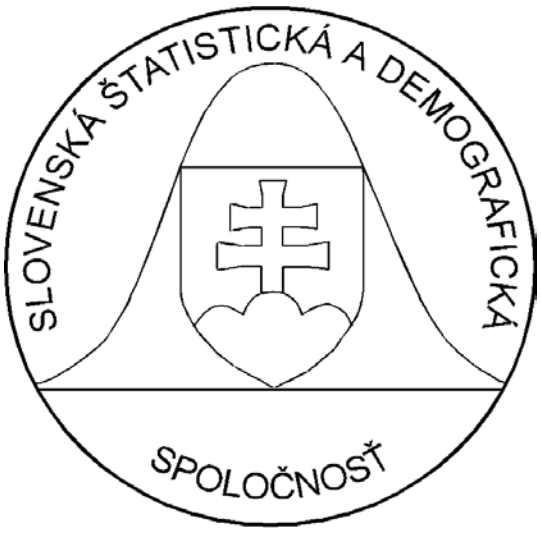
2/2007

FORUM STATISTICUM SLOVACUM



ISSN 1336-7420
72
9 771336 742001

Slovenská štatistická a demografická
spoločnosť
Miletičova 3, 824 67 Bratislava
www.ssds.sk

Naše najbližšie akcie:

(pozri tiež www.ssds.sk, blok Poriadané akcie)

11. SLOVENSKÁ DEMOGRAFICKÁ KONFERENCIA,
tematické zameranie: *Migrácia*
17. – 19. 9. 2007, hotel Čingov, Slovenský raj

FernStat 2007
IV. medzinárodná konferencia aplikovanej štatistiky
(**F**inancie, **E**konomika, **R**iadenie, **N**ázory)
tematické zameranie: *Aplikovaná, demografická, matematická štatistika, štatistické riadenie kvality*.
4. – 5. 10. 2007, hotel Lesák, Tajov pri Banskej Bystrici

16. Medzinárodný seminár VÝPOČTOVÁ ŠTATISTIKA,
6. – 7. 12. 2007, Bratislava

Prehliadka prác mladých štatistikov a demografov
6. 12. 2007, Bratislava

SLÁVNOSTNÁ KONFERENCIA 40 ROKOV SŠDS,
27. 3. 2008, Bratislava

KONFERENCIA POHĽADY NA EKONOMIKU SLOVENSKA 2008,
tematické zameranie: *Vývoj HDP a dlhodobej nezamestnanosti*
15. 4. 2008, Bratislava

EKOMSTAT 2008, 22. škola štatistiky
Tematické zameranie: *Štatistické metódy v praxi*
1. – 6. 6. 2008, Trenčianske Teplice

14. SLOVENSKÁ ŠTATISTICKÁ KONFERENCIA,
tematické zameranie: *Regionálna štatistika*
rok 2008, Žilinský kraj

12. SLOVENSKÁ DEMOGRAFICKÁ KONFERENCIA,
tematické zameranie: *Využitie GIS v demografii*
rok 2009, Trenčiansky kraj

ÚVOD

Vážené kolegyne, vážení kolegovia,

druhé číslo tretieho ročníka časopisu, ktorý vydáva Slovenská štatistická a demografická spoločnosť (SŠDS) je zostavené z príspevkov, ktoré autori pripravili pre konferenciu PRASTAN 2007, ktorá sa uskutočnila v dňoch 4. 6. - 8. 6. 2007 v Banskej Bystrici na Právnickej fakulte Univerzity Mateja Bela. Konferenciu organizovala Slovenská štatistická a demografická spoločnosť v spolupráci so Stavebnou a Strojníckou fakultou STU, Fakultou managementu UK, Fakultou prírodných vied UMB a Štatistickým úradom SR .

Programový a organizačný výbor pracoval v zložení: Martin Kalina, Oľga Nánásiová – predsedovia, členovia: Mária Bohdalová, Michal Greguš, Angela Handlovičová, Jozef Chajdiak, Ivan Janiga, Jozef Komorník, Magda Komorníková, Ján Luha, Alžbeta Michalíková, Mária Minárová, Roman Nedela, Veronika Valenčáková, Beloslav Riečan, Iveta Stankovičová, Jiří Vala, Jan Ámos Víšek, Gejza Wimmer.

Na príprave a zostavení tohto čísla participovali: Mária Bohdalová, Eva Drobná, Angela Handlovičová, Ivan Janiga, Jana Kalická, Martin Kalina, Mária Minárová, Oľga Nánásiová, Iveta Stankovičová,

Editori tohto čísla ďakujú recenzentom za rýchlu a kvalitnú prácu.

Výbor SŠDS

FORUM STATISTICUM SLOVACUM

Vydavateľ
Slovenská štatistická a demografická spoločnosť
Miletičova 3
824 67 Bratislava 24
Slovenská republika

Redakcia
Miletičova 3
824 67 Bratislava 24
Slovenská republika

Fax
02/63812565

e-mail
chajdiak@statis.biz
Jan.Luha@statistics.sk

Registráciu vykonalo
Ministerstvo kultúry Slovenskej republiky

Registračné číslo
3416/2005

Tematická skupina
B1

Dátum registrácie
22. 7. 2005

Objednávky
Slovenská štatistická a demografická spoločnosť
Miletičova 3, 824 67 Bratislava 24
Slovenská republika
IČO: 178764
Číslo účtu: 0011469672/0900

ISSN 1336-7420

Ročník
III.

Číslo
2/2007

Cena výtlačku 500 SKK / 20 EUR
Ročné predplatné 1500 SKK / 60 EUR

OBSAH

Analýza a modelování životní úrovně

Jitka Bartošová

Fakulta managementu

Vysoká škola ekonomická v Praze

Jarošovská 1117/II

377 01 Jindřichův Hradec

bartosov@fm.vse.cz

Abstrakt

K měření životní úrovně obyvatelstva se často používá hodnota čistého příjmu po srážce daně. Příjem lze po vhodných úpravách interpretovat rovněž jako míru produktivity práce a prosperity společnosti. Proto je zkoumání stavu a vývoje příjmů jedinců a domácností v popředí zájmu ekonomů všech vyspělých zemí světa. Důležitým nástrojem tohoto výzkumu jsou pravděpodobnostní modely. Pravděpodobnostní modely rozdělení příjmů umožňují zhodnocení životní úrovně všech obyvatel státu bez rozdílu, stejně jako srovnání životní úrovně příslušníků různých společenských skupin nebo obyvatel různých regionů. Jsou rovněž ukazatelem relativní životní úrovně obyvatelstva státu ve srovnání s jinými státy. Nejčastěji používaným teoretickým modelem rozdělení příjmů je logaritmicko-normální rozdělení (především jeho tříparametrická varianta). Tento článek se zaměřuje na možnost použití tohoto modelu rozdělení příjmů v ČR a SR v r. 2002.

Klíčová slova: logaritmicko-normální model, rozdělení příjmů, test shody rozdělení

1. Úvod

Zjišťování životní úrovně a sociální situace jedinců a domácností je věnováno množství publikací v odborných a vědeckých časopisech a sbornících (domácích i zahraničních) zaměřených na ekonomii a statistiku. Mnohé z nich obsahují výsledky kvalitních statistických analýz dostupných datových zdrojů a současně i cenné informace o použitých metodách a možnostech jejich aplikace při dalších podobných analýzách.

Současný světový výzkum se soustřeďuje především na analyzování dynamiky vývoje příjmů, jeho stabilitu a na odhalování faktorů, které příjmy významně ovlivňují. Tento trend se promítá také do obsahu teoretických i aplikačních příspěvků a článků prezentovaných na mezinárodní úrovni. Na vystižení dynamiky vývoje příjmů prostřednictvím směsí jsou zaměřeny např. práce (Paap, van Dijk, 1998, Pittau, Zelli,

2006, Di Prete, McManus, 2000, Lillard,Willis, 1978), posouzením jeho stability v zemích EU se zabývá článek (Longford, Pittau, 2006). Na odhalení vlivných faktorů a regionálních zvláštností rozdělení příjmů v rámci EU se soustřeďují články (Kneip, Utikal, 2001, Pittau, 2004). Podrobná analýza a modelování stavu a vývoje příjmů v tranzitních ekonomikách byly hlavním cílem disertačních prací dvou účastníků projektu, J. Bartošové (Bartošová, 2006/1) a Ľ. Sipkové (Sipková, 2005/2).

K oblíbeným postupům při modelování rozdělení příjmů se v současné době řadí kromě tradičního metod, použitých např. v pracích (Bartošová, 2007/1, 2006/1, 2006/2) také metody kvantilového modelování (viz např. Pacáková, Sipková, Sodomová, 2005, Sipková, 2005/1). Tento progresivní způsob modelování, založený na využití vlastností kvantilové funkce zobecněného lambda rozdělení (RS GLD) (see Ramberg, Schmeiser, 1974) případně Paretova zobecněného rozdělení (viz. např. Luceno, 2006), vystupuje v poslední době do popředí. Teoretické základy nejmodernějších matematicko-statistických klasifikačních metod (jako jsou např. modelování pomocí směsí, skryté Markovovy modely, zobecněné lineární a aditivní modely) nalezneme v pracích (McCullagh, Nelder, 1994, Fahrmeir, Tutz, 1994, Hastie, Tibshirani, 1996, Hastie, Tibshirani, Friedman, 2001, McLachlan, Peel, 2000).

2. Metodika

Analýzy, modelování a porovnávání životní úrovně v ČR, SR a dalších státech nabývá po vstupu těchto zemí do EU na důležitosti. Jednoduchý a oblíbený přístup k modelování rozdělení příjmů (a tedy i životní úrovně) pomocí jednoduchých distribucí se může v období transformačního procesu, kterými prochází česká i slovenská ekonomika, setkávat s obtížemi způsobenými nestabilitou vývoje příjmů v tomto období. K tomu, abychom mohli objektivně posoudit, zda je vybraný pravděpodobnostní model vhodný, je nezbytné mj. provést odhady parametrů modelu i testy shody s maximální přesností. Proto musí být věnována zvýšená pozornost nejenom volbě modelu, ale také výběru metod konstrukce a optimalizaci jejího provedení.

Zdrojem dat, která byla použita při modelování rozdělení příjmů českých a slovenských domácností v období transformace, jsou výběrová šetření realizovaná Českým statistickým úřadem a Slovenským štatistickým úradom v roce 2002. Při těchto šetřeních nebyl použit jednotný systém dělení domácností do sociálních skupin, takže výsledky nelze přímo srovnávat. Za výchozí je v tomto článku považováno rozdělení českých domácností do devíti sociálních skupin podle typu zaměstnání osoby v čele domácnosti (viz tabulka 1). Z důvodů alespoň částečné srovnatelnosti výsledků bylo ve Slovenské republice použito podobné dělení výběrového souboru – v tomto případě byl soubor rozdělen do osmi sociálních skupin (viz tabulka 2). K charakterizaci životní úrovně obyvatelstva byly použity hodnoty celkových příjmů domácností přepočítané na jednoho člena (tj. příjmy na hlavu).

Hlavním cílem konstrukce teoretického modelu je, jak známo, dosažení maximální přesnosti při vystižení empirického rozdělení četností (Bartošová, 2006/3). Dalším kritériem při výběru modelu je jeho flexibilita a jednoduchost. K pravděpodobnostnímu modelování empirického rozdělení příjmů obyvatelstva byl zvolen tříparametrický logaritmicko-normální model, který je v praxi často používán. Vzhledem k tomu, že výběrové soubory příjmů českých a slovenských domácností z roku 2002 jsou velké (v ČR je to 7973 a v SR 16337 domácností), byla při konstrukci

logaritmicko-normálních modelů s parametry μ, σ^2 a γ, kde γ je teoretické minimum, použita metoda maximální věrohodnosti. Maximálně věrohodné odhady parametru γ, které nelze získat přímo řešením soustavy věrohodnostních rovnic, byly určovány numericky dvěma způsoby:

- vyhledáním *maxima modifikované formy logaritmické věrohodnostní funkce*

$$\tilde{\ell}(\gamma) = -n\left[\hat{\mu}(\gamma) + \frac{1}{2}\ln\hat{\sigma}^2(\gamma)\right],$$

  kde $\hat{\mu}(\gamma)$ a $\hat{\sigma}^2(\gamma)$ jsou maximálně věrohodné odhady parametrů μ a σ^2, γ je hledaná hodnota teoretického minima modelu,

- vyhledáním *minima věrohodnostního poměru*

$$LR(\mu,\sigma^2,\gamma|n) = 2[\ell(\vec{p}|n) - \ell(\vec{\pi}(\mu,\sigma^2,\gamma)|n)],$$

  kde $\vec{p}$ je vektor empirických četností, $\vec{\pi}(\mu,\sigma^2,\gamma)$ je vektor pravděpodobností obsazení jednotlivých tříd, do nichž jsou data setříděna, $\ell(\vec{p}|n)$ a $\ell(\vec{\pi}(\mu,\sigma^2,\gamma)|n)$ jsou hodnoty příslušných logaritmických věrohodnostních funkcí.

Odhad parametru γ (teoretického minima) nabývá v modelech rozdělení příjmů často záporné hodnoty, lze však předpokládat, že interval, na němž se tento odhad nalézá, je zdola omezený. S ohledem na charakter zkoumaného znaku – čisté roční peněžní příjmy domácností – byla za první aproximaci dolní hranice tohoto intervalu autorkou zvolena hodnota $-x_{\max}$. Počáteční rozsah vyhledávání maximální hodnoty funkce $\tilde{\ell}(\gamma)$ je tedy dán intervalem $\langle -x_{\max}, x_{\min}\rangle$. Maximum je zde zjišťováno iteračně na pohyblivé mřížce, jejíž hustota se v každém iteračním kroku zvyšuje. V případě, že funkce $\tilde{\ell}(\gamma)$ dosáhne svého maxima na dolním okraji intervalu, posune se vyhledávací procedura vlevo. Tato autorkou navržená iterační procedura se skládá ze dvou cyklů (vnějšího a vnitřního) a je realizována v programu MatLab (viz Bartošová, 2006/1).

Použitelnost logaritmicko-normálního modelu byla testována prostřednictvím věrohodnostního poměru (statistika LR). Na výsledky testování má, jak známo, vliv také počet tříd, do nichž jsou data rozdělena. Problémem volby optimálního počtu tříd m se zabývá mnoho publikací. V tomto článku byl k výpočtu použit vztah $m = 15 \cdot \sqrt[5]{\left(\frac{n}{100}\right)^2}$, který je vhodný pro velké datové soubory, tj. pro soubory s rozsahem $n > 80$ (viz Williams, 2001).

3. Dosažené výsledky

Výsledky modelování rozdělení příjmů českých a slovenských domácností v r. 2002 pomocí tříparametrického logaritmicko-normálního rozdělení jsou obsahem následujících tabulek. Tabulka 1 zobrazuje informace o výsledcích modelování, které byly dosaženy v ČR, tabulka 2 uvádí informace o výsledcích modelování v SR. V tabulkách jsou uvedeny odhady parametrů logaritmicko-normálních modelů v jednotlivých sociálních skupinách i celkem, bez ohledu na sociální skupinu, dále pak hodnoty testovací statistiky LR a kritické hodnoty (tj. kvantily $\chi^2_{0,95}(m-4)$, kde m je počet tříd, do nichž byla data při testování rozdělena). K odhadu parametru γ byla použita v tomto případě numerická minimalizace věrohodnostního poměru.

