

Informační Bulletin



České Statistické Společnosti

č.1. leden 2002, ročník 13

Motto: "Vždyť je to tak prosté, paní Watsonová!"
(Z rozhlasového detektivního seriálu)

Čím měřit variabilitu okolo harmonického průměru^{*/}

Václav Čermák

Úvod

Harmonický průměr, jak známo, patří do skupiny mocninových průměrů. V učebnicích a příručkách bývají z nich uváděny obvykle tyto čtyři:

A – aritmetický

G – geometrický

H – harmonický

K – kvadratický

Výsadní postavení mezi nimi zaujímá průměr aritmetický, a to nejen jako nejčastěji používaný, ale především proto, že je jádrem takzvaného *funkcionálního průměru*

$$g^{-1}\left(\frac{1}{n}\sum_i^n g(x_i)\right)$$

kde g je nějaká ryze monotónní funkce a g^{-1} je funkce k ní inverzní. Volbami funkce g ve tvarech x , $\log x$, x^{-1} a x^2 dostáváme výše uvedené specifické průměry A, G, H a K. Připomínáme tuto skutečnost proto, že nám za chvíli pomůže při zdůvodnění konstrukce míry rozptýlenosti hodnot okolo harmonického průměru.

^{*/} První verze tohoto příspěvku byla přednesena na mezinárodní vědecké konferenci, konané v Popradu 30.–31. srpna 2000.

Druhá okolnost, která stojí za připomenutí, je kritérium volby druhu průměru při udání polohy (úrovně) řady hodnot x_i . V popisné statistice se často doporučuje takzvaná *určující vlastnost průměru*: při smysluplnosti součtu převratných hodnot $\sum 1/x_i$ - a požadavku zachování jeho velikosti - je namístež užití průměru harmonického, zatímco průměr aritmetický přichází v úvahu při požadavku zachování součtu $\sum x_i$. (Užití průměru geometrického, jak víme, je podmíněno zachováním součinu $\prod x_i$.)

Na závěr této úvodní pasáže ještě připomeňme, že variabilita (rozptýlenost) okolo *aritmetického* průměru se měří nejčastěji pomocí směrodatné odchylky s nebo průměrné absolutní odchylky d , a že v obou případech se přitom vychází z téže metriky, a to $|x_i - A|$. Vzpomeňme, že obvyklé definice těchto měr variability se dají psát

$$d = \frac{1}{n} \sum_i^n |x_i - A| = \frac{1}{n} \sum_i^n |d_i| ,$$

což je aritmetický průměr hodnot $|d_i|$,

$$s = \sqrt{\frac{1}{n} \sum_i^n |x_i - A|^2} = \sqrt{\frac{1}{n} \sum_i^n |d_i|^2} ,$$

což je kvadratický průměr těchto hodnot.

Dále ještě stojí za připomenutí, že obě míry zaujímají důležité postavení nejen v popisné (deskriptivní), ale i v induktivní statistice: při prostém náhodném výběru bývá zvolena - z řady možných statistik (funkcí výběrových pozorování) - jako „rozumný“ odhad parametru σ Gaussova rozdělení právě směrodatná odchylka s , zatímco průměrná odchylka d je analogicky dobrým odhadem parametru δ rozdělení Laplaceova (dvojitě exponenciálního).

Dva příklady

Uvedeme dva klasické ilustrační příklady na použití harmonického průměru: výpočet průměrné pracnosti a výpočet průměrné rychlosti. (V obou případech zatím postačí miniaturní datový soubor, složený toliko ze dvou kladných hodnot; nechť tedy $0 < x_1 < x_2$.)

Příklad 1. V dílně provádějí dva dělníci stejnou pracovní operaci; prvnímu trvá tato operace 2 minuty, druhému 3 minuty. Jak známo, *aritmetický* průměr (2,5 min.) není správným druhem průměru, protože práce dělníků s touto výkonností by znamenala snížení celkového objemu výroby. Naproti tomu *harmonický* průměr (2,4 min.) by objem výroby zachoval. Jinak řečeno, nemá smysl usilovat zde o zachování součtu $x_1 + x_2$, ale součtu $1/x_1 + 1/x_2$. Přejít od hodnot x_i k hodnotám $1/x_i$ se dá interpretovat i věcně, a to jako záměna tzv. nepřímého vyjádření produktivity práce jejím vyjádřením přímým.

Příklad 2. Cyklistický závodník jeden určitou trasu dvakrát: cestou tam rychlostí 30 km/hod., cestou zpět 50 km/hod. (Řekněme při známém závodu Praha-Přibram-Praha.) Průměrná rychlost není 40 km/hod., tedy aritmetický průměr, ale 37,5 km/hod. jakožto hodnota průměru harmonického. (Pouze při jízdě touto rychlostí tam i zpět by celá cesta trvala stejnou dobu!)

Připojme k těmto dvěma elementárním příkladům z oblasti popisné statistiky ještě příklad na harmonický průměr z oblasti počtu pravděpodobnosti. Uvažujme kladnou náhodnou proměnnou X , jejíž rozdělení je udáno distribuční funkcí F . Harmonický průměr je zřejmě definován vzorcem

$$H = E^{-1}(X^{-1}) = \left(\int_0^{\infty} \frac{dF}{x} \right)^{-1},$$

kde integrál zde použitý je v pojetí Stieltjesově. (Tím je dosaženo, že pod jedním zápisem je zahrnut jak případ spojitých, tak i diskrétních proměnných - viz např. Kendallovu učebnici, odst. 1.23.) Místo H se často uvádí definice veličiny $1/H$, tedy

$$\frac{1}{H} = \int_0^{\infty} \frac{dF}{x}.$$

Jde samozřejmě o shodné definice; její druhá podoba nám však bude více vyhovovat při konstruování a zkoumání slíbené míry variability.

V případě *diskrétní* náhodné proměnné, kdy $dF(x) = P(x)$, dostáváme tento definiční vzorec ve tvaru

$$\frac{1}{H} = \int_0^{\infty} \frac{dF}{x} = \sum_x \frac{P(x)}{x},$$

kde $P(x)$ značí funkci masы pravděpodobnosti.

V případě *spojité* náhodné proměnné, kdy $dF(x)/dx = f(x)$,

$$\frac{1}{H} = \int_0^{\infty} \frac{dF}{x} = \int_0^{\infty} \frac{f(x)dx}{x},$$

kde $f(x)$ značí funkci hustoty pravděpodobnosti. (Integrál Stieltjesův zde přechází do svého speciálního případu – do integrálu Cauchyova-Riemannova.)

Příklad 3. Slíbený příklad podáme na variantu druhou, tj. na proměnnou spojitou. Uvažujme beta rozdělení prvního druhu

$$f(x) = \frac{1}{\beta(p, q)} (1-x)^{p-1} x^{q-1}, \quad 0 \leq x \leq 1, \quad p, q > 0.$$

Potom pro $q > 1$ dostáváme

$$\frac{1}{H} = \frac{1}{\beta(p, q)} \int_0^1 (1-x)^{p-1} x^{q-2} dx = \frac{\beta(p, q-1)}{\beta(p, q)} = \frac{p+q-1}{q-1}$$

a tedy

$$H = \frac{q-1}{p+q-1}.$$

Poznámka. Pro porovnání připomeňme, že *aritmetický* průměr (střední hodnota) u tohoto rozdělení činí

$$A = \frac{q}{p+q}.$$

Úkol a jeho řešení

Jak je zřejmé již z názvu celého článku, jde o otázku, čím změřit (charakterizovat) rozptýlenost hodnot okolo *harmonického* průměru. Tato otázka - pokud je nám známo - nebyla dosud v odborné literatuře

řešena (na rozdíl třeba od průměru geometrického). Podáme zde proto jeden námět na odpověď.

