

# Informační Bulletin



č. 3, prosinec 1998

## České Statistické Společnosti

Nebát se a nekrást - ani software.

Josef Tvrdek

Tato obměna známé Masarykovy zásady byla mottem panelové diskuse na semináři Programování v Ostravě zhruba před dvěma roky. Od té doby se naštěstí změnilo mnohé. Obávám se však, že téma uvedené názvem je stále současné. Proto se pokusím k němu učinit pár poznámek. Pro potěšení přátel starých formalismů se přidržím i tehdejší struktury diskuse.

Heslo z nadpisu zapíšeme jako výrokovou formu:

"[Bojíme se] a [krademe software]".

První osoba plurálu mi připadá výstižná. Pokud se najde někdo, koho se tento výrok netýká, předem se mu omlouvám. Výroky v hranatých závorkách označme postupně p,q. Pravdivostní čtyřřádková tabulka výroku  $p \wedge q$  je všeobecně známa, tak se mohou rovnou pustit do poznámek k jednotlivým ohodnocením (řádkům). Pořadí řádků však zvolím tak, jak podle mého soudu odpovídá vývoji v čase. Co jiného než historie nám pamětníkům zbývá.

**10** : Dřevní doby počítačů v naší vlasti. Spíše než bázeň však příčinou nekradení byl nedostatek příležitosti. Většinou nebylo co krást nebo na co krást (nebyly odpovídající počítače). Na strach jsme moc nemuseli myslet, alespoň ne v souvislostech se softwarem.

**01** : "Blažený" stav z doby částečné propustnosti železné opony. Mnoho západních uživatelů by nevěřicně kroutilo hlavou nad obsahem disků mnohých našich péceček. Zdvojení a ztrojení různých programů pro stejné funkce je pojištěno ještě současnou instalací různých verzí téhož programového produktu. Nedostatek manuálů a dalších doprovodných informací k programům nahrazujeme dobrodružstvím individuálního výzkumu. Nejoblíbenější výzkumné metody jsou pokus-omyl a užití geniální intuice. Výzkum je levný, proto-

že jej platí stát, většinou bídě. Šéfové jsou spokojeni, neboť je neobtěžujeme, máme svých starostí dost. Nic to nestojí, protože peníze na mzdy jsou naplánovány. A stejně užitečnější práce pro nás není. Naše práce není bez výsledku: Produktem jsou občas zavírované, ale především "zasoftwarované" počítače, na které se rozumná aplikace nevejde nebo se nedá spolehlivě provozovat.

11 : Žhavá současnost. Železná opona nás už od světa neodděluje, ale ani nechrání. Svět o nás ví, vědí o nás i organizace pro odhalování neoprávněného užití programů. Není příjemné se dočíst, jak byl FAST úspěšný tu i jinde, kolik dolarů musel zaplatit nepoctivý uživatel. A přitom peníze na software nejsou, nedočkavost je veliká. Dlouhodobých pracovních návyků je těžké zbavit se najednou. Tak do bouřlivého běhu událostí přibývá ještě dilema: v bezesných nocích se duševně připravovat na trestní postihy na horní hranici sazby nebo vymýšlet tajné schránky na disky a procvičovat si v duchu přepínače příkazu RESTORE. Prostě současnost žhavá, až pálí.

100 : O výhodách tohoto přístupu nejen k softwaru věděl už náš první prezident. Můžeme se k tomuto přístupu svobodně rozhodnout nebo počkat, až nás okolí (teď už dosti široké) k němu dofackuje. Ale nevidím cestu, jak se mu dlouhodobě vyhnout. Pro ty, ve kterých toto poslední ohodnocení výroku p&q (viz návštějí odstavců) vzbuzuje bezvýhodné představy, uvedu jeden optimistický příklad, který mi vyprávěl před několika léty můj kolega a přítel. Ten má rodinu v Americe a jeho švagr se živí obchodem. Při návštěvě mu můj kolega (podle našich zvyklostí z domova) předhodil, že je škoda utrácet hříšné peníze za software, když přece toto zboží je snadné ukrást a přitom vlastně okradenému nic neubude. Dotyčný Čechoameričan o tomto návrhu uvažoval do druhého dne. Nebyl to tedy žádný pohádkový hrdina, který by odmítnul z etických důvodů. Po zralé úvaze odmítnul z důvodů ekonomických. Usoudil, že by se mu to nevyplatilo. Ztratil by výhody spolehlivosti, aktuální dokumentace, možnosti okamžité telefonické konzultace při potížích, zvýhodněné nabídky nových verzí, pomoci při zavádění atd. Kromě toho při okradání jiných by okradl i sám sebe - o čas, který by musel vynaložit, aby nahradil chybějící služby. Vyspělá společnost je zavedená i na vyspělou dělbu práce. Profesionální úroveň zboží je samozřejmostí. A nejběžnější, nejrychlejší a nejlevnější způsob, jak potřebné zboží získat, je si ho koupit.

Do nového roku přeji nám všem, abychom žili ve společnosti, kde nebát se a nekrást je zcela normální a samozřejmé.

## Zpráva o jedné zahraniční cestě

Jiří Anděl

Letos v září jsem navštívil univerzitu v Bělehradě. Nebudu zde popisovat oficiální program svého pobytu, protože ten neoficiální bývá zajímavější. Při jedné společné večeři se hovořilo i o různých statistických pozoruhodnostech. Jednu velice hezkou úlohu, kterou jsem do té doby neznal, vám popíši.

Nechť  $X_1, X_2, \dots, X_n$  je posloupnost nezávislých stejně rozdělených náhodných veličin, které mají střední hodnotu  $\mu$  a konečný rozptyl  $\sigma^2$ . Z důvodu, který vysvětlím dále, se bude předpokládat, že tyto veličiny jsou nezáporné. Označme

$$S_n = X_1 + X_2 + \dots + X_n$$

jejich součet. K základní abecedě statistické vědy patří poznatek, že

$$\text{var } S_n = n\sigma^2.$$

Mnohý čtenář si možná teď vzpomene, že právě tohle dostal kdysi u zkoušky.

Víme tedy, že rozptyl součtu nezávislých stejně rozdělených náhodných veličin závisí na jejich počtu  $n$  lineárně. A teď přijde ona klíčová otázka. Jak závisí na  $n$  výraz

$$\text{var } \sqrt{S_n} ?$$

Asi by se dalo čekat, že  $\text{var } \sqrt{S_n}$  bude úměrné  $\sqrt{n}$ , nebo snad možná  $n$ . Šokující však je to, že ve skutečnosti  $\text{var } S_n$  je omezeno konstantou. Důkaz je založen na následujícím tvrzení.