Tabulka 1. Porovnání empirického rozdělení s logaritmicko-normálními modely v ČR r. 2002 (Příjmy na hlavu)

Sociální skupina	**Odhady parametrů**			**Statistika** LR	**Kvantil** $\chi^2_{0,95}$(m-4)
	μ	σ^2	γ		
Dělníci	11,728	0,21680	7676	68,148684	69,956832
Samostatně činní	11,357	0,49730	10292	51,892489	49,587884
Zaměstnanci	11,187	0,38765	20416	48,408230	72,443307
Samostatně hospodařící	10,287	0,58180	16993	1,933071	11,344867
Družstevní rolníci	13,471	0,00045	-753192	1,328503	9,210340
Důchodci s EA členy	11,228	0,09971	-5149	30,543463	36,190869
Důchodci bez EA členů	11,597	0,03570	-9825	**717,461299**	**77,385962**
Nezaměstnaní	10,404	0,26975	6940	10,492398	30,577914
Ostatní	10,548	0,68914	11466	19,937594	27,688250
Všichni	11,353	0,24865	6831	**962,734268**	**114,694895**

Tabulka 1. Porovnání empirického rozdělení s logaritmicko-normálními modely v SR r. 2002 (Příjmy na hlavu)

Sociální skupina	**Odhady parametrů**			**Statistika** LR	**Kvantil** $\chi^2_{0,95}$(m-4)
	μ	σ^2	γ		
Zaměstnanci	11,3409	0,56956	3864	**232,400269**	**118,235749**
Podnikatelé se zaměst.	11,3962	0,76043	12808	21,318454	37,566235
Podnikatelé bez zaměst.	11,2368	0,67905	15305	35,458674	50,892181
Družstevníci	11,2447	0,57628	6865	25,501100	29,141238
Důchodci s EA členy	11,2709	0,67044	2563	17,602500	37,566235
Důchodci bez EA členů	10,6123	0,69914	2178	**279,810095**	**102,816314**
Nezaměstnaní	10,9011	0,63036	1286	61,151730	53,485772
Ostatní	10,8337	0,89237	4894	29,448522	37,566235
Všichni	11,0857	0,72664	2017	**296,481445**	**148,570958**

Výsledky testování použitelnosti zkonstruovaných logaritmicko-normálních modelů při modelování rozdělení příjmů českých a slovenských domácností v roce 2002 (viz tabulky 1 a 2) ukazují, že ve většině případů bylo dosaženo dobré shody teoretického a empirického rozdělení četností.

V České republice byla prokázána výrazná neshoda pouze v případě rozdělení příjmů domácností důchodců bez ekonomicky aktivních (EA) členů a všech

domácností. Výraznější odlišnost vykazuje rozdělení příjmů domácnosti důchodci bez EA členů, kde podíl $\frac{LR}{\chi^2_{0,95}(m-4)}$ dosáhl hodnoty 9,2712, zatímco v případě všech domácností to bylo přibližně 8,3939. V Slovenské republice taková výrazná neshoda empirického rozdělení s logaritmicko-normálním modelem prokázána nebyla. I zde byla pozorována výraznější odlišnost v případě rozdělení příjmů domácností důchodců bez EA členů a všech domácností, podíl $\frac{LR}{\chi^2_{0,95}(m-4)}$ však v žádném z těchto případů nepřesáhl hodnotu 2,7215 (pro domácnosti důchodců bez EA členů). Rovněž u domácností zaměstnanců dosáhl tento podíl zřetelně vyšší hodnoty (1,9656).

4. Závěry

Z provedené analýzy vylývá, že tříparametrické logaritmicko-normální rozdělení lze považovat v roce 2002 za vhodný model rozdělení příjmů jednotlivců v převážné většině sociálních skupin, a to jak v Čechách, tak i na Slovensku. Významnější odlišnost lze pozorovat pouze ve skupině domácností důchodců bez EA členů a všech domácností v ČR, kde hodnota testového kritéria dosáhla téměř desetinásobku kritické hodnoty. Rovněž v SR byla v těchto skupinách, a navíc také ve skupině domácností zaměstnanců, zaznamenána určitá odlišnost, která je však méně výrazná. Hodnota testovací statistiky v žádné z těchto sociálních skupin nedosáhla trojnásobku kritické hodnoty.

Příčinu vysoké shody empirického rozdělení příjmů s logaritmicko-normálním modelem lze spatřovat nejenom v modelu samotném, ale také ve vysoké kvalitě jeho konstrukce a využití optimalizace při testování shody pomocí věrohodnostního poměru. Můžeme tedy konstatovat, že navržená optimalizační procedura pro odhad parametrů umožňuje kvalitní konstrukci modelu rozdělení příjmů českých a slovenských domácností, a to i v současném nestabilním období transformace.

5. Použitá literatura

Bartošová, J., Bína, V. 2007/1. Mixture Models of Household Income Distribution in the Czech Republic, In Kováčová, M. (ed.) *6th International Conference APLIMAT 2007, Part I*. Slovak University of Technology, Bratislava. 307-316.

Bartošová, J. 2006/1. *Volba a aplikace metod analýzy stavu rozdělení příjmů domácností v České republice po roce 1990*. Ph.D. Thesis. FIS VŠE, Praha.

Bartošová, J. 2006/2. Logarithmic-Normal Model of Income Distribution in the Czech Republic, *Austrian Journal of Statistics Vol.**35**, Nr.**2&3***. 215-222.

Bartošová, J. 2006/3. Logarithmic-normal Model of Household Income Distribution in the Czech Republic after 1990. *Forum Statisticum Slovacum, **3/2006***, Slovak Statistical and Demographical Society, Bratislava. 3-10.

Di Prete, T. A., McManus, A., 2000, Family change, employment transition, and welfare state: household income dynamics in the United States and Germany. *American Sociological Revue Vol.**65***. 343 – 370.

Fahrmeir, L., Tutz,G. 1994. *Multivariate statistical modeling based on generalized linear models*. Springer, New York.

Hastie, T., Tibshirani, R., Friedman, J. 2001. *The elements of statistical learning: Data mining, inference, and prediction*. Springer, New York.

Hastie, T., Tibshirani, R. 1996. Discriminant analysis by Gaussian mixtures. *Journal of the Royal Statistical Society, Series B* ***58***. 155-176.

Kneip, A., Utikal, K.J. 2001. Inference for Density Families Using Functional Principal Component Analysis. *Journal of the American Statistical Association Vol.****96*** *Nr.****454****, Theory and Methods,* 519-533.

Lillard, L. A., Willis, R. J., 1978. Dynamic aspect of earning mobility. *Econometrica Vol.****46***, 985-1012.

Longford, N.T., Pittau, M.G. 2006. Stability of household income in European countries in the 1990s. *Computational Statistics & Data Analysis* ***51/2006***. 1364-1383.

Luceno, A. 2006. Fitting the generalized Pareto distribution to data using maximum goodness-of-fit estimators. *Computational Statistics & Data Analysis* ***51/2006***. 904-917.

McCullagh, P., Nelder, J.A. 1994. *Generalized Linear Models*. Chapman and Hall, London.

McLachlan, G., Peel, D. 2000. *Finite mixture models*. John Wiley & Sons, New York.

Pacáková, V., Sipková, Ľ., Sodomová, E. 2005. Štatistické modelovanie príjmov domácností v Slovenskej republike. *Ekonomický časopis Vol.****53*** *Nr.****4***. 427-439.

Paap, R, van Dijk, H.K., 1998. Distribution and mobillity of wealth of nation. *European Economic Revue Vol.****42***. 1269 – 1293.

Pittau, M. G., Zelli, R., 2006, Empirical evidence of income dynamics across EU regions. *Journal of Applead Econometrics,* forthcoming.

Pittau, M.G., 2004. Regional income distributions in the European Union. *Oxford Bulletin Economic Statistics Vol.****67***. 135 – 161.

Ramberg, J., Schmeiser, B. 1974. An approximate method for generating asymmetric random variables. *Communications of the ACM Vol.****17*** *Nr.****2***. 78-82.

Rosner, B. 1975. On the detection of many outliers. *Technometrics* 17. 221-227.

Sipková, Ľ. 2005/1. Modelovanie príjmov domácností zovšeobecneným lambda rozdelením. *Ekonomika a informatika Vol.****3*** *Nr.****1***. 90-164.

Sipková, Ľ. 2005/2. *Štatistická analýza rozdelenia príjmov domácností v Slovenskej republike*. Ph.D. Thesis. FHI EU, Bratislava.

Williams, D. 2001. *Wieghing the Odds, A Course in Probability and Statistics*. Cambridge Univ. Press, Cambridge.

Stručné poznámky k finančnej riskmetrike

Martin Boďa

Abstract: Pervading risk symptomatic to finance has always given rise to attempts to measure it. This especially holds for its market form which arises in financial markets and with occurrence of which investors occasionally tackle. The presence of risk and fears of it are the grounds that a system of measures concerning risk have been developed, successfully being employed throughout usual financial or academic practice. Having this in mind, the article offers a light insight into the metrics of financial risk and systematizes common risk measures, with the accent being laid upon the criteria of coherency and consistency.

Key words: (financial) risk, (financial) riskmetrics, coherent risk measures, consistent risk measures

Úvod

Matematické financie (*mathematical finance*) od svojho vzniku ako vedná disciplína v prácach Louisa Bacheliera a Alfreda Cowlesa[1] si postavili do centra svojej pozornosti kategóriu rizika reálne existujúceho na finančných trhoch a ovplyvňujúceho finančnú aktivitu (správanie sa) jednotlivcov. Bola a stále je považovaná za dominujúci motív vstupu jednotlivcov na finančný trh maximalizácia ich subjektívnej užitočnosti v podobe dosiahnutia požadovaného (často maximálneho) výnosu. Zaujatie príslušných pozícií na dosiahnutie takto formulovaného cieľa sprevádza nestabilita finančných trhov, ktorá spôsobuje odklon realizovaných výnosov od ich predstáv a ktorá sa v označení stelesňuje do subjektívneho pojmu riziko. Možno viesť diskusie o tom, čo to vlastne riziko je, a v tejto súvislosti a na účely článku je možné diskutovať o kategórii finančného rizika, ako sa to objavovalo vo finančnej literatúre 20. storočia. V článku sa však pojem finančné riziko definovať nebude a bude sa chápať intuitívne, pretože každý pokus o globálnu definíciu by bol nedokonalý a odsúdený na kritiku.

Napriek tomu bude zvolený operacionalistický prístup k finančnému riziku a bude akceptovaná rozšírená domnienka matematickofinančnej vedy, a síce, že finančné riziko možno kontrolovať, merať a predpovedať. Tomu zodpovedá stotožnenie finančného rizika s vhodnou mierou rizika a prezentácia akademických názorov týkajúcich sa správnych vlastností vhodných mier rizika.

1. Pojem rizika a finančného rizika

Nemá veľký význam a nie je účelné definovať riziko vo všeobecnej rovine. Každá vykonštruovaná definícia sa líši od prípadu k prípadu, vystihuje iba niektoré špecifické prípady a nie je spôsobilá univerzálne vystihnúť vlastnú podstatu rizika. Je zrejmé účelné hovoriť iba o kategóriách rizika ako o jednotlivých triedach prípadov, ktorých sa riziko týka. Napriek tomu je odborná a pseudoodborná literatúra plná pokusov, ako čo najlepšie vystihnúť obsah tohto pojmu, a ako ho najlepšie definovať. Koncízny a systematický pohľad na filozofickú koncepciu rizika je poskytnutý Holtonom (2004), ktorý pri hľadaní obsahu pojmu vierohodne diferencuje v ponímaní rizika

- subjektívny prístup, ktorý je asociovaný s neistotou (neurčitosťou) a zdôrazňuje subjektívne vnímanie individuálnej rizikovej expozície, a

[1] Louis Bachelier publikoval svoj článok *Théorie de la Speculation* v *Annales Scientifiques de l' École Normale Supérieure* v roku 1900 a Alfred Cowles uverejnil v *Econometrica* svoje práce v rokoch 1933 (*Can Stock Market Forecasters Forecast?*) a 1944 (*Stock Market Forecasting*).

- operacionalistický prístup, ktorý vychádza z presvedčenia, že riziko *per se* nejestvuje, ale že ho možno merať, a preto ho často kladie na roveň s konkrétnou číselnou charakteristikou, ktorou sa meria.

Ukazuje sa byť vhodné vzhľadom aj na vyššie uvedené dôvody riziko nedefinovať, ale iba zakladať na jednoduchom axiomatickom systéme obsahujúcom dva tvrdenia.

AXIÓMA Č. 1. Existuje riziko.

AXIÓMA Č. 2. Riziko možno merať.

Spomedzi viacerých uvažovateľných kategórií rizika sa článok koncentruje na bohatú kategóriu finančných rizík, ktoré Jílek (2000) v konzistencii s praxou vymedzuje ako potenciálnu finančnú stratu subjektu. Pre zaujímavosť je možné uviesť, že v odbornej literatúre sa rozlišujú štyri základné subkategórie finančného rizika (porov. napr. Jílek, 2000, Jorion, 2003), ktoré sú priblížené v schéme č. 1a. Z regulačného hľadiska bolo pre oblasť podnikania finančných inštitúcií Novou bazilejskou kapitálovou dohodou (*New Basel Capital Accord / Basel II*) v roku 2004 zavedené rozlišovanie trojdimenzionálnej štruktúry finančného rizika, ktorého praktický obsah je premietnutý v schéme č. 1b.

Hoci sa chápanie rizika vo viacerých dimenziách javí byť ako vhodný prístup ku komplexnému posudzovaniu rizika jednotlivých subjektov, v ďalšom texte sa záber zredukuje iba na trhovú formu finančného rizika. Ak sa preto použije jednoducho výraz finančné riziko, myslí sa tým trhové finančné riziko.

Schéma č. 1a *Kategórie finančného rizika*
Zdroj: Vlastné spracovanie.

Schéma č. 1b *Dimenzie finančného rizika*
Zdroj: Prevzaté z Gallati (2003).

2. Riskmetrika

Z axiómy č. 2 o merateľnosti vyplýva, že musí existovať miera, ktorá kvalifikuje rizikovú expozíciu subjektu. Z praktického hľadiska sú zaujímavé iba miery rizika, ktoré kvalifikujú riziko číselne a vhodným spôsobom numericky ohodnocujú riziko finančných inštrumentov korešpondujúce s možnými stavmi sveta existujúcimi na nejakom pravdepodobnostnom priestore. Na základe toho je možné (v súlade s literatúrou) formulovať pojem miera rizika.

Definícia č. 1 *Nech* V *je neprázdna množina* F *-merateľných reálnych náhodných premenných na pravdepodobnostnom priestore* $(\Omega, \mathsf{F}, \mathsf{P})$. *Mierou rizika nazývame zobrazenie* $\rho: \mathsf{V} \to \mathsf{E} \cup \{+\infty\}$.

Z definície je zrejmé, že miera rizika je nejaké zobrazenie z množiny V reprezentujúcej finančné inštrumenty, ktorým zodpovedajú možné stavy sveta (v kontexte článku realizované výnosy alebo straty), do číselnej osi E rozšírenej o $+\infty$ (čo umožňuje kvalifikovať „výšku“ rizika napr. v podobe výšky potenciálnej straty).