V první řadě je nutno odmítnout myšlenku, že by se rozptýlenost (variabilita) mohla i v těchto případech měřit analogickými charakteristikami, tj. střední kvadratickou odchylkou nebo průměrnou absolutní odchylkou, a to nikoli od aritmetického, leč harmonického průměru. V prvních dvou příkladech, uvedených výše, je totiž dobře patrné, že odchylky $|x_1 - H|$ a $|x_2 - H|$ jsou různě velké. Intuitivně cítíme, že jádrem hledané míry by měla být metrika, která by se touto shodností naopak vyznačovala.

Vyjdeme z faktu, připomenutého výše, že totiž konstrukce H nejprve přemění hodnoty x_i na hodnoty $1/x_i$, které jsou sčitatelné, a z nich se ve druhém kroku utvoří aritmetický průměr. (Ve třetím kroku se poté z něho utvoří převratná hodnota, tj. užije se inverzní funkce g^{-1} .) Analogicky proto postupujeme při konstruování hledané míry rozptýlenosti a utvořme nejprve metriku $|1/x_i - 1/H|$. Aritmetický průměr těchto hodnot

$$\frac{1}{n} \sum_i \left| \frac{1}{x_i} - \frac{1}{H} \right|$$

by byl obdobou průměrné odchylky (okolo A) a jejich kvadratický průměr

$$\sqrt{\frac{1}{n} \sum_i \left(\frac{1}{x_i} - \frac{1}{H} \right)^2}$$

by byl obdobou směrodatné odchylky (okolo A).

Vyvstává otázka: lze tyto míry *absolutní* variability doplnit na míry *relativní* variability? (Motivací pro tento přechod může být požadavek fyzikální bezrozměrnosti hledané míry variability.) Možné to je, ale musí se tak dít důsledně a do jmenovatele poměrného ukazatele dosadit $1/H$; potom dostaneme (názvy považujeme za provizorní)

harmonickou poměrnou odchylku $\frac{1}{n} \sum_i^n \left| \frac{H}{x_i} - 1 \right|$

harmonický variační koeficient $\sqrt{\frac{1}{n} \sum_i^n \left(\frac{H}{x_i} - 1 \right)^2}$

Opodstatněnost této konstrukce podporuje i srovnání s méně sice známými, ale analogickými vzorci měř relativní variability okolo aritmetického průměru. Platí totiž pro "aritmetickou" poměrnou odchylku

$$R = \frac{\frac{1}{n} \sum_i^n |x_i - A|}{A} = \frac{1}{n} \sum_i^n \left| \frac{x_i}{A} - 1 \right|,$$

pro "aritmetický" variační koeficient

$$V = \frac{\sqrt{\frac{1}{n} \sum_i^n (x_i - A)^2}}{A} = \sqrt{\frac{1}{n} \sum_i^n \left(\frac{x_i}{A} - 1 \right)^2}.$$

V ilustračních příkladech na průměrnou pracnost a průměrnou rychlost dostáváme nejprve, že v obou případech platí (vzhledem k $n = 2$)

$$\left| \frac{1}{x_1} - \frac{1}{H} \right| = \left| \frac{1}{x_2} - \frac{1}{H} \right|,$$

jak jsme požadovali výše (číselné hodnoty jsou $1/12$ v prvním příkladu a $1/150$ ve druhém), a dále že obě relativní míry jsou totožné, tj.

$$\frac{1}{2} \left(\left| \frac{H}{x_1} - 1 \right| + \left| \frac{H}{x_2} - 1 \right| \right) = \sqrt{\frac{1}{2} \left[\left(\frac{H}{x_1} - 1 \right)^2 + \left(\frac{H}{x_2} - 1 \right)^2 \right]}$$

Číselné hodnoty jsou $0,2 = 20\%$, resp. $0,25 = 25\%$. (Pro zajímavost uveďme, že "aritmetické" míry by byly $0,5/2,5 = 0,2$, resp. $10/40 = 0,25$, tedy zcela stejné; při $n > 2$ by výsledky byly ovšem většinou rozdílné, ale ne o mnoho.)

Příklad 4. Uvažujme malý datový soubor 2, 2, 4, 8 a porovnejme absolutní i relativní variabilitu, a to jak okolo aritmetického, tak i harmonického průměru.

Míra variability		Okolo A = 4	Okolo $H^{-1} = 0,34375$
průměrná odchylka	- abs.	2,00	0,15625
	- rel.	50,0 %	45,5 %
směrodatná odchylka	- abs.	2,45	0,16238
	- rel.	61,2 %	47,2 %

Připojme ještě případ *náhodné* proměnné včetně ilustračního příkladu č. 3. Definice rozptylu $D^2(X^{-1})$ by mohla být obdobou druhé mocniny "harmonické" směrodatné odchylky empirických hodnot x_i . Definujme tedy rozptyl invertované proměnné

$$D^2(X^{-1}) = E(X^{-1} - H^{-1})^2 = \int_0^{\infty} \left(\frac{1}{x} - \frac{1}{H} \right)^2 dF,$$

speciálně pak u *diskrétní* náhodné proměnné

$$D^2(X^{-1}) = \sum_x \left(\frac{1}{x} - \frac{1}{H} \right)^2 P(x)$$

a u *spojité* náhodné proměnné

$$D^2(X^{-1}) = \int_0^{\infty} \left(\frac{1}{x} - \frac{1}{H} \right)^2 f(x) dx.$$

Odmocněním dostaneme směrodatné odchylky; jejich vydělením střední hodnotou $E(X^{-1}) = 1/H$ bychom dostali *relativní* míry variability. Aplikujeme-li tyto definice na rozdělení beta, máme (pro $q > 2$)

$$D^2(X^{-1}) = \frac{p(p+q-1)}{(q-1)^2(q-2)} \quad \text{a} \quad \frac{D(X^{-1})}{E(X^{-1})} = \sqrt{\frac{p}{(q-2)(p+q-1)}}$$

Pro zajímavost uvedme hodnoty "aritmetických" měr :

$$D^2(X) = \frac{pq}{(p+q)^2(p+q-1)} \quad \text{a} \quad \frac{D(X)}{E(X)} = \sqrt{\frac{p}{q(p+q+1)}}.$$

Následující tabulka přináší číselné hodnoty obou variačních koeficientů pro vybraná p a q :

Tab.1 Harmonické (první řádek) a aritmetické (druhý řádek) variační koeficienty u rozdělení beta

p	q					
	3	4	5	10	20	...
2	0,707	0,447	0,333	0,151	0,073	0
	0,333	0,267	0,224	0,124	0,066	0
3	0,775	0,500	0,378	0,177	0,087	0
	0,378	0,306	0,258	0,146	0,079	0
4	0,816	0,535	0,408	0,196	0,098	0
	0,408	0,333	0,283	0,163	0,089	0
5	0,845	0,559	0,430	0,211	0,108	0
	0,430	0,354	0,302	0,177	0,098	0
10	0,913	0,620	0,488	0,256	0,138	0
	0,488	0,408	0,354	0,218	0,127	0
...						
∞	1	1	1	1	1	1

Populační a výběrové míry

Navržené míry rozptýlenosti okolo harmonického průměru by měly malý význam, kdyby nenašly uplatnění v induktní statistice. K tomu je ovšem nezbytné prozkoumat jejich výběrová rozdělení, zejména rozdělení harmonického průměru a harmonického rozptylu. V tomto článku se omezíme na zkoumání polohy a měnlivosti těchto rozdělení, tj. na stanovení vhodných charakteristik řečených hlavních vlastností.