Tvrzení o rozptylu odmocniny. Nechť  $S$  je nezáporná náhodná veličina s konečným rozptylem  $\text{var } S$  a s kladnou střední hodnotou  $ES$ . Pak platí

$$\text{var } \sqrt{S} \leq (ES)^{-1} \text{var } S.$$

Pro  $S=S_n$ , dostaneme

$$\text{var } S_n = n\sigma^2, \quad ES_n = n\mu,$$

takže z výše uvedeného tvrzení plyne

$$\text{var } \sqrt{S_n} \leq \mu^{-1}\sigma^2.$$

Důkaz toho pozoruhodného tvrzení o rozptylu odmocniny sice neobsahuje nic složitějšího, ale je nepříjemně dlouhý. Odhaduji, že k jeho přednesení bych potřeboval ve škole asi 45 minut, čili jednu vyučovací hodinu. Pokud se někdo ze čtenářů domnívá, že se bez takového důkazu obejde (a sám si ho vyrábět nehodlá), může si ho prostudovat (protože pouhé přečtení nestačí) v článku Banjevič D., Bratičević D. (1983) : Note on dispersion of  $X$ , Publ. Inst. Math., Nouvelle série 33 (47), 23-28.

## Který si vyberete ?

Josef Machek

Co si máte vybírat a z čeho, z jaké nabídky? Bodový odhad, formuli či metodu pro výpočet bodového odhadu z několika možných. Nabídka je omezena jen tím, že 'nevedeme', nemáme na skladě bayesovské odhady a odhady založené na sofistikovaných kritériích moderní teorie rozhodovacích funkcí jako přípustnost atd. (která stejně zpravidla nevedou k žádným pozitivním výsledkům). Výběr je omezen jen na staromódní solidní zboží. Článek je určen k pobavení a snad i k zamyšlení; příklad by měl trochu otřást naší vírou, že vždycky postupujeme tím nejlepším způsobem. Známe přece maximálně věrohodné odhady a nejlepší nestranné odhady, tak co si ještě vybírat?

Představme si však následující velmi jednoduchou situaci. Nechť  $(X_1, X_2, \dots, X_n)$  jsou navzájem nezávislé náhodné veličiny s hustotou pravděpodobnosti

$$f(x, \lambda) = \begin{cases} \exp(-\lambda) & \text{pro } x > 0 \\ 0 & \text{pro } x < 0 \end{cases} \quad (1)$$

kde  $\lambda$  je neznámý parametr,  $\lambda > 0$ . Jinými slovy  $(X_1, X_2, \dots, X_n)$  je náhodný výběr z exponenciálního rozdělení se střední hodnotou  $1/\lambda$ . V mnohých aplikacích je zajímavý a důležitý právě parametr  $\lambda$ , t.j. převrácená hodnota střední hodnoty. Je-li  $X$  doba života (doba do poruchy) výrobku, nebo součásti, je  $\lambda$  známé jako intenzita poruch, je-li  $X$  doba života člověka po určitém lékařském zákroku, nebo doba života experimentálního zvířete po vystavení určitým podmínkám, je  $\lambda$  tzv. intenzita úmrtnosti. Jsou-li  $X_i$  intervaly mezi výskyty určitých událostí tvořících Poissonův proces, je intenzita tohoto procesu, t je střední hodnota počtu výskytů událostí v intervalu délky  $t$ .

Mějme tedy výsledky pozorování  $X_i = x_i$ ;  $i=1, 2, \dots, n$ . Jak odhadnout "co nejlépe" skutečnou hodnotu parametru  $\lambda$ ?

Sdružená hustota náhodného vektoru  $(X_1, X_2, \dots, X_n)$  je

$$f(x_1, x_2, \dots, x_n; \lambda) = \lambda^n \exp[-\lambda \sum_{i=1}^n x_i] = \lambda^n \exp(-\lambda T), \quad (2)$$

označíme-li

$$T = \sum_{i=1}^n x_i. \quad (3)$$

Vzpomeneme-li si na základy teorie odhadu, připomeneme si, že  $T$  je tzv. postačující statistika pro parametr  $\lambda$ . To znamená, zhruba řečeno, že "dobré odhady" stačí hledat mezi funkcemi statistiky  $T$ .

Snadno lze ukázat, že statistika  $T = \sum_{i=1}^n X_i$  má rozdělení s hustotou

$$g(t; \lambda) = \begin{cases} [(n-1)!]^{-1} \lambda^n t^{n-1} \exp(-\lambda t) & \text{pro } t > 0 \\ 0 & \text{pro } t = 0 \end{cases} \quad (4)$$

Odsud plyne, že momenty statistiky  $T$  jsou

$$E(T^k) = \frac{(n+k-1)!}{\lambda^n (n-1)!}, \quad k = \dots, -2, -1, 0, 1, 2, \dots \quad (5)$$

speciálně - což budeme potřebovat dále -

$$E(T^{-1}) = \lambda / (n-1), \quad (6)$$

$$\text{Var}(T^{-1}) = \lambda^2 / [(n-1)2(n-2)]. \quad (7)$$

Ze vztahu (4) je také vidět, že statistika  $T$  má rozdělení jako náhodná veličina

$$\frac{1}{2\lambda} \chi^2(2n), \quad (8)$$

kde  $\chi^2(2n)$  značí náhodnou veličinu s rozdělením chí-kvadrát o  $n$  stupních volnosti. To znamená, že distribuční funkce statistiky  $T$  je

$$F_T(t) = P(T < t) = P(\chi^2(2n) < 2\lambda t). \quad (9)$$

Z (2) snadno zjistíme, že metoda maximální věrohodnosti přikazuje odhadnout parametr pomocí statistiky

$$\hat{\lambda} = n/T = n/\sum_{i=1}^n X_i = 1/\bar{X}. \quad (10)$$

Tento odhad není nevychýlený (nestranný); jeho střední hodnota je (s užitím (6))

$$E(\hat{\lambda}) = n\lambda / (n-1). \quad (11)$$

Má tedy kladné vychýlení rovné

$$B(\hat{\lambda}) = E(\hat{\lambda}) - \lambda = \lambda / (n-1). \quad (12)$$

Rozptyl odhadu (10) je (s užitím (7))

$$\text{Var}(\hat{\lambda}) = n^2 \text{Var}(T^{-1}) = \lambda^2 n^2 / [(n-1)^2 (n-2)] \quad (13)$$

a tzv. střední kvadratická chyba (MSE, Mean Square Error)

$$C(\hat{\lambda}) = E[(\hat{\lambda} - \lambda)^2] = \text{Var}(\hat{\lambda}) + B^2(\hat{\lambda}) = \lambda^2 (n+2) / [(n-1)(n-2)]. \quad (14)$$

