
Data mining ještě snadněji

Vytvoření dataminingového projektu je nyní ještě snadnější. Receptu podobný nástroj *STATISTICA Data Miner Recipes* vás krok po kroku provede při vytěžení vašich dat. Speciálně neuronové sítě lze jednoduše využívat díky novému produktu *STATISTICA Automatizované neuronové sítě*. Uživatel může nejen tvořit vlastní modely neuronových sítí, ale i jednodušeji vybírat či vytvářet modely automaticky.



Podnikový systém pro analýzu dat

Analytický software může nabídnout více než samotnou analýzu dat. *STATISTICA Enterprise* tak kromě rozsáhlých analytických modulů přináší v jednom systému přístup k datům v datových skladech, zabezpečený auditovatelný přístup, rozsáhlé možnosti na poli automatizovaného reportování či on-line monitorování dat, snadné a přehledné prostředí pro účinnou integraci systému do pracovních procesů v malých i velkých společnostech a mnoho dalších možností.



OBSAH BULLETINU

Prof. Ing. Jaroslav Jílek, CSc.	1
<i>Jaromír Běláček</i> Stakan 2007 – Dojmy a reflexe	3
<i>Josef Tvrdlík</i> Grafy a tabulky ve statistice	6
<i>Marek Omelka, Jan Kalina</i> Robust and Nonparametric Stat. Inference ..	20
<i>Gejza Dohnal</i> Third Meeting of National Statistical Societies	22
<i>Gejza Dohnal</i> Prastan 2007	23
<i>Zdeněk Půlpán</i> Nákup zlevněných knih	24
<i>Iveta Stankovičová, Mária Vojtková</i> Informace o nové knize	25
<i>Pavel Stříž</i> Recenze skript Petra Klímka Ekonometrie: Cvičebnice	26
Produkt <i>Statistica 8</i> od StatSoft.com	27

Výroční zasedání ČStS v roce 2008

Výbor ČStS si dovoluje pozvat všechny členy společnosti na pravidelné výroční zasedání, které se bude konat ve čtvrtek 31. ledna 2008 od 13:00 hodiny v prostorách staré budovy VŠE v Praze, náměstí W. Churchilla 4. Místnost bude upřesněna a viditelná zpráva bude umístěna při vstupu do budovy.

Na program bude kromě obvyklých zpráv o činnosti a hospodaření také odborná přednáška manželů Budíkových na téma *Modelování ročního chodu UVB záření v Antarktidě*.

Statistické dny v Liberci

V některém z březnových dnů roku 2008 se uskuteční tradiční Statistické dny, tentokrát v Liberci. Bližší informace budou upřesněny a zveřejněny na webové stránce společnosti.

ISSN 1210 – 8022. Informační Bulletin České statistické společnosti vychází čtyřikrát do roka v českém vydání. Příležitostně i mimořádné české a anglické číslo.

Předseda společnosti: Doc. RNDr. Gejza Dohnal, CSc., ÚTM FS ČVUT v Praze, Karlovo nám. 13, 121 35 Praha 2, e-mail: gejza.dohnal@fs.cvut.cz

Redakce: Ing. Pavel Stříž, Ph.D., e-mail: striz@fame.utb.cz

Uzávěrka: českého čísla 1/2008 je 14. února 2008 půlnoc

FTP: exp.uis.fame.utb.cz; uživatel: csts; heslo: csts

WEB server: <http://www.statspol.cz/>