Uvažujme konečný základní soubor rozsahu N , z něhož bude proveden náhodný výběr rozsahu n . Harmonické průměry v základním a výběrovém souboru označme H a h , tedy definujeme

$$H = \frac{N}{\sum_{i=1}^N \frac{1}{X_i}}, \quad h = \frac{n}{\sum_{i=1}^n \frac{1}{x_i}}.$$

První otázka zní: je h nestranným odhadem průměru H ? Odpovíme opět otázkou: jakou nestrannost máme na mysli? Víme, že kromě obvyklé

nestrannosti, založené na střední hodnotě $E(\cdot)$, existuje např. mediánová nestrannost, založená na $Me(\cdot)$, a mohou se definovat i jiné. Začneme však onou obvyklou a ptejme se, zda platí $E(h) = H$? Stačilo by sestavit jeden protipříklad, abychom prokázali, že obecně nikoliv. Ostatně již pouhý pohled na dvojici vzorců o několik řádků výše většině čtenářů k této odpovědi nepochybně stačí.

Východisko z této poněkud nepříjemné situace je však opět velmi prosté: očividně totiž platí

$$E\left(\frac{1}{h}\right) = \frac{1}{H} \quad \text{neboli} \quad E^{-1}(h^{-1}) = H.$$

Slovy, "harmonická" střední hodnota výběrových harmonických průměrů je rovna základnímu (populačnímu) harmonickému průměru. Jinak řečeno, h je "harmonicky nestranným" odhadem H . Zavedeme-li pro $E^{-1}(X^{-1})$ označení $E_h(X)$, můžeme psát pro harmonické průměry

$$E_h(X) = \left[E\left(\frac{1}{h}\right) \right]^{-1} = \left(\frac{1}{H} \right)^{-1} = H$$

Ted' se již dá lehce domyslet, jak tomu bude s další charakteristikou výběrového rozdělení — s rozptylem $D^2(h^{-1})$. V předchozí části jsme pro náhodnou proměnnou zavedli rozptyl

$D^2(X^{-1}) = E(X^{-1} - H^{-1})^2$. Dosadíme-li za X statistiku h a provedeme obvyklé úpravy, dostaneme pro výběr bez vracení - z konečného základního souboru o rozsahu N - vztah

$$\begin{aligned} D^2(h^{-1}) &= E(1/h - 1/H)^2 = E\left(\frac{1}{n} \sum_{i=1}^n \frac{1}{x_i} - \frac{1}{N} \sum_{i=1}^N \frac{1}{x_i}\right)^2 = \\ &= \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{N-1} \sum_{i=1}^N \left(\frac{1}{x_i} - \frac{1}{H}\right)^2. \end{aligned}$$

Druhá odmocnina tohoto výsledku představuje směrodatnou chybu odhadu h^{-1} a její podíl se střední hodnotou $E(h^{-1}) = H^{-1}$ dává variační koeficient neboli poměrnou chybu

$$\frac{D(h^{-1})}{E(h^{-1})} = \sqrt{\left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{N-1} \sum_i^N \left(\frac{H}{X_i} - 1\right)^2}.$$

Nestranným výběrovým odhadem rozptylu $D^2(h^{-1})$ je zřejmě

$$\text{est } D^2(h^{-1}) = \left(\frac{1}{n} - \frac{1}{N}\right) \frac{1}{n-1} \sum_i^n \left(\frac{1}{x_i} - \frac{1}{h}\right)^2.$$

Některá rozšíření

Závěrem poukážeme na dvě možná rozšíření nevzržených měř, vycházejících z harmonického průměru. První rozšíření je směrem k vyšším momentům, tedy především ke třetímu momentu jakožto charakteristice šikmosti rozdělení. Základní (populační) třetí centrální moment bude mít v dané situaci tvar

$$\frac{N}{(N-1)(N-2)} \sum_i^N \left(\frac{1}{X_i} - \frac{1}{H}\right)^3$$

a jeho nestranný výběrový odhad bude mít tvar

$$\frac{n}{(n-1)(n-2)} \sum_i^n \left(\frac{1}{x_i} - \frac{1}{h}\right)^3.$$

Bezrozměrný koeficient šikmosti se dostane známým způsobem, tj. jako podíl třetího momentu a třetí mocniny směrodatné odchylky.

Druhé rozšíření spočívá v doplnění dvojic momentů (základních a výběrových) o momenty nadsouboru neboli superpopulace; na základní soubor přitom hledíme jako na náhodný výběr z tohoto nadsouboru a na základní momenty jako na nestranné odhady superpopulačních charakteristik (parametrických funkcí). Celkový přehled všech tří sad přináší tabulka 2 na následující stránce.

Tab. 2 Přehled harmonických charakteristik a jejich středních hodnot

Pomyslný nadsoubor	Základní soubor	Výběrový soubor	Střední hodnoty	
			$\overset{N}{E}(\bullet)$	$E(\bullet)^{**}$
$\omega_1 = \frac{1}{\int_x dF}$	$H_1 = \frac{N}{\sum \frac{1}{x_i}}$	$h_1 = \frac{n}{\sum \frac{1}{x_i}}$	$\overset{N}{E}(h_1) = H_1$	$E(h_1) = E(H_1) = \omega_1$
$\frac{1}{\omega_1} = \int_x \frac{1}{x} dF$	$\frac{1}{H_1} = \frac{1}{N} \sum \frac{1}{x_i}$	$\frac{1}{h_1} = \frac{1}{n} \sum \frac{1}{x_i}$	$\overset{N}{E}\left(\frac{1}{h_1}\right) = \frac{1}{H_1}$	$E\left(\frac{1}{h_1}\right) = E\left(\frac{1}{H_1}\right) = \frac{1}{\omega_1}$
$\frac{1}{\omega_2} = \int \left(\frac{1}{x} - \frac{1}{\omega_1}\right)^2 dF$	$\frac{1}{H_2} = \frac{1}{N-1} \sum \left(\frac{1}{x_i} - \frac{1}{H_1}\right)^2$	$\frac{1}{h_2} = \frac{1}{n-1} \sum \left(\frac{1}{x_i} - \frac{1}{h_1}\right)^2$	$\overset{N}{E}\left(\frac{1}{h_2}\right) = \frac{1}{H_2}$	$E\left(\frac{1}{h_2}\right) = E\left(\frac{1}{H_2}\right) = \frac{1}{\omega_2}$
$\frac{1}{\omega_3} = \int \left(\frac{1}{x} - \frac{1}{\omega_1}\right)^3 dF$	$\frac{1}{H_3} = \frac{1}{(N-1)(N-2)} \sum \left(\frac{1}{x_i} - \frac{1}{H_1}\right)^3$	$\frac{1}{h_3} = \frac{1}{(n-1)(n-2)} \sum \left(\frac{1}{x_i} - \frac{1}{h_1}\right)^3$	$\overset{N}{E}\left(\frac{1}{h_3}\right) = \frac{1}{H_3}$	$E\left(\frac{1}{h_3}\right) = E\left(\frac{1}{H_3}\right) = \frac{1}{\omega_3}$
...	...	...	...	...
$\frac{1}{\omega_r} = \int \left(\frac{1}{x} - \frac{1}{\omega_1}\right)^r dF$	$\frac{1}{H_r}$	$\frac{1}{h_r}$	$\overset{N}{E}\left(\frac{1}{h_r}\right) = \frac{1}{H_r}$	$E\left(\frac{1}{h_r}\right) = E\left(\frac{1}{H_r}\right) = \frac{1}{\omega_r}$

^{**} Zatímco $E(\bullet)$ značí střední hodnotu přes nekonečnou posloupnost výběrových, resp. základních souborů, které by bylo možné náhodným výběrem pořídit z pomyslného nadsouboru nekonečného rozsahu, tak $\overset{N}{E}(\bullet)$ zde značí střední hodnotu přes konečnou množinu $\binom{N}{n}$ výběrových souborů, vznikajících při vybírání bez vracení z konečného základního souboru rozsahu N .