Bylo by velmi jednoduché upravit odhad tak, aby byl nevychýlený; tuto podmínku splňuje

$$\tilde{\lambda} = (n-1)/T. \quad (15)$$

Jest zřejmé

$$E(\tilde{\lambda}) = (n-1)E(T^{-1}) = \lambda, \quad (16)$$

tudíž vychýlení

$$B(\tilde{\lambda}) = 0 \quad (17)$$

a rozptyl (který je - vzhledem k nestrannosti - roven střední kvadratické chybě)

$$\text{Var}(\tilde{\lambda}) = C(\tilde{\lambda}) = (n-1)^2 \text{Var}(T^{-1}) = \lambda^2 / (n-2). \quad (18)$$

Srovnáním  $B(\hat{\lambda})$  s  $B(\tilde{\lambda})$ ,  $\text{Var}(\hat{\lambda})$  s  $\text{Var}(\tilde{\lambda})$  a  $C(\hat{\lambda})$  s  $C(\tilde{\lambda})$  (t.j. srovnáním (17) s (12), (18) s (13) a (18) s (14)) zjistíme, že

$$\text{Var}(\hat{\lambda}) > \text{Var}(\tilde{\lambda}), \quad C(\hat{\lambda}) > C(\tilde{\lambda}), \quad |B(\hat{\lambda})| > B(\tilde{\lambda})=0,$$

t.j. odhad  $\tilde{\lambda}$  (který je nejlepší z nevychýlených odhadů) je po všech dosud uvažovaných stránkách "lepší" než maximálně věrohodný  $\hat{\lambda}$ : má nulové vychýlení, menší rozptyl i menší střední kvadratickou chybu.

Tím však není nabídka vyčerpána; zkusme volit i jiné odhady tvaru  $c_n/T$ , kde  $c_n$  by bylo vybráno podle určitých hledisek. Zvolíme-li  $c_n = n-2$ , dostaneme odhad

$$\lambda^* = (n-2)/T. \quad (19)$$

Ten má - na rozdíl od  $\tilde{\lambda}$  - záporné vychýlení

$$B(\lambda^*) = E(\lambda^*) - \lambda = -\lambda / (n-1). \quad (20)$$

Jeho rozptyl a střední kvadratická chyba jsou

$$\text{Var}(\lambda^*) = \lambda^2 (n-2) / (n-1)^2, \quad (21)$$

$$C(\lambda^*) = \lambda^2 / (n-1). \quad (22)$$

Srovnáním zase zjistíme

$$\text{Var}(\lambda^*) < \text{Var}(\tilde{\lambda}) < \text{Var}(\hat{\lambda})$$

$$C(\lambda^*) < C(\tilde{\lambda}) < C(\hat{\lambda})$$

Závěr: z hlediska rozptylu a střední kvadratické chyby je  $\lambda^*$  úplně nejlepší, z hlediska vychýlení horší než  $\tilde{\lambda}$ , ale rovnocenný s  $\hat{\lambda}$  (záleží jen na tom, zda dáme přednost soustavnému "podceňování" či soustavnému "přeceňování" skutečné hodnoty).

Nakonec se o přízeň statistické veřejnosti uchází ještě jeden kandidát z téže skupiny odhadů typu  $c_n/T$ . Je to tzv. *mediánový odhad*

$$\lambda^0 = \chi_{0.5}^2(2n)/(2T), \quad (23)$$

kde  $\chi_{0.5}^2(2n)$  značí 50% kvantil, t.j. medián, rozdělení  $\chi^2$  s  $2n$  stupni volnosti. Může se pochlubit dvěma vlastnostmi :

$$P(\lambda^0 \leq \lambda) = P(\lambda^0 \geq \lambda) = \frac{1}{2}, \quad (24)$$

což je určitá analogie vlastnosti jednostrannosti; skutečná hodnota parametru  $\lambda$  je mediánem rozdělení odhadu  $\lambda^0$ . Lze se o tom přesvědčit snadno užitím vztahu (9):

$$\begin{aligned} P(\lambda^0 \leq \lambda) &= P(\chi_{0.5}^2(2n)/(2T) \leq \lambda) = P(T > \chi_{0.5}^2(2n)/(2\lambda)) = \\ &= 1 - P(\chi^2(2n) > \chi_{0.5}^2(2n)) = \frac{1}{2}. \end{aligned}$$

Dále,  $\lambda^0$  minimalizuje střední hodnotu absolutní odchylky od skutečné hodnoty parametru, t.j.

$$E(|\lambda^0 - \lambda|) \leq E(|\lambda^* - \lambda|)$$

pro jakýkoliv odhad  $\lambda^*$ . Ostatní vlastnosti mediánového odhadu  $\lambda^0$  daného vztahem (24) jsou:

vychýlení  $B(\lambda^0)$  je rovno

$$B(\lambda^0) = E(\lambda^0) - \lambda = \lambda \frac{\chi_{0.5}^2(2n) - 2(n-1)}{2(n-1)}; \quad (25)$$

protože rozdělení  $\chi^2$  má kladnou asymetrii, jeho medián je vždy nejvýše roven střední hodnotě (která je rovna počtu stupňů volnosti), je

$$B(\lambda^0) < \lambda \frac{2n-2(n-1)}{2(n-1)} = \frac{\lambda}{(n-1)} = B(\hat{\lambda}). \quad (28)$$

Dále je zajímavý průběh distribuční funkce uvedených odhadů. Všechny mají tvar  $c_n/T$ , kde  $c_n$  je konstanta závislá na rozsahu výběru a na typu odhadu. Pro libovolný odhad dané třídy,  $\lambda_{est}$  je

$$\begin{aligned} P(\lambda_{est} < 1) &= P(c_n/T < 1) = 1 - P(T < c_n/1) = \\ &= 1 - P(\chi^2(2n) < 2\lambda c_n/1) \end{aligned} \quad (27)$$

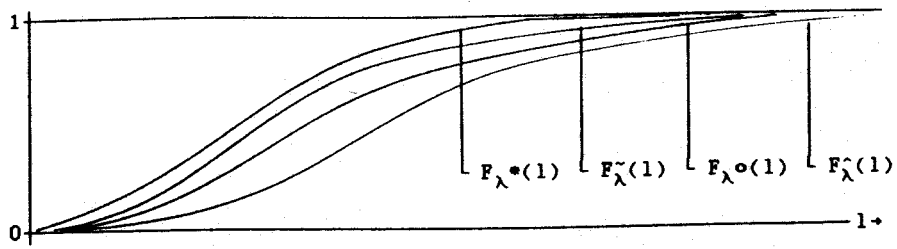
Konstanty  $c_n$  odpovídající různým typům odhadů mají toto pořadí podle velikosti:

$$c_n(\lambda^*) = n-2 < c_n(\tilde{\lambda}) = n-1 < c_n(\lambda^0) = \chi_{0.5}^2(2n)/2 < c(\hat{\lambda}) = n; \quad (28)$$

tudíž distribuční funkce odhadů splňují vztah

$$F_{\lambda^*}(1) > F_{\tilde{\lambda}}(1) > F_{\lambda^0}(1) > F_{\hat{\lambda}}(1) \quad (29)$$

pro všechna  $1$ , jak je vyznačeno na následujícím obrázku. Dokonce platí, že pro jakýkoliv výběr je  $\lambda^* < \tilde{\lambda} < \lambda^0 < \hat{\lambda}$ .