Vymedzenie pojmu miera rizika v zmysle definície č. 1 pripúšťa za mieru rizika pomerne širokú škálu zobrazení a je zrejmé, že nie každé zobrazenie je *rozumné*. Je na mieste otázka, aká je

optimálna miera rizika (teda: aká miera rizika funduje správne rozhodnutie pre investora). V odpovedi sa teoretické prístupy rôznia a roztvárajú sa nožnice medzi teoretickou optimálnosťou na jednej strane a praktickou uplatniteľnosťou a interpretovateľnosťou na druhej strane. Prax skôr inklinuje k používaniu jednoduchých syntetických mier, akými sú volatilita (vyjadrená najčastejšie ako smerodajná odchýlka potenciálnych výnosov a strát) alebo value at risk. Teória ale upozorňuje na skutočnosť, že bežne používané miery nemusia byť vôbec vhodné pre bezpečné rozhodovanie a môžu zakladať rizikové pozície. V tomto možno hľadať aj príčiny pre vyvinutie akejsi teórie miery finančného rizika, ktorá stanovuje kritériá, ktoré by *ideálna* miera rizika mala spĺňať a posudzuje zaužívané varianty zo stanoveného kriteriálneho hľadiska.

Volatilitu, v súčasnosti klasickú a obľúbenú mieru na hodnotenie rizikovosti expozície, zužitkoval už v roku 1900 Louis Bachelier vo svojej *Théorie de la Speculation*. Masívna kritika jej vhodnosti začala objektívne v 50. – 70. rokoch 20. storočia s rozvojom teórie portfólia a v tomto období vznikli prvé úvahy založené na kritériu stochastickej dominancie motivované úsilím zoradiť varianty rozhodovania v prostredí neistoty. Až v období explózie záujmu o value at risk v polovici 90. rokov 20. storočia sa objavili prvé snahy formulovať pojem miera rizika, definovať jej želateľné vlastnosti a identifikovať známe miery rizika z pohľadu stanovených vlastností. Je možné registrovať dva smerodajné pokusy o zadanie *ideálnej* miery rizika:

- V roku 1997 skupina autorov Artzner, Delbaen, Eber a Heath predstavila axiomatickú koncepciu *ideálnej* miery rizika určenej *koherentnosťou.*
- V roku 2006 skupina autorov Daníelsson, Jorgensen, Sarma, de Vries a Zigrand v rámci svojej komparácie niektorých mier rizika ako normatívnu požiadavku ideálnej miery rizika stanovili jej *konzistentnosť* so stochastickou dominanciou.

2.1 Koherentná miera rizika

Koherentná je podľa Artzner, Delbaen, Eber a Heath (1997, 1999) miera rizika, ktorá disponuje axiomatickými vlastnosťami translačnej invariantnosti, subaditivity, pozitívnej homogenity a monotónnosti.

Definícia č. 2 *Nech* V *je neprázdna množina* F *-merateľných reálnych náhodných premenných na pravdepodobnostnom priestore* $(\Omega, \mathsf{F}, \mathsf{P})$. *Koherentnou mierou rizika sa nazýva miera rizika* ρ *vtedy a len vtedy, keď*

(a.) pre všetky $X \in \mathsf{V}$ *a všetky reálne čísla* ε *platí* $\rho(X+\varepsilon) = \rho(X) - \varepsilon$ *(TRANSLAČNÁ INVARIANTNOSŤ),*

(b.) pre všetky $X_1, X_2 \in \mathsf{V}$ *platí* $\rho(X_1 + X_2) \le \rho(X_1) + \rho(X_2)$ *(SUBADIVITIVA),*

(c.) pre všetky $X \in \mathsf{V}$ *a všetky reálne čísla* $\lambda \ge 0$ *platí* $\rho(\lambda X) = \lambda \rho(X)$ *(POZITÍVNA HOMOGENITA),*

(d.) pre všetky $X_1, X_2 \in \mathsf{V}$ *také, že* $X_1 \le X_2$, *platí* $\rho(X_1) \le \rho(X_2)$ *(MONOTÓNNOSŤ).*

Interpretačne axióma translačnej invariantnosti uspokojuje požiadavku, podľa ktorej zaradenie bezrizikového aktíva (napr. hotovosti) do portfólia znižuje (absolútne) celkovú mieru rizika. Subaditivita je vyjadrením princípu diverzifikácie rizika portfólia, tzn., že riziko celého portfólia je menšie (nanajvýš rovnaké) ako súčet individuálnych rizík jeho subportfólií. Požiadavka pozitívnej homogenity (prvého stupňa) vyjadruje, že ak zmeníme proporcionálne objem držaných finančných inštrumentov v portfóliu (pri zachovaní relatívnej štruktúry), rovnako sa proporcionálne zmení i rizikovosť portfólia. Monotónnosť zase vyjadruje, že ak sú straty v jednom portfóliu zastúpené väčšmi ako straty v druhom portfóliu, je i riziko prvého portfólia väčšie ako riziko druhého uvažovaného portfólia.

2.2 Konzistentná miera rizika

Bolo naznačené, že konzistentnosť sa asociuje so stochastickou dominanciou (ako vlastnosťou). Stochastická dominancia je pojmom teórie rozhodovania a ako taký sa používa všeobecne a nedefinuje sa. Kombinuje v sebe prvky pravdepodobnostného prístupu a prístupu založeného na maximalizácii užitočnosti pri hodnotení komparatívnej výhodnosti (zoraďovaní)

riešení rozhodovacích úloh v prostredí neistoty (obzvlášť v oblasti ekonómie a financií). Jej popularita spočíva v tom, že sa vzdáva násilných predpokladov o type rozdelenia možných stavov sveta (tzn. výnosov a strát) a nevyžaduje poznatky o tvare funkcie užitočnosti rozhodovateľa. Bližšie je definovaná a vysvetlená podstata stochastickej dominancie napr. u Mlynaroviča (2001, s. 140-149). V texte sa vystačí iba s definíciou konzistencie miery rizika podľa Daníelsson, Jorgensen, Sarma, de Vries a Zigrand (2006).

Definícia č. 3 *Nech* V *je neprázdna množina* F *-merateľných reálnych náhodných premenných na pravdepodobnostnom priestore* $(\Omega, \mathsf{F}, \mathsf{P})$ *a nech* $X_1, X_2 \in \mathsf{V}$. *Hovoríme, že miera rizika* ρ *je konzistentná podľa [prvého, druhého] rádu stochastickej dominancie vtedy a len vtedy, keď platí, že ak* X_1 *stochasticky dominuje* X_2 *v [prvom, druhom] ráde, potom* $\rho(X_1) \leq \rho(X_2)$.

2.3 Typológia základných mier rizika

V zásade bez ohľadu na diskusiu o *pseudoideálnosti* je možné rozlišovať dva okruhy mier rizika: totalistické (celkové) miery a redukované (podreferenčné) miery[2]. Totalistické miery rizika (*overall risk measures*) v sebe integrujú kladné i záporné odchýlky od bezrizikovej situácie alebo od nejakého referenčného kritéria, tzn. do seba inkorporujú pozitívne realizácie (výnosy) aj negatívne realizácie (straty) finančného aktíva vzhľadom k očakávaniam. Redukované miery rizika (*downside risk measures*) naproti tomu odstraňujú nedostatok totalistických mier v tom, že za zdroj rizika považujú iba odchýlky negatívnym smerom od nejakého referenčného bodu. Kým totalistické miery vzťahujú riziko k celej množine stavov sveta Ω a sú následne vhodné pre symetrické rozdelenia, redukované miery považujú za riziko iba negatívne stavy sveta v porovnaní s očakávaním a fundujú na redukovanej podmnožine negatívnych stavov sveta $\Omega^{neg} \subseteq \Omega$. Prehľad základných (najčastejšie zdôrazňovaných) mier finančného rizika podľa typologického určenia je priblížený v schéme č. 2.[3]

Z uvedeného prehľadu mier finančného rizika sú pre účely posudzovania intenzity expozície najčastejšie diskutované a používané volatilita, value at risk a expected shortfall a je preto vhodné upozorniť na ich spôsobilosť merať riziko.

- Tradičná volatilita je kritizovaná pre svoju symetrickú konštrukciu (vo forme smerodajnej odchýlky), keďže za rizikové považuje ako realizované výnosy, tak aj realizované straty, čo je v rozpore so všeobecnou (primitívnou) koncepciou rizika, podľa ktorej finančné riziko spočíva (iba) v možnosti finančnej straty. Výnosy, ktoré volatilita vníma negatívne ako príspevok k riziku, sú ale obyčajne motiváciou investovania na finančných trhoch.
- Value at risk a expected shortfall poskytujú iný pohľad na riziko. Kým volatilita meria, o koľko sa v priemere výnosy a straty odlišujú od svojej strednej úrovne (inými slovami, aká veľká je menlivosť výnosov a strát), value at risk je kvantilom rozdelenia výnosov a strát portfólia a teda vymedzuje stanovené percento najnepriaznivejších strát portfólia. Expected shortfall je strednou hodnotou najhorších strát portfólia určených stanoveným percentom. Hoci obe miery rizika poskytujú určitú kvalitatívnu úroveň informácií a hoci sa javí byť expected shortfall *kritickejšou* mierou, v oboch prípadoch platí, že ide o vágny (a spochybňovaný) koncept. V praxi sa používa prakticky jednostranne value at risk, pretože expected shortfall pre svoju náročnú technickú implementáciu zostáva skôr akademickým návrhom.

[2] Napr. Daníelsson, Jorgensen, Sarma a de Vries (2005) a Daníelsson, Jorgensen, Sarma, de Vries a Zigrand (2006) v tejto súvislosti používajú terminológiu *overall risk measures* a *downside risk measures*, zatiaľ čo Kaplanski a Kroll (2000) uvažujú v prvej skupine iba disperzné miery rizika (*dispersion risk measures*) a druhú skupinu označujú ako *below-a-reference point risk measures.*

[3] Vo všeobecnosti platí, že čím vyšším počtom parametrov miera rizika disponuje, tým flexibilnejšie je využiteľnejšia pri modelovaní rizika. V prehľade je s najväčším počtom parametrov zastúpený α -t model (teda s dvoma parametrami). Jeho určitým rozšírením je napríklad Stoneov trojparametrický L-model miery rizika $L_{\alpha,t,A} = \int_{\Omega(A)} |r - t|^{\alpha} dF(r)$. Porov. Stone (1973).

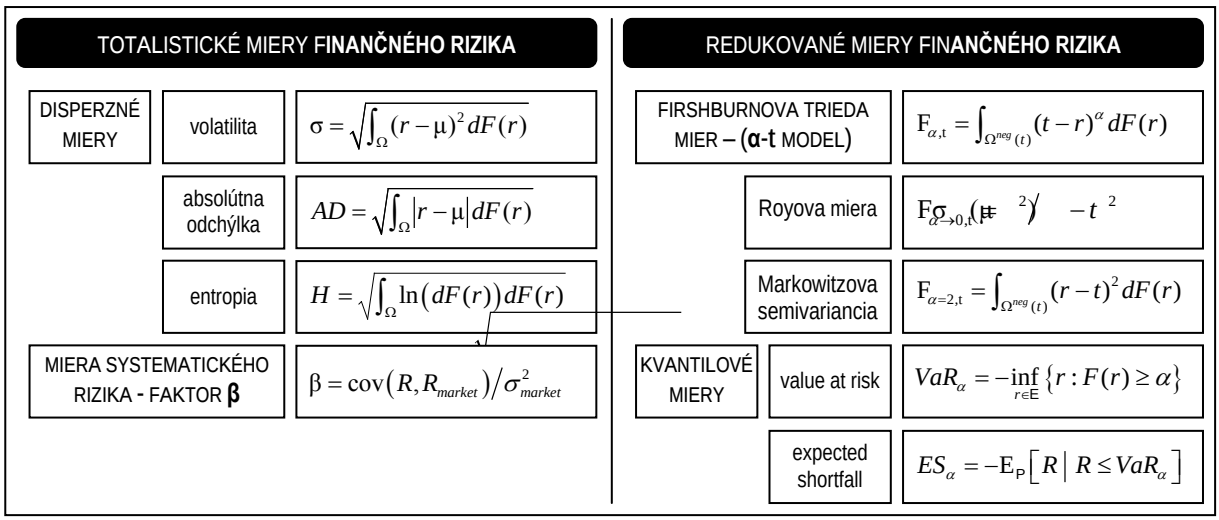


Schéma č. 2 *Prehľad základných mier finančného rizika*
Zdroj: Vlastné spracovanie. (Pozn.: Všetky integrály sú Lebesgueove-Stieltjesove.)

Z hľadiska definovaných kritérií pseudoideálnosti, tzn. koherentnosti v zmysle Artzner, Delbaen, Eber a Heath (1997 a 1999) a konzistentnosti s kritériami stochastickej dominancie sú vytypované miery rizika zhodnotené v schéme č. 3. Pseudoideálnou mierou sa javí byť iba value at risk, ktorá si v bežných prípadoch zachováva konzistenciu, keďže sa volí bežne hodnota parametra $\alpha \in \{0.01, 0.05\}$ a typické finančné aktíva majú rozdelenie ziskov a strát s ťažkými koncami.

MIERA RIZIKA	KONZISTENTNOSŤ S KRITÉRIOM STOCHASTICKEJ DOMINANCIE		KOHERENTNOSŤ
	PRVÉHO RÁDU	DRUHÉHO RÁDU	
VOLATILITA	Konzistentná, ak sa neporovnávajú aktíva, ktorých rozdelenie ziskov / strát patrí do rovnakej triedy rozdelení.	Konzistentná.	Nekoherentná.
VALUE AT RISK	Konzistentná.	Konzistentná pre voliteľné hodnoty parametra $\alpha \in \langle 0, 0.5 \rangle$.	Nekoherentná, ale pri aktívach s rozdelením ziskov / strát s ťažkými koncami *je v koncoch koherentná.*
EXPECTED SHORTFALL	Nekonzistentná.	Nekonzistentná.	Nekoherentná.

Schéma č. 3 *Vlastnosti volatility, value at risk a expected shortfall z hľadiska definovanej pseudoideálnosti*
Zdroj: Vlastné spracovanie podľa Artzner, Delbaen, Eber a Heath (1999), Daníelsson, Jorgensen, Sarma a de Vries (2005), Daníelsson, Jorgensen, Samorodnitsky, Sarma a de Vries (2005), Daníelsson, Jorgensen, Sarma, de Vries a Zigrand (2006), Fishburn (1977), Kaplanski a Kroll (2001).

Záver

V článku boli poskytnuté informácie o aktuálnom stave akademických názorov o *ideálnej* miere rizika. V existencii mier rizika sa zhmotňuje domnienka súčasného finančného manažmentu, že finančné riziko možno operacionalisticky kontrolovať, merať a predpovedať.

Použitá literatúra

ACERBI, Carlo 2002. *Spectral Measures of Risk: a Coherent Representation of Subjective Risk Aversion*. In: *Journal of Banking & Finance*. 2002, č. 7 (júl), roč. 26. S. 1505-1518.

ACERBI, Carlo, TASCHE, Dirk 2001. *Expected Shortfall: a natural coherent alternative to Value at Risk*. In: *Working Papers of Italian Association for Financial Risk Management*. 2001.

ACERBI, Carlo, TASCHE, Dirk 2002. *On the coherence of expected shortfall*. In: *Journal of Banking & Finance.* 2002, č. 7 (júl), roč. 26. S. 1487-1503.

ARTZNER, Philippe, DELBAEN, Freddy, EBER, Jean-Marc, HEATH, David 1997. *Thinking Coherently*. In: *RISK*. 1997, č. 11 (november), roč. 10. S. 68-71.