Literatura

- KENDALL, M. G., STUART, A.: *The Advanced Theory of Statistics, Vol. 1: Distribution Theory*. (4. vyd.) London, Griffin 1981
- ČERMÁK, V., KOZÁK, J.: Rozptyl, geometrický rozptyl a míra V_a . *Statistika* (Praha), č. 8-9/1986, s. 374 — 378.
- ČERMÁK, V.: *Diskrétní a spojitá rozdělení — vzorce, grafy, tabulky*. Příručka. Praha, VŠE 1993

Dodatek

Po sepsání tohoto článku a jeho předání k lektorskému posouzení jsme přece jen našli jednu práci, kde byl učiněn pokus o sestrojení míry rozptýlenosti okolo harmonického průměru. Jde o stat' M. Brown, A measure of variability based on the harmonic mean, and its use in approximations. *Annals of Statistics* 13 (1985), 1239 - 1243. Navržená relativní míra má tvar

$$c^2 = 1 - [E(X)E(X^{-1})]^{-1} \quad \text{neboli v naší symbolice} \quad c = \sqrt{1 - \frac{H}{A}}.$$

Například pro rozdělení beta bychom dostali vzorec

$$c = \sqrt{\frac{p}{q(p+q-1)}}.$$

Číselné hodnoty této veličiny jsou zhruba v polovině mezi "aritmetickým" a "harmonickým" variačním koeficientem (viz tab. 1). Avšak např. u rozdělení ve tvaru geometrické posloupnosti Brow-nova míra nedává již tak uspokojivý výsledek: tak třeba pro posloup-nost 2, 4, 8, 16 vycházejí čtverce aritmetického i harmonického variačního koeficientu shodně $23/45 \doteq 0,51$, zatímco $c^2 = 97/225 \doteq 0,43$. Vedle tohoto „protipříkladu“ je zdrojem určitých pochybností i samotná konstrukce (viz vzorec), kde se jeví, že úkol změřit variabilitu hodnot x_i (resp. hodnot $1/x_i$) byl „zaměněn“ za pouhé porovnání průměrů H a A . Celá otázka by však nepochybně zasluhovala další a hlavně důkladnější analýzu.

NCTM standardy jako výzva pro Českou statistickou společnost

Jan Hendl¹

Bez povšimnutí českých pedagogů a matematiků probíhá již dvě desetítky let ve Spojených státech amerických (USA) reforma výuky matematiky na základních a středních školách. Média Jednoty českých matematiků a fyziků zatím nezaznamenaly nic z procesů, které se svým významem podobají reformačnímu úsilí při zavádění konceptů teorie množin do výuky matematiky. V duchu konzervativní tradice v oblasti pedagogiky ani čeští statistici o reformu v USA nemají zájem. Bude se tedy statistika na nižších stupních stále vyučovat stejně jako před 70 lety (1,2) Naši pedagogové se všemožně snaží podávat znalosti o statistice tím nejzkratovějším způsobem, aby ji tak žákům nadobro znechutili. O inovaci pedagogiky statistiky a pravděpodobnosti se hovoří nejčastěji v souvislosti s výukou na vysokých školách, protože se předpokládá, že aspoň vysokoškolák by měl dokázat analyzovat a interpretovat data. Opomíjí se, že stejný předpoklad by se měl uplatnit u každého gramotného člověka. Pár hodin na střední škole, tak i běžné kurzy statistiky na vysoké škole, které jsou často jenom exhibicí matematického aparátu, nemají na statistickou gramotnost větší vliv, spíše utvrzují příslušnou ngramotnost. Nelze se proto divit, že například mladí učitelé po takovém „proškolení“ nejsou schopni prorazit bludný kruh a vše opakují se svými nebohými svěřenci.

Z USA se k nám valí jedna inovace za druhou. Bez osobních počítačů si nedovede mnoho lidí představit svůj život. Mělo by to tak být i se statistikou? Reforma podle NCTM standardů takovému názoru přitakává. A to je právě důvod, proč bychom se jí měli hlouběji zabývat. Seznámení s myšlenkami NCTM standardů může být inspirující pro volbu strategických cílů české obce statistiků.

V dalším textu naznačím podstatu reformy a uvedu v bodech obsah standardů pro výuku analýzy dat a pravděpodobnosti.

¹ UK FTVS, katedra základů humanitních věd a kinantropologie, José Martí 31,
162 52 Praha 6, e-mail HENDL@FTVS.CUNI.CZ

Podstata reformy podle NCTM

Národní rada učitelů matematiky (*National Council of Teachers of Mathematics*, NCTM) je mohutná americká profesní organizace učitelů matematiky v USA a Kanadě, která má dnes přes 100000 členů. Vznikla v roce 1920 transformací Chicagského Matematického Klubu. Jejím úkolem je „to provide the vision and leadership necessary to ensure a mathematics education of the highest quality for all students“. Její doporučení, přestože ne závazná, mají ohromný vliv na vytváření kurikul a konkrétní výuku v jednotlivých státech USA, na uvolňování peněz grantovými agenturami do příslušného výzkumu a inovativních projektů.

Před dvaceti byla na půdě této organizace iniciovaná diskuse o cílech a obsahu výuky matematiky na základních a středních školách.

V roce 1989 NCTM zpracoval standardy matematického kurikula a pro hodnocení studentů (4). Komise pro vytvoření standardů přitom řešila dva hlavní úkoly:

- Vytvořit koherentní představu, co znamená být matematicky gramotný ve světě, který na jedné straně využívá elektronických kalkulačků k provádění matematických operací a v kterém se na druhé straně matematika široce využívá v nejrůznějších aplikovaných oborech.
- Vytvořit standardy, které by v souvislosti s těmito představami usměrňovaly revizi kurikula školní matematiky i hodnocení studentů.

NCTM Standardy z roku 1989, které se týkají výuky od předškolní výuky (Kindergarten) až po dvanáctou třídu (K-12), formulovaly pět obecných cílů pro všechny studenty a studentky:

1. Mají se naučit vážit si matematiky.
2. Mají nabýt sebedůvěry k využívání matematiky.
3. Mají se stát řešiteli matematických problémů.
4. Mají se naučit komunikovat matematickým jazykem.
5. Mají se naučit matematicky uvažovat.

Základní orientace změn, o které NCTM Standardy usilují, se opírají o obecně didaktické úvahy, které jsou aktuální pro současnou diskusi v této oblasti. V souvislosti se statistikou a pravděpodobností jde o tyto aspekty:

- Zvýšení vlastní iniciativy edukanta pomocí zajímavých didaktických experimentů.
- Zdůraznění aplikačního charakteru matematiky a jejího využití v reálném životě, což zahrnuje reálné problémy a práci z vlastními reálnými daty.