Kvantily odhadů  $\lambda_p^*$ ,  $\tilde{\lambda}_p$ ,  $\lambda_p^o$ ,  $\hat{\lambda}_p$ , t.j. kořeny rovnic typu

$$P(\hat{\lambda} < \lambda_p^o) = p$$

atd., jsou rovny

$$2\lambda c_n / \chi_{1-p}^2(2n), \quad (30)$$

kde  $\chi_{1-p}^2(2n)$  je 100p% kvantil rozdělení  $\chi^2$  o  $2n$  stupních volnosti a  $c_n$  konstanta, na které je založen příslušný odhad.

S užitím kvantilů snadno určíme intervaly, které bychom mohli nazvat "intervaly koncentrace", t.j. intervaly, uvnitř kterých bude odhad s danou pravděpodobností, např. pro  $\hat{\lambda}$  interval  $\hat{\lambda}_\alpha, \hat{\lambda}_{1-\alpha}$  takový, že

$$\alpha = P(\hat{\lambda} < \hat{\lambda}_\alpha) = P(\hat{\lambda} > \hat{\lambda}_{1-\alpha})$$

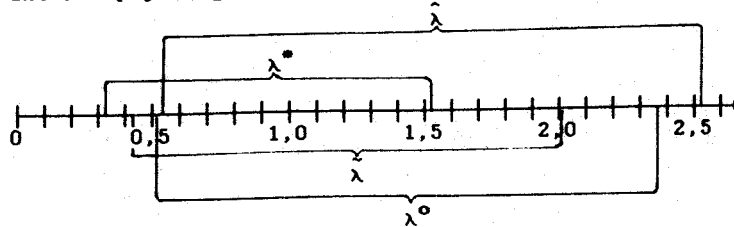
a tudíž

$$P(\hat{\lambda}_\alpha \leq \lambda \leq \hat{\lambda}_{1-\alpha}) = 1-2\alpha.$$

Např. pro  $n=5$  a  $\alpha=0,05$  jsou tyto intervaly uvedeny v následující tabulce:

odhad	dolní hranice	horní hranice
$\hat{\lambda}$	$0,5462\lambda$	$2,5381\lambda$
$\tilde{\lambda}$	$0,4370\lambda$	$2,0305\lambda$
$\lambda^*$	$0,3277\lambda$	$1,5228\lambda$
$\lambda^o$	$0,5103\lambda$	$2,3711\lambda$

Tyto intervaly jsou graficky znázorněny takto:





Tím je ukončeno stručné představení pouhých čtyř uchazečů o funkci odhadu intenzity poruch (intenzity úmrtnosti, intenzity Poissonova procesu, podle toho, o jakou úlohu zrovna běží). Nechme je ještě jednou projít přes podium a uvést předběžně jejich zásluhy:

Jméno	zásluhy		
	vychýlení	rozptyl	stř.kv.chyba
$\hat{\lambda}$ , maximálně věrohodný	$\frac{\lambda}{(n-1)}$	$\frac{\lambda^2 n^2}{[(n-1)^2(n-2)]}$	$\frac{\lambda^2(n+2)}{[(n-1)(n-2)]}$
$\tilde{\lambda}$ , nejlepší nestranný	0	$\frac{\lambda^2}{(n-2)}$	$\frac{\lambda^2}{(n-2)}$
$\lambda^*$ , minimalizující stř.kv.chybu	$\frac{\lambda}{(n-1)}$	$\frac{\lambda^2(n-2)}{(n-1)^2}$	$\frac{\lambda^2}{(n-1)}$
$\lambda^0$ , mediánový	$\left\langle \frac{\lambda}{(n-1)} \right\rangle$	$\left\langle \frac{\lambda^2 n^2}{[(n-1)^2(n-2)]} \right\rangle$	$\left\langle \frac{\lambda^2(n+2)}{[(n-1)^2(n-2)]} \right\rangle$

Pro který budete hlasovat? Já bych asi zvedl ruku pro  $\lambda^*$  nebo  $\tilde{\lambda}$  či  $\lambda^0$ ; maximálně věrohodný by byl až na posledním místě (to je ta zajímavost celého příkladu), přinejmenším do získání dalších informací (např. další intervaly koncentrace apod.).

*P.S. Skoro stejná povídka by se dala vyprávět např. o odhadu směrodatné odchylky normálního rozdělení, jenže by nebylo tak snadné uvádět explicitní výsledky a bylo by nutné uvádět numerické hodnoty. Jiné dotěrné otázky vyvstávají v souvislosti s tzv. cenzorovanými výběry, i když se používá jen jednoduchých tradičních kritérií kvality odhadu. Také by bylo možno uvést zajímavé příklady uvádějící v pochybnost rozumnost požadavku nevychýlenosti či minimálního rozptylu atd. O tom snad někdy příště.*

## Klub nepřátel statistiky

Jiří Žvaček

Zejména na těch početně nejsilnějších statistických katedrách se projevuje určitá nechuť k nově vznikající Statistické společnosti. Část pracovníků právem namítá, že zde není prezentován dostatečně přitažlivý program, který by vyvážil nevýhody klubového života a i jinak se nemohou s programem ztotožnit.

Domnívám se, že tento stav lze připsat spíše neinformovanosti, než samotnému programu. Rád bych upozornil, že počítáme se založením sekce, která ve které naleznou uplatnění i ti, kteří byli ve statistice sice ve většině, ale přesto zatím v ilegalitě.

Mám zato, že větší část odborníků spíše svou prací tíhne k destrukci statistiky, a mohl bych to dokumentovat. Jsem dokonce přesvědčen, že celá československá statistika má tento rys.

Vzhledem k tomu, že každého vedou k této činnosti jiné důvody, a řada odborníků se bude spíše organizovat v rámci sekcí či poboček (ačkoliv svou činností nesporně patří k nám), uvažujeme spíše o klubu bez jednotného vedení. Členy klubu by mohli být perspektivně i nečlenové statistické společnosti, ale myslím, že obecně stěžejí naleznou lepší platformu.

Zda má nebo nemá vzniknout samostatný Klub nepřátel statistiky rozhodne sám život. Nicméně určité potenciální konflikty se již nyní rýsují a nelze tuto možnost zcela vyloučit.

Zatímco zatím například neuvažujeme o čestných členech statistické společnosti, mnozí se naopak návrhy na čestné členy klubu nepřátel statistiky. A co pak s těmi, kteří jsou nesporně čestnými (uvažujeme i o zasloužilých!) členy Klubu nepřátel statistiky již nyní, ale jejich věhlas ještě nestačí na čestné členství ve Společnosti?