ARTZNER, Philippe, DELBAEN, Freddy, EBER, Jean-Marc, HEATH, David 1999. *Coherent Measures of Risk*. In: *Mathematical Finance.* 1999, č. 3, roč. 9. S. 203-228.

BOĎA, Martin 2006. *Value at risk I. Value at risk ako miera rizika, alternatívy, nedostatky a regulačný aspekt.* In: *Forum Statisticum Slovacum.* 2006, č. 4, roč. 2. S. 15-24.

DANÍELSSON, Jón, JORGENSEN, Bjørn N., SARMA, Mandira, VRIES, Casper G. de 2005. *Comparing risk measures.* In: *Eurandom Report Series.* 2005, č.2005-008. 15 s.

DANÍELSSON, Jón, JORGENSEN, Bjørn N., SARMA, Mandira, VRIES, Casper G. de 2006. *Comparing downside risk measures for heavy tailed-distributions.* In: *Economics Letters.* 2006, č. 2 (august), roč. 92. S 202-208.

DANÍELSSON, Jón, JORGENSEN, Bjørn N., SARMA, Mandira, SAMORODNITSKY, Gennady, VRIES, Casper G. de, 2005. *Subadditivity Re-Examined: the Case for Value-at-Risk.* In: *FMG Discussion Papers.* 2005, č. dp549 (november). 20 s.

DANÍELSSON, Jón, JORGENSEN, Bjørn N., SARMA, Mandira, VRIES, Casper G. de, ZIGRAND, Jean-Pierre 2006. *Consistent Measures of Risk.* In: *FMG Discussion Papers.* 2006, č. dp565 (máj). 21 s.

FISHBURN, Peter C. 1977. *Mean-Risk Analysis with Risk Associated with Below-Target Returns.* In: *American Economic Review.* 1977, č. 2 (marec 1977), zv. 67. S 116-126.

GALLATI, Reto 2003. *Risk Management and Capital Adequacy*. New York: McGraw-Hill 2003. 555 s. ISBN 0-07-140763-4.

HARDY, Mary 2003. *Investment guarantees: modeling and risk management for equity-linked life insurance.* New Jersey: Wiley 2003. 286 s. ISBN 0-471-39290-1.

HARMANTZIS, Fotios, MIAO, Linyan, CHIEN Yifan 2005. *Empirical Study of Value-at-Risk and Expected Shortfall Models with Heavy Tails.* [Acrobat ® pdf online]. Hoboken [USA]: Stevens Institute of Technology 2005. [Cit. 29. 09. 2006]. Dostupné na World Wide Web: <http://papers.ssrn.com/sol3/Delivery.cfm/ SSRN_ID788624_ code389663.pdf>.

HOLTON, Glyn A. 2002. *History of Value-at-Risk: 1922 – 1998.* In: *EconWPA / Methods and History of Economic Thought*. 2002, č. 0207001. 27 s.

HOLTON, Glyn A. 2004. *Defining Risk.* In: *Financial Analysts Journal*. 2004, č. 6 (november/december), zv. 60. S 19-25.

JÍLEK, Josef 1997. *Finančí trhy.* Praha: Grada Publishing 1997. 529 s. ISBN 80-7169-453-3.

JÍLEK, Josef 2000. *Finančí rizika.* Praha: Grada Publishing 2000. 635 s. ISBN 80-7169-579-3.

JORION, Philippe 2003. *Financial Risk Manager Handbook.* Second Edition. New Jersey: Wiley 2003. 708 s. ISBN 0-471-43003-X

KAPLANSKI, Guy, KROLL, Yoram 2002. *VaR Risk Measures versus Traditional Risk Measures: an Analysis and Survey*. In: *Journal of Risk*. 2002, č. 3 (jar 2002), zv. 4. 32 s.

MLYNAROVIČ, Vladimír 2001. *Finančné investovanie. Teória a aplikácie.* Bratislava: IURA Edition 2001. 293 s. ISBN 80-89047-16-5.

ROCKAFELLAR, R. Tyrell, URYASEV, Stanislav 2002. *Conditional value-at-risk for general loss distributions*. In: *Journal of Banking & Finance.* 2002, č. 7 (júl), roč. 26. S. 1443-1471.

STONE, Bernell K. 1973. *A General Class of Three-Parameter Risk Measures.* In: *The Journal of Finance.* 1973, č. 3 (jún), roč. 28. S 675-685.

TASCHE, Dirk 2002. *Expected shortfall and beyond*. In: *Journal of Banking & Finance.* 2002, č. 7 (júl), roč. 26. S. 1519-1533.

Adresa autora:

Martin Boďa
Ekonomická fakulta, Katedra kvantitatívnych metód a informatiky
Univerzita Mateja Bela v Banskej Bystrici, Tajovského 10, 975 90 Banská Bystrica
e-mail: *martin.boda@umb.sk*

Porovnanie metód pre odstraňovanie deformácií obrazu

RNDr. Róbert Bohdal
Katedra algebry, geometrie a didaktiky matematiky
Fakulta Matematiky, fyziky a informatiky
Univerzita Komenského, Bratislava
bohdal@fmph.uniba.sk

Abstract

The main goal of this paper is to describe some methods for the picture deformation recovery by the image registration methods using the control points, which describe how parts of an image are transformed. An additional goal is to find the most accurate method among them. We also used a new methodology to determine the accuracy of various methods.

Úvod

Katastrálne mapy vlastníkov pôdy sa donedávna archivovali len v klasickej papierovej podobe. Súčasný stav a možnosti výpočtovej techniky umožňujú uchovávať tieto mapy v digitálnom tvare. Vplyvom zmeny vlhkosti bolo mnoho máp tvarovo deformovaných. Pritom z nich potrebujeme dostať pomerne presnú informáciu o výmerách jednotlivých pozemkov. Našťastie už v dobách minulých sa pri kreslení takýchto máp využívali *identifikačné body*, ktoré označovali polohu významných objektov reálneho sveta (stromy, význačné geografické zvláštnosti, alebo rohy väčších budov). Pomocou presnej polohy týchto bodov je možné vhodnými matematickými metódami čiastočne odstrániť deformácie máp.

V tomto príspevku porovnáme najpoužívanejšie transformačné metódy používané pri odstraňovaní deformácie obrazu a štatistickými mierami určíme presnosť s akou tieto deformácie odstraňujú.

Formulácia problému:

Majme zadané dve množiny bodov $\mathcal{P}$ a $\mathcal{V}$ v rovine E^2, $\mathcal{P} = \{\boldsymbol{p}_i[x_i, y_i] \in E^2;\ i = 1, 2, \dots, n\}$ a $\mathcal{V} = \{\boldsymbol{v}_i[x_i', y_i'] \in E^2;\ i = 1, 2, \dots, n\}$. Budeme hľadať takú transformačnú funkciu $\boldsymbol{f}(\boldsymbol{x})$, pre ktorú platí $\boldsymbol{f}(\boldsymbol{p}_i) = \boldsymbol{v}_i$, kde $i = 1, 2, \dots, n$. Jednotlivé dvojice $(\boldsymbol{p}_i, \boldsymbol{v}_i)$ budeme nazývať *odpovedajúce si body*. Pomocou funkcie $\boldsymbol{f}(\boldsymbol{x})$ budeme transformovať všetky body vstupného obrazu a získame nový obraz.

Jednosegmentové transformačné metódy

Do tejto triedy metód patria tie metódy, ktorých transformačné funkcie sú určené jedinou vektorovou funkciou $\boldsymbol{f}(\boldsymbol{x})$ na celom svojom definičnom obore. Výhodou týchto metód je ich jednoduché vyjadrenie, nenáročná implementácia a možnosť vypočítať hodnoty bodov nového obrazu aj mimo konvexného obalu zadaných odpovedajúcich si bodov. Ich transformačné funkcie majú väčšinou globálny vplyv na transformované body obrazu. Modifikáciou týchto metód môžeme získať transformačné funkcie, ktoré majú lokálny vplyv. V tejto časti uvedieme metódu *tenkostenných splajnov* a *Shepardovu metódu*.

Metóda tenkostenných splajnov

Tenkosplajnová metóda patrí medzi najpoužívanejšie metódy pri odstraňovaní tvarových deformácií obrazu. Interpolačná transformačná funkcia $\boldsymbol{f}(\boldsymbol{x})$ je určená vzťahom [ISKE03]:

$$\boldsymbol{f}(x, y) = \boldsymbol{c}_1 + \boldsymbol{c}_2 x + \boldsymbol{c}_3 y + \frac{1}{2}\sum_{i=1}^{n} \boldsymbol{\lambda}_i r_i^2 \log\left(r_i^2\right), \text{ kde } [x_i, y_i] \in E^2. \qquad (1)$$

Riešenie vzťahu (1) nájdeme, ak parametre $\boldsymbol{\lambda}_i$, $i = 1, \dots, n$ spĺňajú nasledujúce okrajové podmienky:

$$\sum_{i=1}^{n} \boldsymbol{\lambda}_i = \mathbf{0} \quad \text{a} \quad \sum_{i=1}^{n} \boldsymbol{\lambda}_i \boldsymbol{p}_i = \mathbf{0}\,. \tag{2}$$

Použitím podmienok interpolácie $\boldsymbol{f}(\boldsymbol{p}_i) = \boldsymbol{v}_i$, kde $i = 1, 2, \ldots, n$ spolu s okrajovými podmienkami (2) vypočítame neznáme hodnoty $\boldsymbol{c}_1, \boldsymbol{c}_2, \boldsymbol{c}_3$ a $\boldsymbol{\lambda}_i$, $i = 1, 2, \ldots, n$ z nasledujúcej sústavy rovníc:

$$\begin{pmatrix} 0 & 0 & 0 & 1 & 1 & \cdots & 1 \\ 0 & 0 & 0 & x_1 & x_2 & \cdots & x_n \\ 0 & 0 & 0 & y_1 & y_2 & \cdots & y_n \\ 1 & x_1 & y_1 & 0 & r_{21}^2 \log\left(r_{21}^2\right) & \cdots & r_{n1}^2 \log\left(r_{n1}^2\right) \\ 1 & x_2 & y_2 & r_{12}^2 \log\left(r_{12}^2\right) & 0 & \cdots & r_{n2}^2 \log\left(r_{n2}^2\right) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & y_n & r_{1n}^2 \log\left(r_{1n}^2\right) & r_{2n}^2 \log\left(r_{2n}^2\right) & \cdots & 0 \end{pmatrix} \begin{pmatrix} \boldsymbol{c}_1 \\ \boldsymbol{c}_2 \\ \boldsymbol{c}_3 \\ \boldsymbol{\lambda}_1 / 2 \\ \boldsymbol{\lambda}_2 / 2 \\ \vdots \\ \boldsymbol{\lambda}_n / 2 \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{0} \\ \mathbf{0} \\ \boldsymbol{f}_1 \\ \boldsymbol{f}_2 \\ \vdots \\ \boldsymbol{f}_n \end{pmatrix} \tag{3}$$

kde $r_{ij}^2 = r_{ji}^2 = \left(x_j - x_i\right)^2 + \left(y_j - y_i\right)^2$.

Shepardova metóda

Shepardova metóda uvedená v práci [SHEP68], patrí medzi najznámejšie postupy pri riešení interpolačného problému *nerovnomerne rozptýlených dát.*

Rovnako ako v predošlej metóde, hľadáme transformačnú funkciu $\boldsymbol{f}(\boldsymbol{x})$, pre ktorú platí $\boldsymbol{f}(\boldsymbol{p}_i) = \boldsymbol{v}_i$, pre $i = 1, 2, \ldots, n$. Funkciu $\boldsymbol{f}(\boldsymbol{x})$ Shepard definoval ako vážený súčet [HOS93]:

$$\boldsymbol{f}(\boldsymbol{x}) = \sum_{i=1}^{n} \omega_i(\boldsymbol{x}) \boldsymbol{v}_i\,. \tag{4}$$

Váhové funkcie $\omega_i(\boldsymbol{x})$ vo vzťahu (4) sú vyjadrené nasledovne:

$$\omega_i(\boldsymbol{x}) = \frac{\sigma_i(\boldsymbol{x})}{\sum_{j=1}^{n} \sigma_j(\boldsymbol{x})}, \tag{5}$$

kde $\sigma_i(\boldsymbol{x}) = \left\|\boldsymbol{x} - \boldsymbol{p}_i\right\|^{-\mu_i}$, pre $\mu_i > 0$. Parameter μ_i umožňuje modifikovať tvar výslednej plochy v okolí interpolovaných bodov.

Globálny vplyv tejto metódy môžeme redukovať prenásobením funkcií $\sigma_i(\boldsymbol{x})$ tzv. *tlmiacou funkciou* $\lambda_i(\boldsymbol{x})$. Najznámejšia *Frankeova-Littleova* tlmiaca funkcia má nasledovné vyjadrenie [HOS93]:

$$\boldsymbol{\lambda}_i(\boldsymbol{x}) = \left(1 - \frac{d_i(\boldsymbol{x})}{R_i}\right)_+^{\mu}, \text{ kde } d_i(\boldsymbol{x}) = \left\|\boldsymbol{x} - \boldsymbol{p}_i\right\| \text{ a } R_i > 0.$$

Franke a Nielson zvolili hodnotu $R_i = \frac{D}{2}\sqrt{\frac{N_w}{n}}$, kde D je najväčšia vzdialenosť medzi ľubovoľnými dvoma bodmi množiny $\mathcal{P}$ zadaných bodov a N_w je pevne zvolené celé číslo (obyčajne $N_w = 19$).

Franke a Nielson v práci [FRAN80] zovšeobecnili Shepardovu metódu použitím tzv. *lokálnych interpolantov.* Namiesto hodnôt $\boldsymbol{v}_i$ vo vzťahu (4) použili lokálne interpolačné funkcie $\boldsymbol{L}_i(\boldsymbol{x})$ s vlastnosťou $\boldsymbol{L}_i(\boldsymbol{p}_i) = \boldsymbol{v}_i$:

$$\boldsymbol{f}(\boldsymbol{x}) = \sum_{i=1}^{n} \omega_i(\boldsymbol{x}) L_i(\boldsymbol{x})\,. \tag{6}$$

Za interpolačné funkcie zvolili triedu kvadratických polynómov a dosiahli dostatočne hladké plochy s relatívne nízkou výpočtovou náročnosťou.

Modifikovanú kvadratickú Shepardovu metódu môžeme opísať nasledovným postupom: Prepísaním vzťahu (6) do vhodnejšieho tvaru dostávame:

$$\boldsymbol{f}(x, y) = \sum_{i=1}^{n} \omega_i(x, y) Q_i(x, y), \tag{7}$$

kde lokálny kvadratický interpolant $\boldsymbol{Q}_k(x, y)$ je definovaný vzťahom:

$$\begin{aligned} Q_i(x, y) = {} & c_{i,1}(x - x_i)^2 + c_{i,2}(x - x_i)(y - y_i) + \\ & c_{i,3}(y - y_i)^2 + c_{i,4}(x - x_i) + c_{i,5}(y - y_i) + v_i . \end{aligned} \tag{8}$$

Koeficienty pre $\boldsymbol{Q}_i(x, y)$ vypočítame pomocou metódy najmenších štvorcov z podmienky:

$$\sum_{k=1,\, k\neq i}^{n} \omega_k(x_i, y_i)\left[c_{i,1}(x_k - x_i)^2 + \cdots + c_{i,5}(y_k - y_i) + f_i - f_k\right]^2 \to 0, \tag{9}$$

kde $\omega_k(x, y) = \left(\dfrac{R_q - d_k(x, y)}{R_q d_k(x, y)} \right)_+^2$, pričom $d_k(x, y) = \sqrt{(x - x_k)^2 + (y - y_k)^2}$. Symbol R_q označuje pevne daný polomer vplyvu pre bod $\boldsymbol{p}_i[x_i, y_i]$ a k nemu zodpovedajúci lokálny interpolant $\boldsymbol{Q}_i(x, y)$.