- Integraci statistiky, přesněji obhospodařování a analýzy dat, do obecného matematického kurikula již od prvních ročníků školy.
- Odklon od klasické inferenční statistiky jako rituálu metod směrem k explorační analýze z důrazem na aktivní, experimentující a explorační aktivity.
- Odklon od výuky pravděpodobnosti pomocí urnových schémat a kombinatoriky k explorativnímu přezkušování reálných dat a simulačních scénářů.

V roce 2000 se objevil nový dokument *Principles and Standards for School Mathematics*, upravená forma standardů, která obsahuje některé inovace a zohledňuje kritické připomínky k verzi standardů z roku 1989 (7). Nový text Standardů opakuje v modifikované podobě volání po reformě výuky matematiky. Zjednodušuje prezentaci požadavků, jednotlivé principy se vyjasňují pomocí příkladů problémů, studentských úkolů a modelových dialogů ve třídě.

Knižní publikace NCTM Standardů 2000 má 402 stránek a její obsah je organizována do osmi kapitol. Předmluva a první kapitola uvádějí účel a všeobecný záměr Standardů. V druhé kapitole se probírá šest principů, které tvoří předpoklady pro implementaci dokumentu. Princip *rovnosti* vyjadřuje požadavek, že všichni studenti mají mít možnost přístupu ke stejně náročné a hodnotné výuce matematiky. Principy *kurikula, vyučování a učení* ozřejmují potřebu koherentního programu ve všech věkových kategoriích, který poskytne studentům znalosti významných částí matematiky, přičemž se bude vycházet z integrace mimoškolních znalostí a zkušeností studentů se zkušenostmi získávanými ve třídě. Princip *hodnocení* předpokládá takový způsob evaluace, který bude podporovat výuku a proces učení. Standardy předpokládají ústup od testování pomocí zaškrťování volitelných odpovědí. Řešené úlohy mají vycházet z autentických problémů. Hodnocení si všímá správného matematického myšlení, vyjadřování a konečného výsledku. Konečně princip *technologie* zdůrazňuje důležitost a vliv technologie na výuku a učení matematiky.

Ve třetí kapitole se na obecné úrovni popisuje deset standardů matematiky, které mají studenti zvládnout v průběhu školní docházky. Prvních pět se považuje za obsahové standardy, protože se týkají základních témat školní matematiky. Nazývají se obsahové linie (*strands*):

- | | |
|---------------------------|---|
| 1. Čísla a jejich spojení | 4. Míry a měření |
| 2. Funkce a algebra | 5. Analýza dat, statistika a pravděpodobnost. |
| 3. Geometrie | |

Poznamenejme, že na rozdíl od poslední verze Standardy 1989 pracovaly s obsahovými liniemi Číslo a operace, Konfigurace a funkce, Algebra, Statistika, Geometrie, Diskrétní matematika.

Jednotlivé obsahové linie nepředstavují oddělené disciplíny, vzájemně se mají podporovat, při řešení problému se používají příslušné prostředky jednotlivých obsahových linií integrovaným způsobem. V této kapitole se probírají také základní procesy, které mají studenti zvládnout, když pracují v rámci dané obsahové linie. V pěti standardech se pojednává o podobě řešení problémů nebo dokazování a odvozování a o celkovém stylu práce a matematické argumentaci. Studenti se mají naučit vyjadřovat pomocí matematického jazyka a vytvářet spojení mezi různými oblastmi uvnitř a vně matematiky. Tato část dokumentu je tvořena odstavci:

1. Matematika jako řešení problémů
2. Matematika jako racionální zdůvodnění a dokazování
3. Matematika jako komunikace
4. Matematika jako stavba propojených znalostí
5. Matematika jako reprezentace znalostí

Implementace těchto standardů tvoří kouzlo NCTM doporučení a je také největším zdrojem diskusí v komunitě matematiků a pedagogů. Obsahové linie nevyvolávají tolik sporů jako realizace výukového procesu. Některé státy Unie (např. Kalifornie) kvůli nim reformu zcela zamítly. Na tomto místě se jimi nebudeme hlouběji zabývat. Pouze ocitujeme výňatek z vyjádření odpůrců reformy ze skupiny Mathematically Correct, který charakterizuje jejich pocity:

„V zemi se změnil způsob vyučování i úprava učebnic. Výuka je často organizována v malých skupinách, kde si studenti kladou navzájem otázky a učitel se nemá míchat do jejich diskuse. Při výuce algebry se používají se kostky a jiné manipulativní objekty jako v mateřské školce. Základní procedurální dovednosti se procvičují čím dál tím méně. Kalkulačky se využívají při každé příležitosti a studenti se proto nemusí učit provádět výpočetní operace bez nich. Učebnice, pokud je studenti mají, jsou plné barevných obrázků a historek, ale ne plné matematiky. Neobsahují často explicitní definice nebo procedury. Je to proto, aby studenti objevovali celou matematiku sami pro sebe. Manuální dělení se ztrácuje a když se jim to podaří, tak odbourají i násobení.“

Obsah NCTM Standardů dále pokračuje čtyřmi kapitolami, které se rozpracovávají doporučení pro jednotlivé věkové kategorie: předškolní výuka až 2. třída, třídy 3 - 5, třídy 6 - 8, třídy 9 - 12. Každá z těchto kapitol obsahuje podrobný výklad k jednotlivým obsahovým liniím i

procesům. Také se uvádějí návrhy, jak je možné doporučení aplikovat v rámci výuky ve třídě. Pro každý z obsahových standardů se uvádějí specifická „očekávání“ znalostí a dovedností. Závěrečná kapitola diskutuje, jak jednotlivé skupiny obyvatelstva (učitelé, školní úředníci, politikové, vědci, studenti, rodičové) mohou spolupracovat na zlepšování matematické výuky.

Analýza dat a pravděpodobnost

NCTM Standardy pro analýzu data a počet pravděpodobnost doporučují, aby žákyně a žáci se zabývali otázkami, které je možné zodpovědět pomocí dat, sběrem dat a jejich smysluplným zpracováním. Žákyně a žáci se mají učit, jak se data shromažďují, analyzují a jak je možné je pomocí grafů a obrázků znázorňovat tak, že to umožní zodpovědět položené otázky. Tento standard také zahrnuje zvládnutí některých metod numerické analýzy dat a inferenční statistiky. Jeho obsahem jsou také základní koncepty a použití počtu pravděpodobnosti, přičemž se zvláště poukazuje na vztah mezi počtem pravděpodobnosti a statistikou.

Základní cíle výuky pro obsahovou linii Analýza dat, statistika a pravděpodobnost NCTM Standardy definují jako získání schopností:

1. *Formulovat otázky, které je možné zodpovědět pomocí dat a k tomu účelu zvládnout sběr, zpracování a prezentaci relevantních údajů*
2. *Zvolit a použít vhodné statistické metody pro analýzu dat*
3. *Navrhnout a vyhodnotit závěry (inference) a predikce pomocí dat*
4. *Pochopit a použít základní pojmy teorie pravděpodobnosti*

Ve zdůvodnění k tomuto standardu se uvádí, že v informační společnosti musí patřit mezi cíle školního vzdělání připravit studenty na život ve světě, který je stále více ovlivněn informačními technologiemi, analýzou empirických dat a přívalem statistických údajů z médií. Edukanti se mají postupně během školní docházky seznamovat s celou škálou statistických prostředků a konceptů, které by jim umožnili lépe se rozhodovat v praktickém životě.