Tyto a další okruhy otázek lze také řešit v nově vznikající Statistické společnosti!

Jako předběžný plán činnosti navrhuje zcela konkrétní akce:

1. Každoroční veřejné vyhlášení cen pro největší statistickou pitomost ve třech oblastech
  - a) Údaj roku, věnovaný té organizaci nebo jednotlivci, který nejvíce zkreslí konkrétní údaj,
  - b) Interpretace roku, věnovaný nejpovednějšimu zkreslení statistických údajů ve veřejných sdělovacích prostředcích.
  - c) Práce roku, věnovaná té publikaci, ve které bude relativně největší počet chyb.

Členové klubu jsou povinni registrovat potenciální kandidáty a předávat materiály k vyhodnocení komisi pro tento účel vytvořené.

2. Pro dubnová čísla organizovat pro renomované časopisy články, které budou pod pláštěm serióznosti zesměšňovat statistiku a její metody.

3. Jako nejvyšší poctu společnosti prosadit titul Statistik roku. Soutěže o nejvyšší poctu se může zúčastnit kromě článků publikovaných v dubnových číslech i články předem nahlášené vedení klubu (jinak může být zařazen pouze do soutěže o práci roku). Informovaní členové klubu jsou vázáni v těchto případech mlčenlivostí a musí usilovat v rámci svých možností o publikaci těchto příspěvků.

4. Jako klub podávat průběžně zprávy o své činnosti v rámci bulletinu statistické společnosti.

Uvítáme i všechny pro činnost Klubu i když základní program již nejenom existuje, ale ve své aplikační části se i plní.

Přihlášky přijímá

Jirka Žváček

## COMPSTAT '90 - programy prezentované účastníky

Jiří Militky

O symposiu COMPSTAT '90 je v Bulletinech uvedena řada sdělení. Ve svém příspěvku bych se chtěl zabývat statistickým software prezentovaným některými účastníky. Většina těchto softwarových produktů je stručně popsána v katalogu software, který všichni obdrželi spolu se sborníkem (vydavatel Physica Verlag) a dvoustránkovými abstrakty krátkých sdělení. Vzhledem k různorodosti software a jeho úrovně se zaměřím spíše na to, co mě nějak zaujalo, resp. to, co je možné získat od jednotlivých autorů bez velkých finančních nákladů.

### Presentace software

Vlastní presentace software probíhala v PC-roomu odděleně od ostatního dění a paralelně s řádným programem konference. Bylo k dispozici celkem pět osobních počítačů se 40 Mbyte hard disky a šachtami pro 3,5 a 5,25 palcové diskety. Pro řadu účastníků se jako oříšek ukázala grafická karta typu Hercules. Někteří nebyli schopni provést příslušné úpravy svých programů, takže nic neprezentovali. Vlastní provoz v presentační místnosti nebyl nijak řízen. To mělo za následek, že v prvních dnech bylo obtížné vůbec něco někomu ukázat. Většina návštěvníků spíše průběžně kontrolovala, zda se náhodou na discích nevyskytuje něco zajímavého. Postupně však ti, co chtěli něco prezentovat, zamluvili jednotlivá PC na určité hodiny a skupinky účastníků konference měly možnost výběru co a kdy chtějí vidět. Zajímavým postřehem je, že mezi lovce software nepatřili pouze účastníci zemí Východní Evropy, ale řada ostatních, včetně známých statistiků. Na druhé straně se zde prakticky neobjevil žádný z Američanů (kromě předvádějících). Vlastní předváděný software lze rozčlenit dle několika hledisek. Pro řadu účastníků bylo v prvním kole rozhodující, zda jde o software "free of charge", který si mohou okopírovat, nebo zda jde o ukázkou programů řádně prodávaných. Pro účely tohoto příspěvku je vhodné rozdělit software na programové soubory a izolované programy.

Z tohoto dělení se poněkud vymykají softwarové doplňky k učebnicím Robinsona a Bowmana, které jsou označeny "Počítačem ilustrované texty". Jde o soubor programů ilustrujících jednotlivé techniky statistické analýzy. Využívá se zejména

postupů animace a simulace spolu s možnostmi interaktivní grafiky. Jednotlivé grafy jsou ilustrovány "krok po kroku" tak, aby mohl uživatel postupně ověřovat své znalosti.

Počítač má tři hlavní úkoly:

- elektronická "tabule" pro interaktivní řešení úloh,
- "učitel" pro ověřování znalostí probírané látky
- běžný software pro statistické úlohy.

Na základě předvedených demonstrací tohoto software se zdá, že bude vhodný zejména pro výuku statistiky a pravděpodobnosti na základní úrovni pro začátečníky. Zatím jsou k dispozici tři knihy s tímto atraktivním druhem software (pravděpodobnost, statistika a regrese s analýzou rozptylu).

#### Programové soubory

V této kategorii software byly prezentovány jako obecné statistické programové pakety (IDAMS, ADSTAT), tak i pakety pro speciální aplikace (MESOSAURUS, DISTAN, STATXACT). Vzhledem k problémům s grafickou kartou nebyl prezentován systém TSTAT-16 pro analýzu časových řad, i když je uveden v katalogu.

#### IDAMS

Je obecný statistický paket vyvinutý v UNESCO především pro analýzu rozsáhlých datových souborů. Obsahuje řadu metod zejména pro předzpracování dat, základní statistickou analýzu dat a regresi. Je vytvořen v jazyce FORTRAN a využívá minimálně možnosti PC. Jeho hlavní předností je zřejmě to, že ho lze získat pro nekomerční účely zdarma. Postačí napsat na adresu:

G. Del Bigio  
Chief Management Systems Section  
UNESCO  
7 place de FONTENDY  
757 00 PARIS  
FRANCE

Na základě vlastních testů však mám dojem, že nebude pro řadu statistiků prakticky žádným přínosem.

#### ADSTAT

Jde o rozsáhlý interaktivní statistický paket určený zejména pro řešení technických úloh. Vzhledem k tomu, že jsem ho prezentoval, nechci podrobněji uvádět jeho výhody a nevýhody. Připravujeme však pro některé další číslo Bulletinu informativní příspěvek o tomto původním interaktivním tuzemském produktu.

Zájemci o ADSTAT mohou získat podrobnější informace včetně demonstrační diskety na adrese:

TriloByte s.r.o.  
Pošt. schránka 31  
160 00 Praha 6  
tel. 02-763030

#### MESOSAURUS

Demonstrační verze programového balíku pro analýzu časových řad MESOSAURUS 1.0 vypracovaného Kuzněcovem a kol. (SSSR) asi řadu účastníků prezentace velmi příjemně překvapila. Kromě jednoduché obsluhy (hierarchický systém nabídek) a obsáhlých statistických možností (od modelů trendu přes vyhlazování, filtraci a spektrální analýzu až k modelům typu ARIMA), obsahuje balíček kvalitní grafický editor. Zajímavá je také nabídka "REFERENCE", obsahující popis metod a algoritmů s klíčovými slovy (podobně jako jsou nabídky "HELP" produktů firmy Borland). Lze konstatovat, že úroveň provedení se tento balíček blížil špičkovým profesionálním produktům. Pouze manuál, kteří autoři ukazovali, působil spíše jako výzkumná práce. Podle předběžných zpráv je snahou autorů využít ke komerční distribuci firmu SYSTAT.