Viacsegmentové metódy

Viacsegmentové metódy (známe ako *metódy konečných prvkov*) sú založené na triangulácii konvexného obalu vstupných odpovedajúcich si bodov množiny $\mathcal{P}$. Vo všetkých nasledujúcich metódach je nevyhnutné, aby celý obraz, ktorý chceme transformovať, ležal vnútri konvexného obalu bodov $\boldsymbol{p}_i$ množiny $\mathcal{P}$.

Cloughova-Tocherova metóda

Nevýhodou po častiach lineárnych interpolačných plôch je to, že sú iba C^0-spojité. C^1-spojitosť vyžaduje konštruovať interpolačné plochy zložené z polynómov vyššieho stupňa ako 1. Na zaručenie C^1-spojitosti pozdĺž hraníc potrebujeme poznať nielen súradnice $[x_i, y_i]$ vrcholov a ich hodnoty z_i, ale aj dotykovú rovinu (resp. gradient) v danom vrchole, ako i derivácie v priečnom smere prislúchajúce jednotlivým hranám trojuholníkovej siete. Keďže hodnoty gradientov nie sú obyčajne známe, musia byť nejakým spôsobom vypočítané zo zadaných súradníc vrcholov. Detailnejší opis niektorých metód na odhad derivácií môžeme nájsť v [STE84] a [AKI84]. Pri konštrukcií C^1-spojitého kubického Cloughoveho-Tocheroveho interpolantu rozdelíme každý trojuholník pôvodnej triangulácie na tri *minitrojuholníky* spojením každého vrcholu trojuholníka s jeho ťažiskom.

Hľadanou interpolačnou funkciou bude C^1-spojitá plocha zložená z *Bézierovych trojuholníkových záplat* stupňa 3 nad každým minitrojuholníkom.

Cloughova-Tocherova metóda používa kubické Bézierove trojuholníkové záplaty:

$$\begin{aligned} \boldsymbol{X}(u, v, w) = {} & \boldsymbol{b}_{300}u^3 + 3\boldsymbol{b}_{210}u^2v + 3\boldsymbol{b}_{120}uv^2 + \\ & \boldsymbol{b}_{030}v^3 + 3\boldsymbol{b}_{021}v^2w + 3\boldsymbol{b}_{012}vw^2 + \\ & \boldsymbol{b}_{003}w^3 + 3\boldsymbol{b}_{102}w^2u + 3\boldsymbol{b}_{201}wu^2 + 6\boldsymbol{b}_{111}uvw. \end{aligned} \tag{10}$$

Vyčíslenie riadiacich bodov

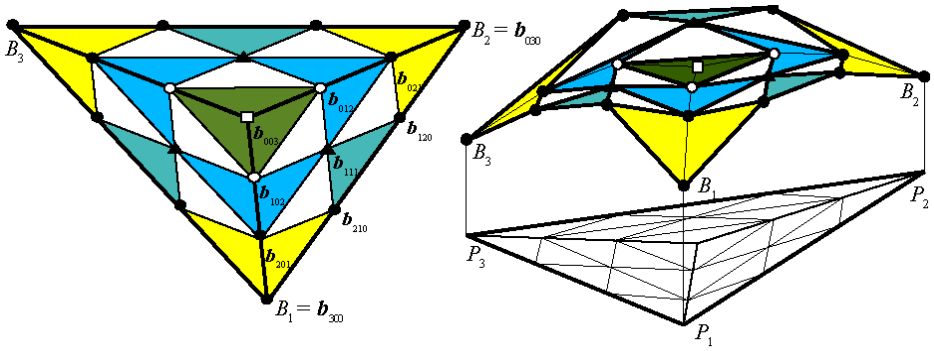
Bézierove vrcholy riadiacej siete, troch stýkajúcich sa trojuholníkových záplat, vyčíslime nasledovným postupom [AMID02]:

Súradnice [*x*, *y*] Bézierovych vrcholov nad každým minitrojuholníkom sú určené súradnicami [*x*, *y*] vrcholov minitrojuholníka, ďalej súradnicami bodov ležiacich v 1/3 a 2/3 každej jeho strany a nakoniec súradnicami ťažiska minitrojuholníka.

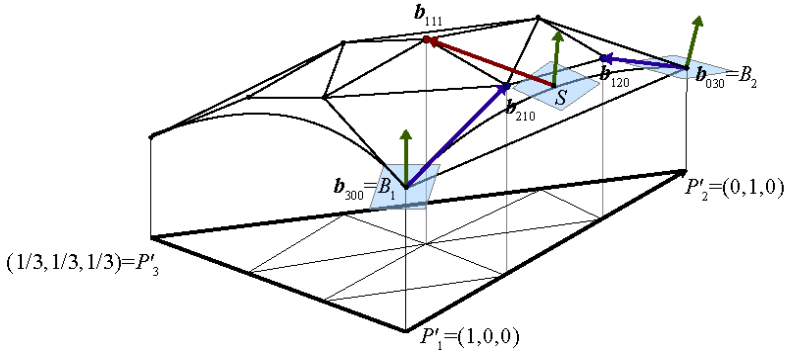
Hodnoty z-ových súradníc Bézierovych vrcholov určíme pomocou nasledovných krokov:

1. Hodnoty z Bézierovych vrcholov (nad P_1 a P_2) označených „●“ sú *z*-ové hodnoty súradníc bodov B_1 a B_2 z danej triangulácie (pozri obrázok 1 vpravo).
2. Hodnoty z vrcholov označených „●“ ležiacich na okraji záplaty vypočítame z podmienky, že tieto vrcholy ležia v dotykovej rovine určenej bodom B_1 alebo B_2 a normálovým vektorom v príslušnom bode (pozri obrázok 1 vpravo a obrázok 2).
3. Hodnoty z vrcholov označených „●“ ležiacich na spojnici ťažiska trojuholníka s jeho vrcholmi vypočítame z podmienky, že ležia v rovine určenej dvomi už vypočítanými vrcholmi „●“ a jedným daným bodom B_i.
4. Hodnoty z troch vrcholov označených „▲“ určíme z odhadnutých derivácií v priečnom smere v strede každej z troch hrán trojuholníka $B_1B_2B_3$. Tieto vrcholy budú ležať v rovine určenej bodom *S* Bézierovej krivky (pre hodnotu parametra t = ½) danej riadiacimi bodmi $\boldsymbol{b}_{300}$, $\boldsymbol{b}_{210}$, $\boldsymbol{b}_{120}$, $\boldsymbol{b}_{030}$ a normálovým vektorom vypočítaným z lineárnej kombinácie normálových vektorov vo vrcholoch strany trojuholníka (pozri obrázok 2 a obrázok 1 vľavo).
5. Hodnoty z ďalších troch vrcholov označených „○“ vypočítame z podmienky, že ležia v rovine určenej dvomi vrcholmi „▲“ a jedným vnútorným vrcholom „●“ (pretože dva susedné riadiace mikrotrojuholníky s vrcholmi „▲, ●, ○“ musia byť komplanárne, pozri obrázok 1).
6. Posledný hľadaný Bézierov vrchol „□“, ležiaci nad ťažiskom trojuholníka $P_1P_2P_3$, leží v rovine určenej tromi vrcholmi „○“, pretože tri „prostredné“ trojuholníky musia byť komplanárne.

Ak už poznáme hodnoty súradníc vrcholov $\boldsymbol{b}_{ijk}$ môžeme použiť vzťah (10) na vyčíslenie bodov Bézierovej záplaty nad konkrétnym minitrojuholníkom. Rovnakým spôsobom budeme postupovať pri všetkých ostatných trojuholníkoch danej triangulácie, čím získame C^1-spojitú interpolačnú plochu.



Obrázok 1: Konštrukcia siete riadiacich Bézierovych vrcholov nad tromi minitrojuholníkmi.



Obrázok 2: Normály a derivácia v priečnom smere určujú polohu Bézierovych bodov $\boldsymbol{b}_{210}$, $\boldsymbol{b}_{120}$ a $\boldsymbol{b}_{111}$.

Zhodnotenie jednotlivých metód

Presnosť jednotlivých metód pre odstraňovanie deformácií obrazu sme určili na základe obrázku obsahujúceho čiernobielu mriežku a jej troch deformácií. Prvý deformovaný obrázok bol vytvorený stiahnutím stredov okrajov mriežky smerom k jej stredu. Druhý obrázok bol zdeformovaný transformáciou, ktorá zvlnila okraje mriežky. Posledný obrázok bol vytvorený štyrmi lokálnymi deformáciami, z ktorých dve posunuli časti obrazu, ďalšia zväčšila a posledná zmenšila časť obrazu.

Na každý z deformovaných obrázkov sme aplikovali jednotlivé metódy pre odstránenie deformácií obrazov. Pre odstránenie deformácie v prvom obrázku sme použili 13, v druhom 37 a v treťom 86 dvojíc odpovedajúcich si bodov. Počet bodov sme zvolili podľa povahy jednotlivých deformácií.

Presnosť s akou jednotlivé metódy odstránili deformáciu obrazu sme určili pomocou:

- *odmocniny strednej odchýlky - RMSE:*

$$RMSE = \sqrt{\frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(x_{ij} - y_{ij}\right)^2},$$

kde symbolom x_{ij} sú označené hodnoty pixlov vzorového obrazu a symbolom y_{ij} hodnoty pixlov porovnávaného obrazu,

- *pomeru signálu k šumu - SNR*:

$SNR = 10 * \log_{10}\left(\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}^2}{MSE}\right)$, kde symbol *MSE* označuje strednú kvadratickú odchýlku,

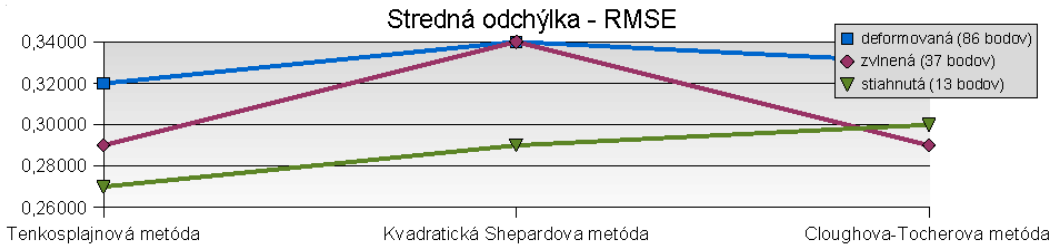
- *koeficientu krížovej korelácie - CC*:

$$CC = \left|\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}y_{ij} - mn\bar{x}\bar{y}}{\sqrt{\left(\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}^2 - mn\bar{x}^2\right)\left(\sum_{i=1}^{m}\sum_{j=1}^{n} y_{ij}^2 - mn\bar{y}^2\right)}}\right|,$$

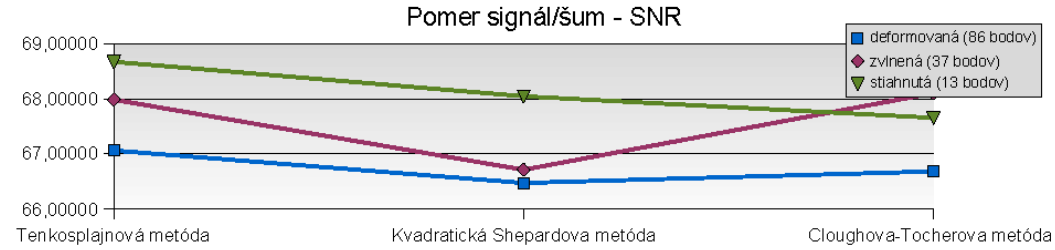
kde $\bar{x}$, $\bar{y}$ sú priemerné hodnoty: $\bar{x} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} x_{ij}$ a $\bar{y} = \frac{1}{mn}\sum_{i=1}^{m}\sum_{j=1}^{n} y_{ij}$.

		deformovaná (86 bodov)	zvlnená (37 bodov)	vtiahnutá (13 bodov)
Tenkosplajnová metóda	*RMSE*	**0,31974**	0,29178	**0,26966**
	SNR	**67,05630**	67,98896	**68,68383**
	CC	**0,77261**	0,80886	**0,83655**
Kvadratická Shepardova metóda	*RMSE*	0,34258	0,33808	0,29000
	SNR	66,47147	66,71043	68,04685
	CC	0,73866	0,74335	0,81108
Cloughova-Tocherova metóda	*RMSE*	0,33223	**0,28750**	0,30271
	SNR	66,68221	**68,10920**	67,65690
	CC	0,75509	**0,81454**	0,79438

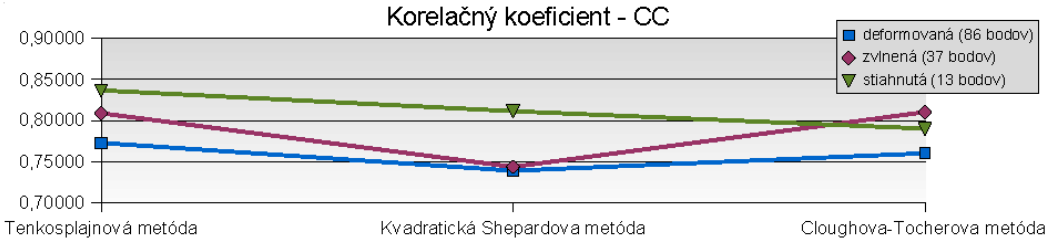
Tabuľka 1: Porovnanie metód pre odstraňovanie deformácií obrazu zvlnenej, stiahnutej a deformovanej mriežky.



Graf 1: Porovnanie metód pre odstraňovanie deformácií obrazu zvlnenej, stiahnutej a deformovanej mriežky. Koeficient *RMSE*.



Graf 2: Porovnanie metód pre odstraňovanie deformácií obrazu zvlnenej, stiahnutej a deformovanej mriežky. Koeficient *SNR*.



Graf 3: Porovnanie metód pre odstraňovanie deformácií obrazu zvlnenej, stiahnutej a deformovanej mriežky. Koeficient *CC*.

Záver

Presnosť jednotlivých metód pre odstránenie deformácie obrazu, sme určili pomocou viacerých štatistických mier (*RMSE*, *SNR*, *CC*). Porovnávali sme vždy pôvodný, nedeformovaný obrázok s obrázkom, v ktorom bola odstránená deformácia konkrétnou metódou. Výsledky týchto porovnaní sú uvedené v tabuľke č. 1 a v grafe č. 1 až 3. Zo všetkých metód na odstránenie deformácií obrazu bola vo všeobecnosti najpresnejšia tenkosplajnová metóda. Práve túto metódu pomerne často používajú autori pri vývoji aplikácií, ktoré vykonávajú transformáciu obrazu.

Literatúra

[AKI84] H. Akima: *On Estimating partial derivatives for bivariate interpolation of scattered*. Rocky Mountain Journal of Mathematics 1(14), 1984, (str. 41-52).