„Aby porozuměli základům statistických myšlenek, musejí žákyně a žáci v průběhu školní docházky poznávat nové ideje a postupy místo toho, aby se stále vraceli ke stejným tématům a aktivitám. Tato obsahová linie o datech a statistice umožňuje vyučujícím, žákyním a žákům vytvořit celou řadu spojení mezi matematickým koncepty a postupy v jiných

oblastech jako např. k pojmu čísla, k otázkám měření a k problémům z algebry a geometrie. Práce v oblasti analýzy dat a počtu pravděpodobnosti poskytuje žákyním a žákům přirozené možnosti spojit a propojit matematické znalosti s ostatními školními předměty a také se svými zkušenostmi z každodenního života.

Navíc jsou procesy, které se používají při argumentaci s daty a statistikou, velmi užitečné ve světě práce a v životě. Něco, co poznávají děti ve škole, se jim bude zdát předem dané a určené pravidly. Při studiu dat a statistiky se ale také naučí, že řešení mnohých problémů závisí na nějakých předpokladech a skrývá v sobě určitou nejistotu.“

Uvedeme podrobné cíle standardů pro obsahovou linii o analýze dat a pravděpodobnosti v bodech, kdy pro každý hlavní cíl jsou definovány požadavky pro jednotlivé věkové kategorie K -12. Tím budou patrnější vývojové rysy budoucího kurikula.

Cíl 1: Formulovat otázky, které je možné zodpovědět pomocí dat a k tomu účelu zvládnout sběr, zpracování a prezentaci relevantních údajů

Cíle pro věkovou kategorii žáků a žákyní v předškolní výchově až po 2. třídu

Všichni žáci a žákyně by měli

- klást otázky a shromažďovat data o sobě a svém okolí,
- třídit a klasifikovat objekty podle jejich vlastností a zpracovávat údaje o objektech,
- prezentovat údaje užitím konkrétních objektů, obrázků a grafů.

Cíle pro věkovou kategorii v 3. až 5. třídě

Všichni žáci a žákyně by měli

- navrhnout pozorování s cílem vyřešit určitou otázku a posuzovat jak metody sběru dat ovlivňují vlastnosti získaných údajů,
- sbírat data pomocí pozorování, šetření a experimentu,
- zobrazovat data pomocí tabulek a grafů (čárové vývojové grafy a sloupkové grafy),
- rozpoznávat rozdíly v prezentaci kategoriálních a číselných údajů.

Cíle pro věkovou kategorii v 6. až 8. třídě

Všichni žáci a žákyně by měli

- formulovat otázky, navrhnout studie a shromažďovat data o jedné charakteristice ve dvou populacích nebo o různých charakteristikách v jedné populaci,
- volit, vytvářet a používat vhodné grafické zobrazení dat (histogram, krabíčkový graf, shlukový graf).

Cíle pro věkovou kategorii v 9. až 12. třídě

Všichni žáci a žákyně by měli

- rozumět rozdílům mezi různými typy zkoumání a typům inference, které jim odpovídají,
- znát charakteristiky dobře navržených studií, včetně roli randomizace v šetření a experimentu,
- porozumět významu číselných a kategoriálních údajů, jednorozměrných a dvourozměrných dat a pojmu proměnná,
- porozumět pojmu histogram, paralelní krabíčkový graf, shlukový graf a používat tyto grafy při zobrazení dat,
- počítat základní statistiky a porozumět rozdílům mezi statistikou a parametrem.

Cíl 2: Zvolit a použít vhodné statistické metody pro analýzu dat

Cíle pro věkovou kategorii žáků a žákyň v předškolní výchově až po 2. ročník

Všichni žáci a žákyně by měli

- popisovat části dat a množinu dat jako celek s cílem ukázat, co data vypovídají

Cíle pro věkovou kategorii v 3. až 5. třídě

Všichni žáci a žákyně by měli

- popisovat tvar a důležité vlastnosti množiny údajů a porovnávat množiny údajů s důrazem na to, jak jsou data rozloženy,
- používat míry centrální tendence s důrazem na medián a porozumět, co tyto míry znamenají a co nevypovídají o datech,
- srovnávat různé znázornění stejných dat a vyhodnocovat, jak každé znázornění ukazuje důležité vlastnosti dat.

Cíle pro věkovou kategorii v 6. až 8. třídě

Všichni žáci a žákyně by měli

- volit, používat a interpretovat míry centrální tendence a rozptylu, včetně průměru a interkvartilového rozpětí,
- diskutovat vztah mezi množinou dat a její grafickou reprezentací, zvláště u histogramu, grafu lodyhy a listu, krabíčkového grafu a shlukového grafu.

Cíle pro věkovou kategorii v 9. až 12. třídě

Všichni žáci a žákyně by měli

- být schopni pro jednorozměrná data znázornit jejich rozložení, popsat jejich tvar a zvolit a vypočítat sumární statistiky,
- být schopni pro dvojrozměrná data sestavit shlukový graf, popsat jeho tvar a určit regresní koeficienty, regresní rovnici a korelační koeficient pomocí kalkulatoru,
- zobrazit a diskutovat dvojrozměrná data, kde aspoň jedna proměnná je kategoriální,
- rozpoznat, jak lineární transformace jednorozměrných dat ovlivňuje jejich tvar, průměr a rozptyl,
- identifikovat trend pro dvojrozměrná data a nalézt funkci, která vztah modeluje nebo dokázat data transformovat, tak aby je bylo možné modelovat.

Cíl 3: Navrhnout a vyhodnotit závěry (inference) a predikce pomocí data

Cíle pro věkovou kategorii žáků a žákyní v předškolní výchově až po 2. ročník

Všichni žáci a žákyně by měli

- diskutovat jevy ze zkušenosti žáků a žákyní s ohledem na jejich možnost nebo nemožnost výskytu za daných okolností.

Cíle pro věkovou kategorii v 3. až 5. třídě

Všichni žáci a žákyně by měli

- navrhovat a zdůvodnit závěry a predikce, které jsou založeny na datech a navrhnout zkoumání k dalšímu sledování závěrů a predikcí.

Cíle pro věkovou kategorii v 6. až 8. třídě

Všichni žáci a žákyně by měli

- posuzovat data dvou nebo více výběrů a provádět závěry o populacích, z kterých data pocházejí,
- dělat závěry o možných vztazích mezi dvěma charakteristikami výběrových jednotek na základě shlukového grafu a přibližného proložení přímkou,
- využít předběžné závěry k formulování nových otázek a plánování nových zkoumání.

Cíle pro věkovou kategorii v 9. až 12. třídě

Všichni žáci a žákyně by měli

- používat simulace ke zkoumání variability výběrových statistik ze známé populace a konstruovat jejich výběrová rozložení,
- porozumět tomu, jak výběrová statistika odráží hodnotu populačního parametru a používat výběrové rozložení jako základ pro neformální inferenci,
- vyhodnocovat publikované statistické výsledky zkoumáním plánu studie, vhodnosti použité analýzy a validity závěrů,
- rozumět, jak základní statistické techniky se používají k monitorování charakteristik procesu v reálném světě.

Cíl 4: Pochopit a použít základní pojmy teorie pravděpodobnosti

Cíle pro věkovou kategorii žáků a žákyní v předškolní výchově až po 2. ročník

(neformulováno)

Cíle pro věkovou kategorii v 3. až 5. třídě

Všichni žáci a žákyně by měli

- popsat jevy pomocí pojmů pravděpodobné a nepravděpodobné a diskutovat stupeň pravděpodobnosti slovy jako jistý, stejně pravděpodobný a nemožný,
- predikovat pravděpodobnost výsledků jednoduchých experimentů a posuzovat predikce,
- porozumět, že míra pravděpodobnosti nějakého jevu může být zachycena číslem mezi nulou a jedničkou.