#### DISTAN

Tento balíček je produktem maďarského informačního centra sociálních věd (T. Rudas). Je určen pro manipulaci, transformaci a analýzu nominálních dat. Obsahuje tyto základní možnosti:

- maximálně věrohodné odhady pro smíšené lineární a multiplikační modely,
- využití vážené MNC pro modely s ordinálními proměnnými,
- booleovskou faktorovou analýzu,
- vícerozměrnou korespondenční analýzu,
- rozklad směsí,
- binární shlukovou analýzu.

V přípravě jsou ještě další programové moduly zejména pro průzkumovou analýzu ordinálních dat. Pro ukládání dat a manipulaci se používá speciální systém souborů. Balíček je ovládán systémem nabídek. Není však zdaleka tak uživatelsky přátelský jako kvalitní programy pro PC. Také jeho ovládání je poněkud těžkopádnější. Systém lze ve verzi 1.0 obdržet zdarma na adrese:

T. Rudas  
Social Science Informatic Center  
H. 1027 Frankel Leó 11, Budapest.

### STATXACT

Jde o statistický paket umožňující konstrukci neparametrických testů s exaktním výpočtem kvantilů pomocí permutačních metod. To umožňuje použití neparametrických testů i pro malé výběry, které mohou být silně sešikmené. Autorem algoritmů je C. Mehta z USA. Paket se prodává řádově za stovky dolarů.

### Jednotlivé programy

V této skupině byly prezentovány programy pro řešení speciálnějších úloh jako je konstrukce přejímacích plánů (SESAMPLA), grafické modelování vícerozměrných diskrétních dat (DIGRAM), konstrukci experimentálních plánů (KEYFINDER), analýzu ordinálních dat (CATANA) a realizaci technik typu Bootstrap (BOJA).

Většinou jde o programy vyvinuté jednotlivci a jejich cena se pohybuje kolem 100 dolarů.

Na velmi dobré úrovni z hlediska ovládání je BOJA a CATANA. Oba tyto programy používají systém nabídek typu "Borland like". Program CATANA je vlastně doplňkem knihy B. Andersena: The Analysis of Categorical Data, Springer 1990.

Na druhé straně programy DIGRAM a KEYFINDER nejsou uživateli příliš přátelské a nemají postačující systém nápovědy. Pro detailnější zhodnocení pouhá prezentace zdaleka nestačila. Vážnější zájemci mohou kontaktovat přímo autory, jejichž adresy mají všichni účastníci symposia COMPSTATu v katalogu programů. Ostatním je na požádání mohu zaslat.

Kromě výše uvedených byla v rámci symposia COMPSTAT šířena některými účastníky nabídka řady dalších programů pro PC. Zajímavá byla zejména nabídka programů holandského centrálního statistického úřadu obsahující řadu izolovaných statistických programů v ceně kolem 100 dolarů. Některé z nich (např. BLAISE a SCATTER) byly demonstrovány již na předchozím COMPSTATu v Kodani.

### Závěr

Z tohoto, ne zcela úplného, přehledu je patrné, že také v oblasti software prezentovaného účastníky symposia byla řada zajímavých věcí. Lze očekávat, že i tento druh software nalezne u nás své řádné uživatele.

Jiří Žváček

Zatímco někteří vesele statističili na konferencích, museli jiní nést prapor pokroku v domácím statistickém prostředí. Takže se mi přirozeně zdá, že COMPSTAT trochu upadá. Buď nám něco tají, nebo nevědí co s těmi počítači dělat.

Stejně bych ale rád k tomu něco řekl. Třeba proč je třeba si sborníky od účastníků půjčit a o čem je třeba přemýšlet - co se mi zkrátka v nich líbilo.

#### a) Proceedings

**Algoritmy.** Hodně pozornosti se věnuje shlukové analýze (líbilo se mi třeba, že náhodně volí centra shluků ve Stochastic Algorithms for Clustering), ale jinak se spíše popisují algoritmy paketů (DISTAN, CLUSE).

**Pakety.** Je trochu popsán BOJA (\*), realizující bootstrap a jack-knife. (zájemcům sděluji, že mám RESAMP, který může sloužit k elementární výuce simulace a bootstrapování a že jsem jej ochoten poskytnout tomu, kdo si pošle disketu).

S language (\*) má všechno, ale pouze ukazovat obrázky dynamické grafiky je pro učence trochu málo. Zato informace o modelech v připravované nové verzi (Statistical Models in S) ukazuje, kam se asi bude vyvíjet statistický jazyk. Kdo má UNIX s S nebo se na to chystá, ať se ozve. Vidím v tom budoucnost a založíme klub uživatelů.

**Expertní systémy.** Trochu se mi líbil článek Estimation Procedures for Language Context. Úloha je praktická - oprava překlepů.

WAMASTEK (\*) je zpráva o fungujícím systému v oblasti, které ještě řada statistiků trochu rozumí - výběru jednorozměrných popisných charakteristik (sám jsem zkoušel expertní systém v této oblasti). Potěšily mne výsledky experimentu, srovnávajícího systém s názorem expertů a to v obou směrech.

Z analýzy expertů mj. vyplývá, že

- statistik se řídí spíše svou zkušeností, než daty,
- používají se pouze notoricky známé míry (kam nepatří šikmost, špičatost, kvartilové míry),



- míry se volí individuálně, bez ohledu na jejich vztahy.

To, že je dobrá shoda mezi expertem a strojem již pak nepřekvapí.

Jinak většina autorů nových metod nabízí spolu s novou metodou i programy a stále častěji i nabídkově řízené pakety. U některých metod se hovoří i o příkazovém jazyku.

#### b) Short Communications

Zaujala mne (již opakovaně) myšlenka Spreadsheet as a Statistical Platform. Chlapíci sice používají WingZ na Macintoshi, ale zdá se mi to idea dobrá a diplomanti již dělají malý paket pro sezónní očišťování v Quattro Pro (doporučuji).

Abacus je speciální program pro rychlý a kvalitní výstup tabulek (zrychlení oproti SPSS 300 krát, camera ready). Ještě stále jsou věci, které se dají zlepšit!

Více pozornosti je konečně věnováno databankovým systémům. Zatím zřejmě do povědomí ještě zcela neproniklo, co to vlastně je a k čemu všemu se tento způsob myšlení dá použít.