[AMID02] I. Amidror: *Scattered data interpolation methods for electronic imaging systems*. Journal of Electronic Imaging 2(11), 2002, (str. 157-176).

[FRAN80] R. Franke, G. Nielson: *Smooth interpolation of large sets of scattered data*. Intern. Journal for Numerical Methods in Engineering (15), 1980, (str. 1691-1704).

[HOS93] J. Hoschek, D. Lasser: *Fundamentals of Computer Aided Geometric Design*. A K Peters, Wellesley, MA, 1993, (str. 388-421).

[ISKE03] A. Iske: *Radial basis functions: basics, advanced topics and meshfree methods for Transport Problem*. Seminar of Mathematics, 2003, (str. 247-274).

[SHEP68] D. Shepard: *A two dimensional interpolation function for irregular spaced data*. Proceedings 23rd ACM National Conference, 1968, (str. 517-524).

[STE84] S. Stead: *Estimation of gradients from scattered data*. Rocky Mountain Journal of Mathematics 1(14), 1984, (str. 265-279).

The using of the PCA method for measuring risk of the financial portfolios[1]

Mária Bohdalová, Iveta Stankovičová

Abstract: The correlations are essential driving portfolio risk. When the number of assets or the number of the risk factors is large, however, the measurement of the covariance matrix becomes increasingly difficult. The number of correlations increases geometrically with number of assets. For large portfolios, this causes real problems: the portfolio VaR may not be positive and correlations may be estimated imprecisely. This paper examines how can be these problems solved to aid Principal Component Analysis method.

Key words: PCA methods, models VaR, Monte Carlo simulation method, measuring of the risks of the financial portfolios

1. Introduction

Principal component analysis (PCA) is a mainstay of modern data analysis that is widely used in many practical applications. The goal of this paper is present how this method may be used in the identifications independent sources of financial risk within a large system. Bohdalová and Stankovičová [2006] have shown the base principle of this method and they showed how can be used PCA in the analyse of the risk factors of the options portfolios. This paper extends these ideas to two other important areas: firstly the efficient computation of large positive semi-definite covariance matrices, and secondly the modeling of multivariate scenarios for computing risk measures VaR[2] (Value at Risk). Both problems have immediate applications to internal models for measuring market risk. The outlines the theory and methodology for using a few key market risk factors that represent only the most important independent sources of information to generate large covariance matrices. These matrices will be positive semi-definite, relative stable over time, and may be computed easily using sophisticated models that have many advantages, but they are too complex for a direct application to large systems.

2. Identifications of the Key Risk Factors

Suppose a set of data with $N+1$ observations on k asset or risk factor returns is summarized in a $(N+1)\times k$ covariance matrix $\boldsymbol{Y}$. Principal component analysis[3] will give up to k uncorrelated stationary variables, called the principal component of $\boldsymbol{Y}$, each component being a simple linear combination of the original returns as in (1) below. At the same time it is stated exactly how much of the total variation in the original system of risk factors is

[1] This work supported by Science and Technology Assistance Agency under contracts No. VEGA-1/3014/06, APVV-1/4024/07 and VEGA-0375-06

[2] VaR (Value at Risk) summarizes the expected maximum loss (or worst loss) over a target horizon within a given confidence interval [JOR00, p. 108].

[3] Alexander, C., O.: *Key Market Risk Factors: Identification and Applications.* The Q-Group Seminar on Risk, Florida, 2-5 April 2000. p.3

explained by each principal component, and the components are ordered according to the amount of variation they explain.

The first step in principal component analysis is to normalize the data in a $(N+1)\times k$ matrix $\boldsymbol{X}$, that represents the same variables as $\boldsymbol{Y}$, but in $\boldsymbol{X}$ each column is standardized to have mean zero and variance 1. So if the i-th risk factor or asset return in the system is y_i, then the normalized variables are $x_i=(y_i - \mu_i)/\sigma_i$, where μ_i a σ_i are the mean and standard deviation of y_i for $i=1, 2, ..., k$. Now let $\boldsymbol{W}$ be the matrix of eingenvectors of $\boldsymbol{X}^T\boldsymbol{X}$ and $\boldsymbol{\Lambda}$ be the associated diagonal matrix of eigenvalues Λ_i's, ordered according to decreasing magnitude of eigenvalue[4]. The principal components of $\boldsymbol{Y}$ are gave by the matrix of the order $(N+1)\times k$

$$\boldsymbol{P} = \boldsymbol{XW}. \tag{1}$$

Thus a linear transformation of the original risk factor returns has been made in such a way, that the transformed risk factors are orthogonal, that is, they have zero correlation[5].
The new risk factors are ordered by the amount of the variation they explain[6]. Hence only the new risk few, the most important factors may be chosen to represent the system as follows: Since $\boldsymbol{W}$ is orthogonal, (1) is equivalent to $\boldsymbol{X}=\boldsymbol{PW}^T$, that is

$$\boldsymbol{x}_i = \boldsymbol{w}_{i1}\boldsymbol{p}_1 + \boldsymbol{w}_{i2}\boldsymbol{p}_2 + ... + \boldsymbol{w}_{ik}\boldsymbol{p}_k. \tag{2}$$

The matrix $\boldsymbol{W}$ is called as the matrix of "weights factor". In terms of the original variables $\boldsymbol{Y}$ the representation (2) is equivalent to

$$\boldsymbol{Y}_i = \mu_i\boldsymbol{\omega}^*_{i1}\boldsymbol{p}_1 + \boldsymbol{\omega}^*_{i2}\boldsymbol{p}_2 + ... \boldsymbol{\omega}^*_{im}\boldsymbol{p}_m + \varepsilon_i, \tag{3}$$

where $\boldsymbol{\omega}^*_{ij} = \boldsymbol{\omega}^*_{ij}\boldsymbol{\sigma_i}$ and the error term in (3) picks up the approximation from using only the first m of the k principal components. These m principal components are the "*key*" risk factors of the system, and the rest of the variation is ascribed to "noise" in the error term. The representation (3) indicates how, covariance or scenario calculations are based only on the most important principal components, the effect may be easily translated back to the original system through a simple linear transformation[7].

3. Efficient Computation of Positive Semi-Definite Covariance Matrices

Since principal components are orthogonal their covariance matrix is simply the diagonal matrix of their variances. These variances can be quickly transformed into a covariance matrix of the original system using the factor weight as follows:
Taking variance of (3) gives

$$\boldsymbol{V} = \boldsymbol{ADA}^T + \boldsymbol{V}_{\boldsymbol{\varepsilon}}, \tag{4}$$

where $\boldsymbol{A} = (\boldsymbol{\omega}^*_{ij})$ is the $k\times m$ matrix of normalized factor weights,
$\boldsymbol{D}=\boldsymbol{diag}(\boldsymbol{V}(\boldsymbol{P}_1), ..., \boldsymbol{V}(\boldsymbol{P}_m))$ is the diagonal matrix of variances of principal components and $\boldsymbol{V}_{\boldsymbol{\varepsilon}}$ is the covariance matrix of the errors.
Ignoring $\boldsymbol{V}_{\boldsymbol{\varepsilon}}$ gives the approximation

$$\boldsymbol{V} \approx \boldsymbol{ADA}^T \tag{5}$$

with an accuracy that is controlled by choosing more or less components to represent the system. This show, that the full $k\times k$ covariance matrix of asset or risk factor returns $\boldsymbol{V}$ is obtained from a just a few estimated of the variances of the principal components.

---

[4] Thus $\boldsymbol{X}^T\boldsymbol{XW}=\boldsymbol{W\Lambda}$
[5] Note that $\boldsymbol{P}^T\boldsymbol{P}=\boldsymbol{W}^T\boldsymbol{X}^T\boldsymbol{XW}=\boldsymbol{W}^T\boldsymbol{W\Lambda}$, but $\boldsymbol{W}$ is an orthogonal matrix so $\boldsymbol{P}^T\boldsymbol{P}=\boldsymbol{\Lambda}$, a diagonal matrix.
[6] The proportion of the total variation in $\boldsymbol{X}$ that is explained by the m-th principal component and the column labeling in $\boldsymbol{W}$ has been chosen so that $\lambda_1>\lambda_2> ... >\lambda_k$.
[7] Jorion, P.: *Value at Risk: The Benchmark for Controlling Market Risk.* Blaklick, OH, USA: McGraw-Hill Professional Book Group, 2000, p. 179-181.

Note, that $\boldsymbol{V}$ will be positive semi-definite, but it may be not strictly positive definite unless $m=k$[8]. Although $\boldsymbol{D}$ is positive definite because it is a diagonal matrix with positive elements, there is nothing to guarantee that $\boldsymbol{ADA}^T$ will be positive definite when $m<k$. To see this write

$$\boldsymbol{x}^T\boldsymbol{ADA}^T\boldsymbol{x}=\boldsymbol{y}^T\boldsymbol{Dy}, \qquad (6)$$

where $\boldsymbol{A}^T\boldsymbol{x}=\boldsymbol{y}$. Since $\boldsymbol{y}$ can be zero for some non-zero $\boldsymbol{x}$, $\boldsymbol{x}^T\boldsymbol{ADA}^T\boldsymbol{x}$ will not be strictly positive for all non-zero $\boldsymbol{x}$. It may be zero, and so $\boldsymbol{ADA}^T$ is only positive semi-definite. When covariance matrices are based on (5) with $m<k$, they should be run through an eigenvalue check to ensure strict positive definiteness. However it is reasonable to expect that the approximation (5) will give a strictly positive definite covariance matrix if the representation (3) is made with a high degree of accuracy[9].

The advantages of using this type of orthogonal transformation for generate risk factor covariance matrices is clear. There is a very high degree of computational efficiency in calculating only m variances and covariances of the original system. Exponentially weighted moving averages of the squares and cross products of returns are a standard method for generating covariances matrices. But a limitation of this type of direct application of exponentially weighted moving averages is, that the covariance matrix is only guaranteed to be positive semi-definite if the same smoothing constant is used for all the data.

4. Modeling of multivariate scenarios for computing risk measures VaR

Many portfolios consist of the financial instruments, which pay some portion of their return in the future. Because the risk factors used to value these instruments vary over time, the value of the portfolio at a future date is uncertain. Many analysts use Value at Risk (VaR) as a measure of the potential profit or loss at a specific point in the future. Put simplify, the VaR for a portfolio is a quantile value of the distribution of the change in the portfolio value relative to the base case value. They are known several methods for calculating VaR [JOR00, JIL02]. In this paper, we use Monte Carlo simulation method.

Monte Carlo simulation is made of three steps:

- simulation of the future state of the world,
- state variable transformations, and
- pricing the portfolio.

Because the future state of the world is unknown, models of the state variables are used to forecast the future values of the state variables. Historical values of the state variables can also be used to develop possible future states but we will focus on market models for forecasting the future.

Let the state of the world is represented by a vector of state variables, for example, interest rates, exchange rates, and stock prices. Even with very good models for the state variables the forecasts of these values have a very wide confidence band. Instead of viewing the forecast as a point estimate, the forecast is viewed as a probability density. These models describe the multi-dimensional probability density function of the future. In order to get a picture of this multi-dimensional distribution and how it affects the value on a portfolio, the three steps of Monte Carlo simulation are repeated sufficient times to explore the state space. Usually 1000 to 10000 replications are used. The resulting values of the portfolio are sorted, and order statistics are used to obtain the VaR.

---

[8] A symmetric matrix $\boldsymbol{A}$ is positive definite if $\boldsymbol{x}^T\boldsymbol{Ax}>0$ for all non-zero $\boldsymbol{x}$. If $\boldsymbol{w}$ is a vector of portfolio weights and $\boldsymbol{V}$ is the covariance matrix of asset returns, then the portfolio variance is $\boldsymbol{w}^T\boldsymbol{Vw}$. So covariance matrices must always be positive definite, otherwise some portfolios may have non-positive variance.

[9] Alexander, C., O.: *Key Market Risk Factors: Identification and Applications.* The Q-Group Seminar on Risk, Florida, 2-5 April 2000. p. 6

In order for Monte Carlo simulation to be effective, the dimension of the risk factors should be as small as possible, and it is important for the models to be as accurate as possible. The dimension of the state space directly affects the number of replications that are required to explore the multi-dimensional space. The accuracy of the model determines how well the multi-dimensional probability density function describes the actual probabilities of future events. We may use a principle components analysis for reducing the dimension of the generated covariance matrix.

5. Results

Consider, that we are holding a portfolio composed of three types of the coupon bonds[10] (Euro Corporation, British Corporation and American Corporation). Risk factors influencing on this portfolio are the short-term (six months) LIBOR interest rates (EUR_6M, GBP_6M, USD_6M), the intermediate-term (twelve months) LIBOR interest rates (EUR_12M, GBP_12M, USD_12M)[11], and the exchange rates (EUR /USD, EUR/GBP)[12].

Our task is to determine VaR of our portfolio using Monte Carlo simulation method and compare it with VaR computed after applying PCA method on the risk factors.

The major steps are as follows:

Step 1: To collect monthly historical data of the k risk factors – k time series spanning $N+1$. We denote this data as $x_{i,0}, x_{i,1}, \ldots, x_{i,N}$, for $i=1, 2, \ldots, k$, where actual data are given by the $x_{i,N}$'s.

Step 2: Assuming $x_{i,j} \neq 0$, to compute the geometric rates of changes in the data:

$$r_{i,j} = \ln \frac{X_{i,j}}{X_{i,j-1}}, \tag{7}$$

for $i = 1, 2, \ldots, n$ and $j = 2, \ldots, N$, the $r_{i,1}, r_{i,2}, \ldots, r_{i,N}$ will be considered as elements belonging to a random variable r_i.

Step 3: To compute correlation matrix of the rates r_i (see (7)) of the risk factors. PCA method can be used if some correlation coefficients are significant[13].

Rates of the risk factors are significant and mutually correlated, in our example (see Table 1).

Step 4: To determine the number of m - principal components using Principal Components Analysis method.

If we use Kaiser's rule[14], then number of the principal components is equal to number of the eigenvalues of the correlation matrix greater to one.

If we use proportion criterion, then we take into account, what fraction of the variance is clarified by selected number of the principal component.

We see, from Table 2, that we can choose two (by Kaiser's rule) or three (by proportion criterion) principal components, in our example[15].

---

[10] Coupon bonds are bonds that entitle the bond's holder to receive cash payments in future time periods.

[11] The monthly's datas about interest rates follows from http://www.bba.org.uk/bba/jsp from January 2004 to April 2007

[12] The datas about exchange rates follows from http://www.nbs.sk from January 2004 to April 2007

[13] BOHDALOVÁ, M. - STANKOVIČOVÁ, I.*: Using the PCA in the Analyse of the risk Factors of the investment Portfolio*. In: Forum Statisticum Slovakum, 3/2006, p. 41-52, ISSN 1336-7420

[14] BOHDALOVÁ, M. - STANKOVIČOVÁ, I.*: Using the PCA in the Analyse of the risk Factors of the investment Portfolio*. In: Forum Statisticum Slovakum, 3/2006, s.41-52, ISSN 1336-7420

[15] We use SAS® EG v.4 software for principal components analysis.

Step 5: Firstly, to compute VaR using Monte Carlo simulation method for all risk factors. Secondly, to compute VaR using Monte Carlo simulation method for two and consequently for three principal components[16].