Cíle pro věkovou kategorii v 6. až 8. třídě

Všichni žáci a žákyně by měli

- porozumět a používat vhodnou terminologii k popsaní doplňkových a vzájemně se vylučujících jevů,
- používat relativní četnost a základní poznatky o pravděpodobnosti, aby mohli provádět a testovat úvahy o výsledcích experimentů a simulací,
- počítat pravděpodobnosti pro jednoduché složené jevy použitím metod jako seřazený seznam, stromový diagram a plošné modely.

Cíle pro věkovou kategorii v 9. až 12. třídě

Všichni žáci a žákyně by měli

- porozumět pojmem jako výběrový prostor a pravděpodobnostní rozložení a v jednoduchých případech konstruovat výběrový prostor a rozložení,
- používat simulace ke konstrukci empirického pravděpodobnostního rozložení,
- počítat a interpretovat očekávanou hodnotu náhodné proměnné v jednoduchých případech,
- porozumět pojmu podmíněné pravděpodobnosti a nezávislých jevů,
- porozumět, jak počítat pravděpodobnost složených jevů.

Závěrečné poznámky

1. V diskusi o zlepšování výuky základů statistiky na vysokých školách se uvádí několik možných přístupů:

- zvýšení hodinové dotace,
- zlepšení didaktiky,

- přesun části výuky do předmětných disciplín s cílem provázat matematicko-statistické postupy s výzkumnou praxí v daném oboru.

Nehovoří se o možnosti provést "kopernikovský obrat" a převést část této výuky na základní a střední školy. Logicky totiž plyne, že za 12 let výuky na těchto stupních se lze naučit základům statistiky a citu pro statistické problémy jistě lépe než za semestr na vysoké škole. NCTM Standardy se pouští tímto směrem. Jedním z hlavních argumentů je úvaha, že jedinec s maturitou, aby mohl fungovat jako řádný občan a volič, musí být statisticky gramotný.

2. Někdy se namítá, že realita na základních a středních školách v USA je horší než na evropských školách, tudíž, že sotva se můžeme co poučit. Z toho ale neplyne, že dokument o NCTM Standardech nepředstavuje inspirativní mezník v nazírání na výuku matematiky. Důkladný proces vývoje a schvalování dokumentu a podíl mnoha stovek pracovníků na jeho vzniku dává zároveň záruku, že tento dokument má mnohem lepší obsah ve srovnání se všemi podobnými dokumenty vytvořenými v Česku.
3. Podobně jako naše standardy výuky matematiky na základních a středních školách se nezrodily americké standardy na základě ministerského příkazu, ale vytvořily se spontánně na půdě NCTM. Byly výsledkem participativního procesu všech zúčastněných. Vytvořené dokumenty slouží jako přesvědčivý důkaz fungování demokracie.
4. Ignorance našich matematiků a pedagogů je o to nepochopitelnější, že na rozdíl od dřívějších let existuje za podpory elektronických médií svobodná výměna informací bez zábrán. Většina dokumentů o celém procesu vytváření a kritické diskusi o NCTM Standardech (3-7) je k dispozici na Internetu. Stačí do prohlédáváče napsat „NCTM Standards“.

Literatura

- Calda, E., Dupač, V.: *Matematika pro gymnázia – Kombinatorika, pravděpodobnost, statistika*. Prometheus, 1993
- Fuchs, E., Hrubý, D. a kol.: *Standardy a testové úkoly z matematiky pro základní školy a nižší ročníky gymnázií*. Prometheus, 2000.
- NCTM: *An agenda for action recommendations for school mathematics*. Reston, Va.: National Council of Teachers of Mathematics, 1980.
- NCTM: *Curriculum and Evaluation Standards for School Mathematics*. Reston, Va.: National Council of Teachers of Mathematics, 1989.
- NCTM: *Professional Standards for Teaching Mathematics*. Reston, Va.: National Council of Teachers of Mathematics, 1991.
- NCTM: *Assessment Standards for School Mathematics*. Reston, Va.: National Council of Teachers of Mathematics, 1995.
- NCTM: *Principles and Standards for School Mathematics*. Reston, Va.: National Council of Teachers of Mathematics, 2000 (standards-e.nctm.org).

Internetové stránky k tématu NCMT Standardy

Kromě základních odkazů (standards-e.nctm.org a www.nctm.org) skupina Mathematics Forum (forum.swarthmore.edu) uvádí mnoho spojů na zdroje o matematickém vyučování včetně stránek skupin proti reformě. Organizace National Science Foundation (NSF) založila čtyři centra pro podporu tvorby kurikul podle reformy NCTM:

Centrum pro všechny stupně K-12 na stránce www.edc.org/mcc

Centrum pro základní stupeň na místě www.arccenter.comap.com

Centrum pro střední stupeň na místě <http://shomecenter.missouri.edu>

Centrum pro střední školy na stránce www.ithaca.edu/compass

Skupina zaměřená proti reformě vystavuje na stránce www.mathematicallycorrect.com.

Objektivní příčiny nespokojenosti v divadelních souborech

Josef Tvrdlík, OU

Na oslavu padesátin našeho kamaráda-učitele napsal jiný kamarád, právník, veršovanou jednoaktovku nazvanou Komenského hledání. Ve hře vystupuje sedm postav a jejich obsazení vzniklo vcelku spontánně. Roli nejhlavnější samozřejmě uchvátil autor a současně režisér pro sebe. Jelikož jsem dostal jednu z pěti vedlejších (ale významných) rolí, považoval jsem obsazení za správné. Narozeninová premiéra měla velmi příznivý ohlas. Celkem pochopitelně mezi neúčinkujícími pak zazněly poznámky o jejich nespravedlivém opomenutí. Proč zrovna oněch sedm bylo obsazeno, když kmenový stav potenciálních herců byl 22? Jistě bylo možné i obsazení jiné, které by oni považovali za lepší. Těch jiných obsazení je opravdu mnoho. Hra nevyžaduje, aby dámské role hrály dámy a pánské role pánové, dokonce ani při premiéře to tak nebylo, takže sedmička herců mohla být vybrána $\binom{22}{7} = 170\,544$ způsoby. Z těchto sedmic lze

pak sestavit. $\binom{22}{7} 7! = 859\,541\,760$ různých obsazení.

V opravdových divadelních souborech je situace pro spokojenost členů souboru většinou trochu příznivější. Herci i role jsou rozděleni do kategorií podle věku, pohlaví atd. Uvažujme, že je k takových kategorií, v kategorii je m_i rolí a soubor disponuje n_i herci v kategorii, $0 \leq m_i \leq n_i$, $i=1, 2, \dots, k$. Počet možných obsazení N je pak

$$N = \prod_{i=1}^k \binom{m_i}{n_i}.$$

Vezmeme-li dva docela realistické příklady s číselnými hodnotami z následující tabulky, dostaneme $N_1 = 2\,592$, resp. $N_2 = 54\,000$ možných obsazení.

i	1	2	3	4	Příklad
m_i	2	3	1	1	1,2
n_i	4	4	3	3	1
n_i	6	5	6	5	2

Naděje na spokojenost snad má divadlo jednoho herce, soubory s jedním hercem v každé kategorii a tomu přizpůsobeným repertoárem nebo možná soubory nehrající vůbec nic. V ostatních případech je často mnoho možných obsazení a mnoho nespokojených herců.