Stále častěji se objevuje název Hyper Card. Je to exkluzivní multi-media (něco více než databankový) systém pro Apple s velmi jednoduchou obsluhou. Na jeho bázi je možno dělat zejména krásné systémy s vysvětlením - ale to se musí vidět a ne číst od Hietyly.

Častým problémem je identifikace algoritmu generování náhodných čísel (\*). Je uveden příklad.

#### c) Software catalogue

Docela se mi líbila koncepce DMS (\*), což je dynamický vícesektorový model (zřejmě francouzské) ekonomiky. Lze v něm volit různé varianty okolí a rozhodnutí, produkční funkce - pokud to funguje, je to dobrý nápad jak spojit data s modelem!

Popis systému pro vyučování statistiky je nekonkrétní.

-----  
Pokud má někdo zájem o konkrétní věc (označenou hvězdičkou - BOJA, RESAMP, S, WAMASTEX, identifikace, DMS), pošlu mu kopii příspěvku (když se ozve).

Jinak jsou sborníky u účastníků.

Jirka Žváček

## NEJLEPŠÍ Z ROBUSTA

Připravujeme vydání sborníku vybraných článků z konferencí ROBUST 82-90. Tyto konference pořádá každé dva roky Jednota československých matematiků a fyziků a jsou věnovány netradičním metodám matematické statistiky. Velká pozornost je vždy soustředěna na výpočetní aspekty statistických metod a na problémy spojené s jejich aplikací. Navrhovaný výběr článků obsahuje především články věnované buď obecnějšímu přehledu určité třídy metod či obecnějším problémům aplikací. Podmínkou výběru byla samozřejmě i srozumitelnost širší statistické veřejnosti. Články z konferencí ROBUST, které vyšly ve sbornících jen pro účastníky těchto konferencí byly vždy vyhledávaným zdrojem informací - většina témat nebyla jinde v české literatuře zpracována. Především navržený výběr obsahuje práce J. Antocha o neparametrických odhadech regresních křivek, T. Cipry o nových přístupech k analýze časových řad, M. Huškové o detekci změny regrese, T. Havránka o mnoho-rozměrných kontingenčních tabulkách a výběru modelů, J. Jurečkové o robustnosti, práce J. Kláčky "O jedné dosti obecné a velmi účinné metodě statistického důkazu (čehokoliv)" a "O zákonu počáteční hodnoty", práce V. Lánské o Coxově modelu v analýze přežívání, J. Machka o analýze experimentů s kvalitativní odpovědí, D. Pokorného o studiu závislosti v kontingenčních tabulkách a o paradoxech v kontingenčních tabulkách, J. Řeháka o analýze ordinálních dat, L. Tomáška o odhadu řádu autoregresního modelu, I. Vajdy o statistických odhadech s minimální divergencí, práci P. Vopěnky, D. Pokorného a J. Šimka "Metafora o výzkumu multifakto-

riálně podmíněných dat" spolu s dosud nepublikovanou odpovědí K. Bendové, práce K. Zváry o identifikaci odlehých pozorování v lineární regresi, J. Žváčka o aplikaci částečné a segmentové regrese a J. Žváčka a H. Řezankové o statistických programovacích jazycích. Všechny práce budou recenzovány a doplněny pokud možno o poznatky, které přibyly od jejich publikování. Doufáme, že tento sborník bude užitečným i zábavným čtením nejen pro členy naší společnosti.

(Ha)

#### Krátké zprávy

#### O 11. Pražské konferenci o teorii informace, statistických rozhodovacích funkcích a náhodných procesech.

Kromě tradičního organizátora - sektoru Teorie informace ÚTIA ČSAV - se na zajištění zdárného průběhu konference podílela nemalou měrou i katedra pravděpodobnosti a matematické statistiky MFF UK, v jejímž areálu v Troji se konference ve dnech 27.8 - 31.8. 1990 uskutečnila. Studentské prostředí kolejí a menzy omlazujícím způsobem vlévalo nové síly a dobrou náladu do zhruba 170 účastníků, kteří vytrvaleckým stylem absolvovali více jak 120 přednášek (včetně 25 pozvaných). Nejtěžší zkouškou prošli účastníci navštěvující sekci "Robustní statistiky", která probíhala po celý týden na rozdíl od sekcí "Teorie pravděpodobnosti", "Umělé inteligence", "Náhodných procesů", "Teorie informace" a "Teorie her a multikriteriálního rozhodování", jejichž účastníci si alespoň občas mohli vydechnout, či se dokonce jít podívat na některou (pozvanou, přehledovou) přednášku do jiných sekcí. Organizátoři programu jsou si vědomi přinejmenším potenciálního zájmu účastníků vyslechnout některé zvané přednášky v příbuzných sekcích se pokusili zvládnout tento fakt v uspořádání programu. Přesto však z hlediska profesního profilu některých účastníků došlo ke kolizi atraktivních přednášek, a tak byla prakticky ověřena stará

pravda, že ke každé verzi programu kterékoliv konference existuje neprázdná třída účastníků, kteří by jej sestavili lépe. Ačkoliv pozvanými řečníky byli renomovaní odborníci z výše uvedených oborů, ozvaly se hlasy, že ne o všech lze říci, že splnili (možná přehnaná) očekávání. Avšak ať už si účastníci pochvalovali či kritizovali náplň a úroveň přednášek, drtivá většina se shodla na tom, že konference pro ně znamenala přínos také v diskusích, probíhajících mimo nalinkovaný program. Mnozí se upřímně těšili ze společně prožívaných chvil s dlouholetými přáteli z celého světa či ze vzdálenějších míst naší vlasti. Opětovně jsme se utvrdili v radostném faktu, že nás s našimi kolegy ze Slovenska spojují daleko podstatnější věci než spojovník v názvu republiky.

Bylo by jistě předem ztracenou snahou chtít alespoň telegraficky popsat nové výsledky ohlášené na konferenci, či se dokonce pokusit vybrat některé nejzávažnější. Počkejme si tedy na sborník konference, jehož vydání očekáváme v druhé polovině příštího roku a těšíme se již nyní na další konferenci, již 12. Pražskou.

J. A. Višek

Ze světa

Dne 14.3.1990 se uskutečnilo řádné Valné shromáždění Slovenskej štatistickej a demografickej spoločnosti. Tehdy ještě Slovenskej demografickej a štatistickej spoločnosti, neboť právě na tomto shromáždění došlo k přejmenování ("jako vedlejší produkt našeho úsilí" <sup>1</sup>) a k některým změnám ve stanovách společnosti. SŠDS vydává obdobu našeho bulletinu, tzv. "Informácie", jejichž 2. číslo (říjen 1990) jsme v listopadu dostali. Z tohoto čísla citujeme:



<sup>1</sup> J. Žváček, IB č. 1

VZ (Valné zhromaždenie) SDŠS v tajných voľbách volilo predsedu spoločnosti, členov výboru a revízorov spoločnosti. Za predsedu SDŠS bol zvolený Doc. Ing. Miroslav Abrahám, CSc.