SAS® Risk Dimensions® software enable to use Monte Carlo simulation with dynamic risk factor modelling[17]. Dynamic risk factor modeling enables risk management practitioners to efficiently fit many risk factor models simultaneously. Risk factors models can be fitted using different distributional specifications, including non-normal distributions. This multivariate simulation process captures and maintains the dependence structure of the risk factors modelled separately. To accomplish this, the simulation engine uses a framework based on the statistical concept of a copula. A copula is a function that combines marginal distributions of the variables (risk factors) into a specific multivariate distribution in which all of its one-dimensional marginal are the cumulative distribution functions (CDF's) of the risk factors. We use mean and difference model to fitting our risks factors. First model - mean model assume that the rates of change of each risk factor is a constant. Second model – difference model assume that the exchange rates values depend on interest rates differentials. The reasoning is that countries with higher interest rates tend to attract capital, difference in interest rates between countries produce capital flows. These capital flows, in turn, generate movements in exchange rates.

The results summarize Table 3. We obtained negative 95 % VaR for all estimates. Negative VaR implies a gain in the portfolio is no less than 214,894.97 EUR in 95 % of series of repeated trials, if we take into account all risk factors and we fit risk factors by difference model. If we take into account 2 or 3 principal components we obtain comparable results (208,580.52 EUR and 210,732.16 EUR) for difference model. Analogously results are obtained, if we fit risks factors by mean models. The 95 % VaR is estimated for all risk factors on 275,454.92 EUR, for 2 PC on 276,587.94 EUR and for 3 PC on 276,736.48 EUR.

6. Conclusion

The main advantages of using the PCA method for computing VaR are well known. Principal components analysis attempts to find a series if independent combinations of the original variables, that provides the best explanation of the diagonal terms of the covariance matrix. The PCA method makes the computing process of VaR faster. It makes easier process of the generation multivariate random number scenarios in the Monte Carlo simulation. Both advantage have immediate applications to internal models for measuring market risk.

Table 1: Correlation matrix of the returns of the risk factors

Correlation Matrix								
	Ret_EUR_6M	**Ret_EUR_12M**	**Ret_GBP_6M**	**Ret_GBP_12M**	**Ret_USD_6M**	**Ret_USD_12M**	**Ret_EUR_GBP**	**Ret_EUR_USD**
Ret_EUR_6M	1.000	0.992	0.756	0.793	0.799	0.749	0.491	0.050
Ret_EUR_12M	0.992	1.000	0.738	0.789	0.822	0.788	0.464	0.082
Ret_GBP_6M	0.756	0.738	1.000	0.981	0.444	0.419	0.512	-0.285
Ret_GBP_12M	0.793	0.789	0.981	1.000	0.483	0.476	0.501	-0.259
Ret_USD_6M	0.799	0.822	0.444	0.483	1.000	0.979	0.513	0.419
Ret_USD_12M	0.749	0.788	0.419	0.476	0.979	1.000	0.479	0.422
Ret_EUR_GBP	0.491	0.464	0.512	0.501	0.513	0.479	1.000	0.353
Ret_EUR_USD	0.050	0.082	-0.285	-0.259	0.419	0.422	0.353	1.000

[16] We use SAS® Risk Dimensions® software for computing VaR.
[17] SAS® RISK Dimensions®: *Dynamic Risk factor Modeling Methodology*. White Paper, http://www.riskadvisory.com/pdfs/sasriskdimensionsriskfactor.pdf, visited 20.9.2006

Table 2: Eigenvalues of the Correlation Matrix

Eigenvalues of the Correlation Matrix				
	Eigenvalue	**Difference**	**Proportion**	**Cumulative**
1	**5.067**	**3.318**	**0.633**	**0.633**
2	**1.750**	**0.997**	**0.219**	**0.852**
3	0.753	0.527	0.094	0.946
4	0.226	0.058	0.028	0.974
5	0.168	0.139	0.021	0.995
6	0.028	0.022	0.004	0.999
7	0.006	0.004	0.001	1.000
8	0.002		0.000	1.000

Table 3: VaR

		Base Case Date	**Mark to Market Value (EUR)**	**At-Risk Value (VaR) (EUR)**
MC difference model	All risk faktors (8)	19APR2007	112,051.17	-214,894.97
	2 PC	19APR2007	112,051.17	-208,580.52
	3 PC	19APR2007	112,051.17	-210,732.16
MC mean model	All risk faktors (8)	19APR2007	112,051.17	-275,454.92
	2 PC	19APR2007	112,051.17	-276,587.94
	3 PC	19APR2007	112,051.17	-276,736.48

References

[ALE00] ALEXANDER, C. O.: *Key Market Risk Factors: Identification and Applications.* The Q-Group Seminar on Risk, Florida, 2-5 April 2000

[BOHS06] BOHDALOVÁ, M. - STANKOVIČOVÁ, I.*: Using the PCA in the Analyse of the risk Factors of the investment Portfolio*. In: Forum Statisticum Slovakum, 3/2006, p. 41-52, ISSN 1336-7420

[JIL00] JÍLEK, J.: *Finanční rizika.* Praha, GRADA Publishing, 2000, 635 p., ISBN 80-7169-579-3

[JOR00] JORION, P.: *Value at Risk: The Benchmark for Controlling Market Risk*. Blacklick, OH, USA:McGraw-Hill Professional Book Group, 2000, 535p.,ISBN 0-07-137921-5

[URB06] URBANÍKOVÁ, M.: *Využitie matematiky pri riadení rizík.* In *Acta Mathematica 9: Zborník zo VI. nitrianskej matematickej konferencie*. Nitra: FPV UKF, 2006, p. 193 - 202. ISBN 80-8094-036-3

SAS® documentations available on www.sas.com

SAS® RISK Dimensions®: *Dynamic Risk factor Modeling Methodology*. White Paper, available on http:www.riskadvisory.com/pdfs/sasriskdimensionsriskfactor. pdf

Address of authors:

RNDr. Mária Bohdalová, PhD.
Katedra informačných systémov
Fakulta managementu
Univerzita Komenského Bratislava
Odbojárov 10, P.O.Box 95, 820 05 Bratislava
maria.bohdalova@fm.uniba.sk

Ing. Iveta Stankovičová, PhD.
Katedra informačných systémov
Fakulta managementu
Univerzita Komenského Bratislava
Odbojárov 10, P.O.Box 95, 820 05 Bratislava
iveta.stankovicova@fm.uniba.sk

Využitie algoritmu GRG2 pri zisťovaní prechodových charakteristík pneumatických umelých svalov / The Using of Algorithm GRG2 for Searching the Transient Response of Pneumatic Artificial Muscles

Jana Boržíková

Fakulta výrobných technológií so sídlom v Prešove

Technická univerzita v Košiciach

Bayerova 1

080 01 Prešov

borzikova.jana@fvt.sk

Abstrakt

Článok sa zaoberá algoritmom GRG2, ktorý bol vyvinutý ako optimalizačný kód Generalized Reduced Gradient Methods. Algoritmus je súčasťou bežne dostupného MS Excelu ako Solver (Riešiteľ), no v súčasnosti sa vyskytujú na trhu viaceré novšie, prípadne rozšírené produkty firmy Frontline (Napríklad Premium Solver alebo celé platformy Premium Solver platform). Tieto novšie verzie sú rozšírené jednak čo do počtu možných premenných či podmienok, ale aj do typu funkcií a možných riešení problémov. Základný algoritmus GRG2 je v článku využitý na numerickú identifikáciu prechodovej funkcie ako dynamickej charakteristiky pneumatického umelého svalu. Po zistení prechodovej funkcie je možné nájsť lineárny dynamický systém, ktorý popisuje vybraný umelý sval.

Kľúčové slová: / Key words: numerická aproximácia, algoritmus GRG2, prechodová funkcia pneumatického umelého svalu / Numerical Approximation, Algorithm GRG2, Transient Response of Pneumatic Artificial Muscles

1 MS Excel a Solver

Pre riešenie numerických a úloh vo vysokoškolskej matematike, pri riešení úloh optimalizácie a lineárneho programovania sa využíva Solver (Riešiteľ). Tento známy nástroj využíva na riešenie spomenutých úloh algoritmus GRG2. Algoritmus bol pôvodne spracovaný v 70-tych rokoch v jazyku Fortran.
Algoritmus využívame v nelineárnych optimalizačných úlohách, kde minimalizujeme (maximalizujeme) funkciu $g_p(X)$ za obmedzujúcich podmienok:

$$g^{lb_i} \le g_i(X) \le g^{ub_i}, \text{ pre } i = 1, \ldots, m, \; i \neq p \qquad (1)$$

$$x^{lb_j} \leq x_j \leq x^{ub_j}, \text{ pre } j = 1, \ldots, n \qquad (2)$$

kde X je vektor n premenných, $x_1,...,x_n$ a funckia $g_1,...,g_n$ je závislá od X

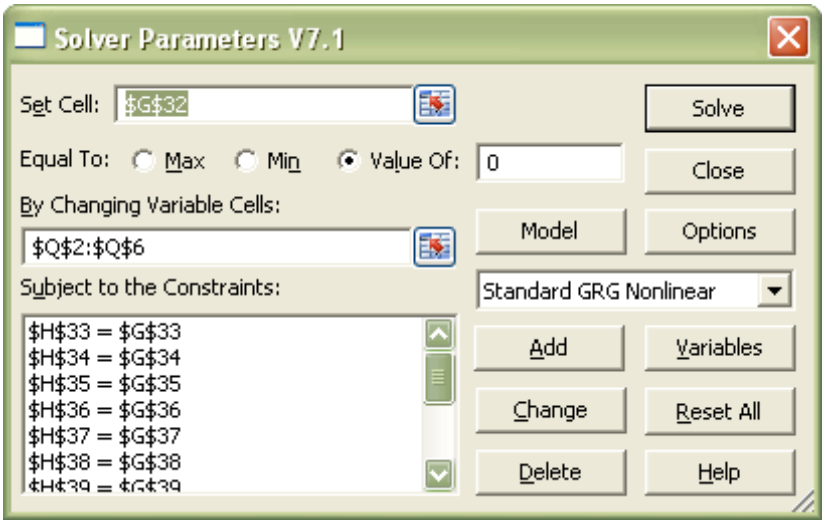
Ľubovoľná z daných funkcií (1) a (2) môže byť nelineárna, prípadne úplne vynechaná. Horná a dolná hranica premennej je ľubovoľná a ak existuje, nie je považovaná ako súčasť obmedzenia, ale je riešená samostatne. GRG2 používa prvé parciálne derivácie každej funkcie podľa jednotlivých premenných a tie sú vo výpočtoch nahradzované numerickým výpočtom (konečnými diferenciami). Po zadaní vstupných hodnôt algoritmus pracuje v dvoch fázach. Ak užívateľom zadané vstupné hodnoty premenných nevyhovujú podmienkam g_i, začne fáza I. Tá je ukončená správou, že je problém neriešiteľný, alebo správou o nájdenom riešení. Je potrebné brať na zreteľ, že správa o neriešiteľnosti je aj v prípadoch, že sa program zastavil na lokálnom minime funckie, no pritom problém má vhodné riešenie. V týchto prípadoch autori odporúčajú zadať iné počiatočné podmienky pre premenné a úlohu opäť riešiť.

Fáza II začína vhodným riešením nájdeným vo fáze I alebo užívateľom zabezpečenými vhodnými počiatočnými podmienkami a pokračuje optimalizáciou účelovej funkcie. V závere fázy II je ukončený cyklus optimalizácie a zverejnený výstup. Podrobnejšie môžete pozrieť priamo u autorov. (Lasdon L. S.)

2 Premium Solver Platform

Možnosti bežného riešiteľa sú obmedzené na 200 možných premenných a 10 podmienok. Firma Frontline ponúkla na trh komerčné riešenie Riešiteľa a to Premium Solver Platform.

Táto platforma rozširuje možnosti predchádzajúcich verzií a umožňuje riešiť hladké nelineárne problémy a nehladké optimalizačné problémy 500 premennými a 250 obmedzujúcimi podmienkami. Problémy lineárneho programovania môžu obsahovať až 2000 premenných a 8000 obmedzujúcich podmienok a dochádza k značnému urýchleniu výpočtov. Navyše umožňuje ďalšie nastavenia: „nematematické“ predpisy účelovej funkcie, výber numerickej derivácie a ďalšie iné. (Lípa, Hynek, 2004)

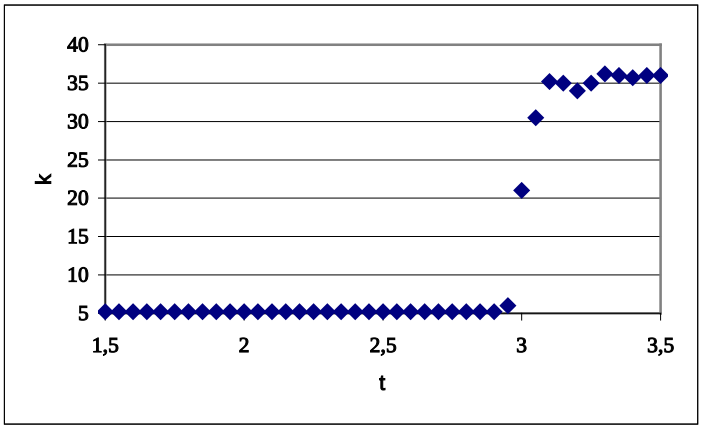


Obrázok 1 Premium Solver Platform

3 Aplikácia numerickej aproximácie pri zisťovaní prechodovej funkcie

Vrámci riešenia výskumných úloh v oblasti mechatroniky sme riešili problém aproximovať namerané hodnoty funkcie $k = f(t)$ (Obrázok 2) čím by sme dostali prechodovú funkciu ako dynamickú charakteristiku. Tá bude následne podkladom pre úplný popis dynamickej sústavy na báze pneumatických umelých svalov systémom diferenciálnych rovníc.

Hodnoty boli zaznamenané pre umelý sval pod záťažou (45 N ≅ 4,5 kg) a pri konštantnom tlaku napĺňania (p = 3,5 bar). Os x vyjadruje hodnoty času v sekundách



Obrázok 2 Namerané a normované vstupné hodnoty funkcie k = f(t).

z intervalu ⟨1,5; 3,5⟩. Na os y sú hodnoty kontrakcie svalu v %. Zo skúmaného javu je zrejmé, že interval $t \in \langle 1{,}5;\ 2{,}9\rangle$ je pásmo necitlivosti, t. j. umelý sval nereaguje na napúšťaný vzduch. Na intervale $t \in \langle 2{,}9;\ 3{,}5\rangle$ je pásmo napĺňania, t. j. umelý sval reaguje na napĺňanie vzduchom. Hodnota kontrakcie sa nakoniec ustáli na hodnote k_0 = 36.