ROBUST'2002

Jaromír Antoch, Gejza Dohnal

Ve dnech 21.-25.1.2002 se konala další z řady škol Jednoty českých matematiků a fyziků (nejen) o robustní statistice, tentokrát s přívlastkem „zimní“. Ani tato, v pořadí už dvanáctá, nezůstala nic dlužna svému podrouškou statistiky ukrytému vlastivědnému poslání. Tentokrát se konala v malebném prostředí Frýdlantského výběžku, pod hřebeny Jizerských hor v místě zvaném Hejnice. Podle legendy zde ve 12. století měl chudý řemeslník beznadějně nemocnou ženu a dítě, které byly zázračně uzdraveny na přimluvu Panny Marie. Jako výraz poděkování připevnil tento řemeslník milostnou sošku na Strom staré lípy. Tak vzniklo jedno z nejznámějších poutních míst, k němuž se každoročně vydávali poutníci z blízkého okolí z Čech, ze Slezska, Lužice, Polska a Saska. V místě byla postavena kaplička, později gotický kostel a v 18. století barokní chrám, jemuž dala jméno právě ona soška – „*Mater formosa*“ (t.j. „*Matka spanilá*“). Od konce 17. století zde stojí františkánský klášter, jehož obyvatelé

se starali o statisíce poutníků udílením svátosti a duchovní útěchy. Františkáni tu už nejsou, nicméně klášter znovu slouží v celé své kráse jako „Mezinárodní centrum duchovní obnovy“. A právě zde probíhala 12. zimní škola JČMF Robust'2002.

Akci připravil organizační výbor ve složení doc. RNDr J. Antoch, CSc (MFF UK Praha) - předseda, doc. RNDr M. Brzezina, CSc (PedF TUL) a doc. RNDr G. Dohnal, CSc. (FS ČVUT Praha). Konference proběhla pod záštitou MVS JČMF a za pomoci KPMS MFF UK, PedF TUL a České statistické společnosti. Účastníky konference přijel mj. pozdravit rektor TUL prof. RNDr D. Lukáš, CSc. a slavnostního zahájení se zúčastnili prorektor TUL prof. Ing. Jiří Militký, CSc, proděkan PedF TUL doc. RNDr J. Vild, prezident Mezinárodní asociace pro výpočetní statistiku (IASC) prof. dr. J. Hinde z Univerzity v Exeteru, Velká Británie, a prezident-elect IASC prof. dr. G. Galmacci z Univerzity v Perugii, Itálie.

Zájem o Robust neustále roste. Letos se jej zúčastnilo okolo 90 účastníků, z toho sedm zahraničních; dva z Německa, dva ze Slovenska a po jednom z Belgie, Velké Británie a Itálie. Předneseno bylo okolo padesáti příspěvků. Zvané přednášky přednesli (v abecedním pořadí) prof. RNDr P. Hájek, DrSc, ÚI ČAV, kolektiv pracovníků Státního zdravotního ústavu pod vedením RNDr M. Malého, CSc, doc. RNDr Z. Strakoš, DrSc, ÚI ČAV a prof. RNDr J. Štěpán, DrSc, MFF UK.

Již tradičně byl jeden večer věnován programovému vybavení pro (statistickou) analýzu dat, jehož se zúčastnili zástupci firem Elkan, MDTex, Trilobyte a StatSoft. Dále byla připravena beseda se zástupci IASC o horkých tématech výpočetní statistiky ve světě a u nás.

Paralelně s Robustem'2002 se uskutečnil 1st Czech – German Seminar on Semiparametric Models se zaměřením na semiparametrické modelování, především pak v oblasti ekonometrie. Tento seminář připravila doc. RNDr J. Antoch, CSc (MFF UK Praha) a Mgr. L. Čížková, Dr (Humboldtova Univerzita v Berlíně).

Pro organizátory bylo velkým potěšením, že se akce zúčastnilo 24 doktorandů a studentů. Jejich osobní údaje v databázi účastníků pozitivně přispějí k podpoře hypotézy, vyslovené před dvěma lety kolegou Tvrdíkem o rychlosti stárnutí účastníků této konference (viz IB č.1, ročník 11, 2000, str. 11). Vystoupení doktorandů a studentů byl věnován páteční celodenní program. Odborná porota ve složení prof. RNDr. J. Štěpán, DrSc (předseda), doc. RNDr J. Antoch, CSc, doc. RNDr J. Hurt, CSc, doc. RNDr Z. Prášková, CSc. a RNDr. I. Saxl, CSc, vybrala jako nejlepší vystoupení (a práci) s výpočetním zaměřením přednášku Mgr.

M. Betince (MFF UK). Oceněný získal program Mathematica, který věnovala fi Elkan, Praha.

Kulturní a vlastivědný program zahrnoval několik významných akcí. Především to byl varhanní koncert z díla Bachova, Brahmsa a Janáčkovy v chrámu Navštívení Panny Marie, který připravil a zahrál prof. RNDr P. Hájek, DrSc. Dále to byl zajímavý večer věnovaný historii Frýdlandska, který připravil RNDr I. Saxl, DrSc., seznámení se s historií kláštera a řádu františkánů, v podání Mgr. M. Betince a pátera Ing. ThDr M. Rabana. Páter Raban nás také provedl všemi zákoutími (a podzemími) kláštera a chrámu. Počasí po dobu Robustu nebylo zrovna ideální, nicméně střední výlet do okolí (do nedalekých lázní Libverda a na hřebeny Jizerek) se vydařil. Nicméně krásy Jizerských hor a jejich rašeliníšť jsme mohli obdivovat večer z diapoзитivů a vyprávění dlouholetého člena horské služby pana Jecha.

Zimní ROBUST'2002 je již za námi, ať žije letní ROBUST'2004! Kam nás tentokrát zavede?

Pedagogický software 2002

5. - 6. června 2002 České Budějovice - Čtyři Dvory areál Jihočeské univerzity a AV ČR

Jihočeská univerzita v Českých Budějovicích si Vás dovoluje pozvat k aktivnímu vystoupení na 8. ročníku mezinárodní konference *Pedagogický software 2002*. První informace včetně elektronické přihlášky najdete na <http://www.zf.jcu.cz/aktuality/ps2002/1inf>. Konference je již tradičně orientována na široké spektrum problémů souvisejících s modernizací, racionalizací a zvyšováním efektivnosti výuky jednotlivých disciplín. Klade si proto za cíl vytvořit prostor pro výměnu zkušeností a získání nových informací o způsobech výuky jednotlivých disciplín a přispět tak k hledání optimálních forem výuky a jejich racionalizace s důrazem na výuku elektronickou, využití výpočetní techniky, informačních systémů, internetu a dalších progresivních metod pedagogické práce.

<i>Václav Čermák, Čím měřit variabilitu okolo harmonického průměru.....</i>	1
<i>Jan Hendl, NCTM standardy jako výzva pro Českou statistickou společnost</i>	13
<i>Josef Tvrdlík, Objektivní příčiny nespokojenosti v divadelních souborech</i>	24
<i>Jaromír Antoch, Gejza Dohnal, ROBUST'2002</i>	25

Informační Bulletin České statistické společnosti vychází čtyřikrát do roka v českém vydání. Předseda společnosti: Doc. RNDr. Jaromír Antoch, CSc., KPMS MFF UK Praha, Sokolovská 83, 186 75 Praha 8, e-mail: jaromir.antoch@kartin.mff.cuni.cz.
ISSN 1210 – 8022
Redakce: Doc. RNDr. Gejza Dohnal, CSc., Jeronýmova 7, 130 00 Praha 3, e-mail: dohnal@fsik.cvut.cz