Za členov Výboru boli zvolení všetci navrhnutí:

Abrahám Miroslav	Chajdiak Jozef	Krč Rudolf	Sojková Zlatica
Bakytová Hedviga	Janiga Ivan	Kvetko Juraj	Štrauch Dušan
Brezák Jozef	Kevická Benée	Luha Ján	Thurzo Igor
Finková Zuzana	Komorník Jozef	Mišíková Ivica	Volná Anna
Halašová Eva	Kovačka Milan	Očovský Štefan	Wimmer Gejza
Hanzlík Ján	Kozaňák Jozef	Pastor Karol	
Holičková Edita	Kozíková Emília	Potocký Rastislav	
Za revízorov boli zvolení:		Tirpák Michal	Šlauková Iveta
		Pardelová Růžena	

#### Štruktúra Spoločnosti:

Slovenská štatistická a demografická spoločnosť má ku dňu 1.9.1990 342 členov (z toho 8 z Prahy - pozn. red.). Spoločnosť riadi výbor. Funkcionármi výboru sú: predseda, štyria podpredsedovia, vedecký tajomník, hospodár, predsedovia pobočiek.

Od roku 1990 sú v spoločnosti organizované štyri sekcie. Prácu jednotlivých sekcií riadia podpredsedovia Spoločnosti:

- aplikovanej štatistiky - Ing. Jozef Brezák,  
(VUSEIaR Bratislava),
- matematickej štatistiky - Doc. RNDr. Rastislav Potocký, CSc.,  
(KTP a MŠ MFF UK Bratislava),
- výpočtovej štatistiky - RNDr. Ivan Janiga, CSc.,  
(SF SVŠT Bratislava)
- demografie - Doc. Mudr. Anna Volná, CSc.,  
(LF UK Bratislava).

Spoločnosť je územne členená na centrum v Bratislave a pobočky. Už založené pobočky pracujú v Nitre, v Banskej Bystrici a v Košiciach. Propravuje sa založenie pobočky v Žiline. Predsedami výborov v pobočkách Spoločnosti sú

- v Nitre - Ing. Zlatica Sojková, CSc.,
- v Banskej Bystrici - Doc. Ing. Miroslav Abrahám, CSc.,
- v Košiciach - Doc. Ing. Pavol Kováč, CSc.

**Plán práce Spoločnosti:**

Každý druhý rok (párny) sa organizuje x-tá Slovenská konferencia štatistikov. Konferencia je trojdňová, prebieha v DVP Smolenice.

Každý druhý rok (nepárny) sa organizuje x-tá demografická konferencia. Konferencia je dvojdňová, prebieha v DVP Smolenice.

Každý rok sa organizuje x-tá škola štatistiky - Využitie štatistických metód v sociálno-ekonomickej praxi - EKOMSTAT. Škola je týždenná.

Počínajúc nasledujúcim rokom sa ako spoluorganizátori podielame na Letnej škole pravdepodobnosti a matematickej štatistiky - PROBASTAT. Škola je týždenná.

S nepravidelnou periodicitou prebieha dvojdňový seminár Výpočtová štatistika.

Každoročne sa organizuje niekoľko kurzov v rámci cyklu Štatistický software pre osobné počítače triedy XI a AT.

Každý rok sa uskutočňuje prehliadka diplomových prác a prác SVOČ z oblasti štatistiky a demografie. Spoločnosť podľa možnosti alebo potreby organizuje aj ďalšie akcie.

Predbežne sa organizujú prednášky.

V najbližšom období sa plánujú nasledovné väčšie akcie:

- 11.- 12.12.1990 - Tretí seminár Výpočtová štatistika  
(v Bratislave)
- máj 1991 - prehliadka prác študentov a mladých vedeckých pracovníkov z oblasti štatistiky a demografie  
(v Bratislave)
- 3.- 7. 6.1991 - 4. škola štatistiky EKOMSTAT'91  
(v Trenčianskych Tepliciach)
- 26.- 30. 8.1991 - Konferencia o pravdepodobnosti a matematickej štatistike - PROBASTAT'91  
(v Bratislave)
- 16.- 17. 9.1991 - 3. demografická konferencia  
(v Smoleniciach)

## Vánoční dárek Stanislava Komendy

---

### Klasická definice :

Mimořádným je profesor, který nedokázal udělat nic řádného.

Řádným je profesor, kterému se nepodařilo vykonat nic mimořádného.

★

Je pravděpodobné, že vědecké zkoumání se neobejde bez mozku;  
Je jisté, že se neobejde bez zadku, na kterém je třeba sedět  
u pracovního stolu ...

★

**Nula** je to, co přidáno k čemukoli, cokoli nezmění.

**Nula** je to, co násobením čehokoli cokoli činí sebou.

**Nula** je to, co dělením čehokoli cokoli činí nekonečným

★

### Neřešitelný

je problém, na jehož řešení došly prostředky.

★

### Statistika

vypovídá o tom, co je.

### Matematická statistika

vypovídá o tom, co by být mohlo,

vezme-li se v úvahu, že to, co je, je právě takové,  
jaké je.

★

**Sókrates** (po absolvování kurzu statistiky) pochybuje:

"Mám tušení, že by tu mohlo být něco, o čem něco nevím."

**Statistik:**

"Vím, že nic nevím - ale nevím to úplně jistě."

★

Udělat chybu je nepříjemné. Opravdu nebezpečná je však jenom  
taková chyba, jejíž příčinu se nepodařilo najít.

★

### Zamítnutí hypotézy

není důvod ke smutku - je to důvod k vymyšlení nových hypotéz.

- \* Omlouváme tyto nepřesnosti v říjnovém čísle (vzniklé přepisem příspěvku z COMPSTATU, který byl přepsán z autorova rukopisu, aby se stihla uzávěrka):

na celkový počet 500 účastníků  
grafové modely  
program CATANA ke knize Andersona  
C. Mehta

Postřehy o programovém vybavení jsou od Tvrdíka (příspěvek měl být kolektivním dílem).

- \*\* Obdrželi jsme od firmy TRILOBYTE Inc. demonstrační verze programů ADSTAT a SCISTAT. Diskety spolu s některými informačními materiály jsou k dispozici u vědeckého tajemníka Společnosti Dr. G. Dohnala. Podrobnější informace o uvedených systémech lze obdržet od Doc. Ing. J. Militkého, CSc., VŠST, Hálkova 6, 460 00 Liberec.

23. ledna 1991 se bude v aule VŠE v Praze konat

### 1. výroční konference

## České Statistické Společnosti

Předběžný program :

13.00 - 14.30 Zpráva výboru, organizace voleb  
14.30 - 15.00 Přestávka a volby  
15.00 - 16.00 Přednášky o stavu oboru