Vzhľadom na tvar budúcej krivky sme vylúčili aproximáciu bežnými spojitými lineárnymi a nelineárnymi funkciami. Pristúpme k samotnému numerickému riešeniu. Podľa poznatkov analýzy funkcie jednej premennej je vhodným nahradením funkcia:

$$k = a\left(1 - e^{bt+c}.\sin(dt + f)\right) \tag{3}$$

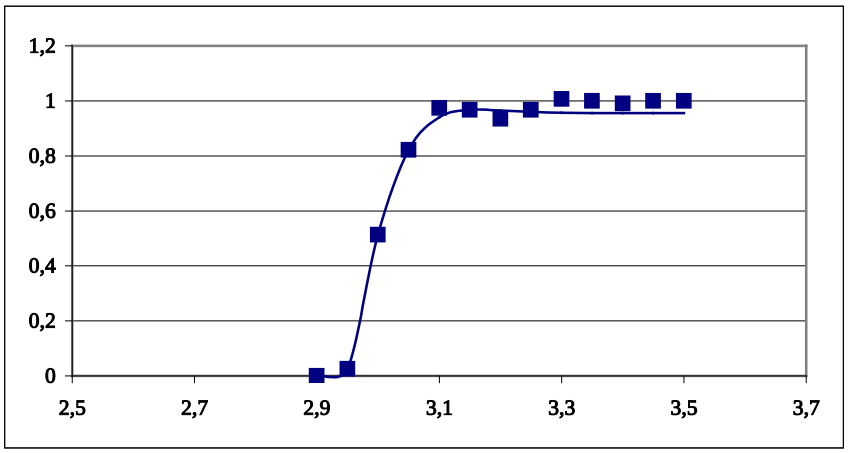
s neznámymi parametrami a, b, c, d, f. Pre zjednodušenie boli hodnoty k normované, aby bola ustálená hodnota rovná 1, t. j. bola zavedená substitúcia:

$$v_i = \frac{k_i - 5{,}2}{30{,}8} \text{ pre všetky } i \tag{4}$$

Následne boli hodnoty $[t_i - 2{,}9,\ v_i]$ dosadené do vzťahu (3), čím vznikla sústava m = 13 nelineárnych rovníc s n = 5 neznámymi. Sústava je predefinovaná, jej riešenie získame iba numerickými alebo optimalizačnými metódami a o kvalite budeme rozhodovať podľa vypočítaného indexu korelácie. Využijeme Riešiteľa v MS Excel. Nájdené riešenie

$$v = 1 - e^{(-18{,}2296.t + 1{,}055821)}.\sin(12{,}9171.t + 0{,}35549) \tag{5}$$

malo najvyšší index korelácie IK = 0,99 a teda je najvhodnejšie. (Obrázok 3)



Obrázok 3: Porovnanie nameraných a aproximovaných hodnôt podľa (3).

Využijúc poznatky z teórie riadenia, z hodnôt prechodovej charakteristiky vieme určiť parametre ξ a T:

$$y(t) = 1 - \frac{1}{\sqrt{1-\xi^2}}\, e^{-\frac{\xi}{T}.t}.\sin\left(\frac{\sqrt{1-\xi^2}}{T}.t + arctg\frac{\sqrt{1-\xi^2}}{\xi}\right) \tag{6}$$

na základe ktorých vieme nájsť prenos:

$$G(s) = \frac{y(s)}{u(s)} = \frac{1}{T_0^2 s^2 + 2\xi\, T_0 s + 1} \tag{7}$$

čo je vlastne obraz diferenciálnej rovnice v Laplaceovej transformácii, ktorá popisuje dynamický systém:

$$T_0^2\, y'' + 2\, T_0 \xi\, y' + y = u(t) \tag{6}$$

Keďže systém vykazuje aj dopravné oneskorenie, tak podľa vety o posunutí vzoru bude mať prenos získaný numericky podľa vzťahu (5) predpis:

$$G(s) = \frac{1}{0{,}000726027\, s^2 + 0{,}050520369\, s + 1}.e^{-2{,}9t}.$$

Čo je obraz lineárnej diferenciálnej rovnice 2 rádu s konštantnými koeficientmi.

4. Záver

Článok sa zaoberá výsledkom numerickej aproximácie prechodovej funkcie a možnosťami využitia rozšírených komerčných produktov napr. Premium Solver Platform. Vzhľadom na tieto nové možnosti odporúčam ďalej v projekte riešiť úlohu aj inými podpornými prostriedkami a algoritmami a vzhľadom na získaný typ diferenciálnej rovnice hľadať algoritmus na získanie nelineárnej diferenciálnej rovnice. Problém bol riešení s podporou inštitucionálnej úlohy č. 1/2007 „Výskum dynamických systémov a možnosti zdokonaľovania ich syntézy“.

Literatúra

1. Lasdon L. S.: *User's Guide for GRG2 Optimization Library*. http://www.maxthis.com/Grg2ug.htm #Windward
2. Lípa T., Hynek J. 2004. Excel + Evolver = Solver. In: *Kvalita, inovácia, prosperita.* Ročník VIII., č. 1/2004, s. 39-46. ISSN 1335-1745

Porovnanie kointegračnej a korelačnej analýzy s aplikáciou na časové rady miery inflácie krajín V4.

Eva Brestovanská

Univerzita Komenského, Fakulta managementu, Bratislava

1. ÚVOD DO PROBLEMATIKY.

Makroekonomické časové rady sú charakteristické tým, že ich vývoj je zviazaný určitými vzťahmi. Pri skúmaní vzťahov medzi časovými radmi pomocou regresnej analýzy často vzniká stav, ktorý sa označuje ako *zdanlivá regresia*. Je to situácia, keď regresný model sa potvrdí ako štatisticky významný, pričom časové rady spolu v skutočnosti nesúvisia. Riešením tohto problému je koncept kointegrácie, ktorý zaviedol Granger (pozri [1]) v prvej polovici 80. rokov minulého storočia. V tejto práci budem skúmať vzájomné vzťahy medzi časovými radmi miery inflácie krajín V4 klasicky pomocou regresnej (korelačnej) analýzy a porovnám ich s výsledkami nového prístupu založeného na kointegračnej analýze. Ďalej som skúmala, či existuje spoločný vývoj miery inflácie v krajinách V4. Časové rady obsahujú dáta od januára 1993 do apríla 2007, ako zdroj dát som použila internetovú stránku Organizácie pre hospodársku spoluprácu a rozvoj OECD http://www.oecd.org.

2. TEORETICKÉ ZÁKLADY Z KOINTEGRAČNEJ ANALÝZY.

2.1. Integrované a kointegrované procesy.

Nestacionárny časový rad, ktorý môžeme transformovať prvými diferenciami na stacionárny časový rad nazývame Integrovaný proces prvého rádu a označujeme ako I(1). Stacionárny proces označujeme I(0). Z definície Engle-Grangera (pozri [1]) pre k integrovaných vektorov obsahujúcich procesy $\{Y_{1,t}\}...\{Y_{k,t}\}$ platí: ak je každý z týchto procesov typu I(1) a ak existuje vektor β (k x 1) tak, že $\xi = \beta.\{\mathbf{Y}_t\}$ je typu I(0), potom sú tieto procesy kointegrované. Vektor β sa nazýva kointegračný vektor. **Princíp kointegrácie** sa v súčasnosti považuje za ústrednú myšlienku modelovania integrovaných časových radov najmä z nasledujúcich dôvodov (pozri [1]):

- Analýza vzťahov medzi integrovanými časovými radmi má zmysel len vtedy, ak sú tieto časové rady kointegrované, t. j. sú spojené spoločným stochastickým trendom. Ak tomu tak nie je, časové rady majú iný smer vývoja. Pri skúmaní vzťahov medzi takýmito časovými radmi pomocou regresnej analýzy vzniká tzv. zdanlivá regresia (pozri [3]). Test kointegrácie časových radov je teda súčasne metódou pre rozlíšenie medzi pravou a zdanlivou regresiou.
- Strednú hodnotu stacionárnej lineárnej kombinácie integrovaných časových radov je možné chápať ako ekvilibrium, ktoré spája uvažované časové rady.
- Kointegrované časové rady je možné popísať modelom korekcie chyby (error model correction), pomocou ktorého môžeme rozlíšiť vzťahy dlhodobé (medzi nediferencovanými procesmi) a krátkodobé (medzi diferencovanými procesmi). Tento model obsahuje parametre, ktoré charakterizujú mieru vychýlenia sa systému od dlhodobo prevažujúceho rovnovážneho stavu.

2.2. Dickey - Fullerov test stacionarity - (test jednotkového koreňa AR polynómu) (pozri [1, 4, 5]).

Tento test označujeme v skratke ADF a testujeme ním hypotézu H_0: časový rad je typu I(1), oproti H_1: časový rad je typu I(0). Uvažujme proces AR(p) s AR polynómom: $\phi_p(B) = 1 - \phi_1 B - \ldots - \phi_p B^p$. Ak je časový rad typu I(1), t. j. platí $\phi_p(1) = 0$, potom môžeme polynóm $\phi_p(B)$ písať v tvare:

$$\phi_p(B) = (1 - \phi_1 - \ldots - \phi_p) B^i + \phi^*_{p-1}(B)(1 - B) \text{ pre ľubovoľné } i = 1, 2, \ldots, p \qquad (1)$$

kde $\phi^*_{p-1}(B) = \phi^*_0 + \phi^*_1 B + \cdots + \phi^*_{p-1} B^{p-1}$ je polynóm p -1 rádu. Možno ukázať (pozri v [6]), že zavedením pomocnej regresnej funkcie

$$\Delta Y_t = \rho Y_{t-1} - \alpha^*_1 \Delta Y_{t-1} - \cdots - \alpha^*_{p-1} \Delta Y_{t-(p-1)} + \varepsilon_t, \text{ kde } \alpha^*_i = -\phi^*_i, \; i = 1, ..., p - 1 \qquad (2)$$

je možné ADF - test stacionarity previesť na T - test, ktorým testujeme významnosť regresného parametra, t.j. testujeme nulovú hypotézu H_0: $\rho = 0$, proti alternatívnej hypotéze H_1: $\rho < 0$. Fuller publikoval v [6] kritické hodnoty rozdelenia štatistiky T pre tri typy modelov. Pri záverečnom zhodnotení testu porovnáme vypočítanú štatistiku T pre koeficient ρ (pri Y_{t-1}) s tabuľkovou kritickou hodnotou**.** Ak štatistika T leží v intervale $t_{1-\alpha/2} > T > t_{\alpha/2}$ (kde $t_{1-\alpha/2}$ je dolná a $t_{\alpha/2}$ je horná tabuľková kritická hodnota), H_0 nezamietame, t.j. proces Y_t je typu I(1) a obsahuje stochastický trend.

2.3. KPSS test stacionarity (Kwiatkowski, Phillips, Schmidt, Shin) (pozri [1, 4, 5]) .

V 1992 bol v [5] publikovaný ďalší test stacionarity, ktorý má rovnakú funkciu ako predchádzajúci test. Avšak na rozdiel od neho testujeme nulovú hypotézu: H_0: časový rad je typu I(0) oproti H_1: časový rad je typu I(1). Testovaciu štatistiku η porovnáme s kritickou hodnotou η_{krit} pre KPSS test (tabelované napr. v [1]). Ak $\eta > \eta_{krit} \Rightarrow$ nulovú hypotézu H_0 zamietame $\Rightarrow \{Y_t\}$ je typu I(1), t. j. je nestacionárny a obsahuje stochastický trend.

2.4. VAR(p) (Vector Autoregression Model) (pozri [1, 4]).

V prípade m-rozmerných časových radov je m – rozmerný vektorový VAR(p) model daný vzťahom: $\mathbf{Y}_t = \boldsymbol{\mu} + \delta t + \boldsymbol{\phi}_1 \mathbf{Y}_{t-1} + \ldots + \boldsymbol{\phi}_p \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t$ (3)

kde $\boldsymbol{\mu} = (\mu_1, \cdots, \mu_m)'$ a $\delta = (\delta_1, \cdots, \delta_m)'$ sú m-rozmerné konštantné vektory, $\boldsymbol{\phi}_1$, …, $\boldsymbol{\phi}_p$ sú matice typu m x m, $\mathbf{Y}_t = (Y_{t1}, Y_{t2}, \ldots, Y_{tm})'$ je m-rozmerný náhodný vektor a $\boldsymbol{\varepsilon}_t$ je m-rozmerný proces bieleho šumu, pre ktorý $E(\boldsymbol{\varepsilon}_t) = \mathbf{0}$, $\text{cov}(\boldsymbol{\varepsilon}_t, \boldsymbol{\varepsilon}_s) = 0$ pre $t \neq s$, $\text{var}(\boldsymbol{\varepsilon}_t) = \boldsymbol{\Sigma}_{\boldsymbol{\varepsilon}}$ je kovariančná matica typu m x m. Model VAR(p) môžeme prepísať do tvaru vektorového modelu korekcie chyby VECM: $\Delta \mathbf{Y}_t = \boldsymbol{\mu} + \delta t + \boldsymbol{\Gamma}_1 \Delta \mathbf{Y}_{t-1} + ... + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{Y}_{t-p+1} + \boldsymbol{\Pi} \mathbf{Y}_{t-p} + \boldsymbol{\varepsilon}_t$. (3a)

kde $\boldsymbol{\Gamma}_i = (\boldsymbol{\phi}_1 + ... + \boldsymbol{\phi}_i) - \mathbf{E}_m$, i = 1, ..., p – 1 a.$\boldsymbol{\Pi} = (\boldsymbol{\phi}_1 + ... + \boldsymbol{\phi}_p) - \mathbf{E}_m$, ($\mathbf{E}_m$ je jednotková matica typu m x m). VECM (vector error correction model) obsahuje krátkodobé vzťahy medzi procesmi (vzťahy medzi diferencovanými stacionárnymi procesmi). Na druhej strane obsahuje aj vzťahy dlhodobé (t.j. vzťahy medzi nediferencovanými procesmi). Informácie o dlhodobých vzťahoch obsahuje matica $\boldsymbol{\Pi}$. Pritom VECM je konštruovaný tak, že tieto dva druhy vzťahov môžeme oddeliť a skúmať ich samostatne.

2.5. Kointegrácia v procese VAR(p).

Pre hodnosť matice $\boldsymbol{\Pi}$ platí $h(\boldsymbol{\Pi}) = r$, $0 \leq r \leq m$. Môžu nastať tri prípady:

- $h(\boldsymbol{\Pi}) = m \Rightarrow$ Matica $\boldsymbol{\Pi}$ je regulárna, nemá zmysel hovoriť o kointegrácii.

- h(**Π**) = 0 ⇒ Matica **Π** je nulová, medzi časovými radmi dlhodobý vzťah neexistuje.
- 0< h(**Π**) = r < m ⇒ Matica **Π** má hodnosť r. V tomto prípade model (3a) obsahuje nediferencovaný člen $\mathbf{Y}_{t-p}$, ale súčasne nemôžeme považovať proces $\{\mathbf{Y}_t\}$ za stacionárny. Pretože matica **Π** nie je nulová, môžeme medzi jednotlivými časovými radmi nájsť dlhodobý vzťah, o ktorý by sme individuálnym diferencovaním jednotlivých časových radov prišli. Niektoré časové rady, ktoré nie sú v dlhodobom vzťahu s inými časovými radmi môžeme stacionarizovať individuálnym diferencovaním. Ostanú časové rady, ktorých lineárne kombinácie s inými časovými radmi sú stacionárne, t.j. sú kointegrované. Prvé dve situácie sú na prvý pohľad zrejmé a ich vysvetlenie je logické. Tretiu situáciu objasnil C.W.J. Granger (pozri [2]) v známej Grangerovej vete.

2.6. Testovanie rádu kointegrácie, Johansenov test (pozri [1, 4]).

Z Grangerovej vety vieme, že hodnosť matice **Π** určuje aj počet kointegračných vektorov, preto nás v ďalšom kroku bude zaujímať práve určenie hodnosti matice **Π**. Jedným z testov určujúcich hodnosť matice **Π** je Johansenov test. Spočíva v rozšírení jednorozmerného ADF testu do mnohorozmerného prípadu za predpokladu, že ε_t je m-rozmerný Gaussovský proces bieleho šumu.

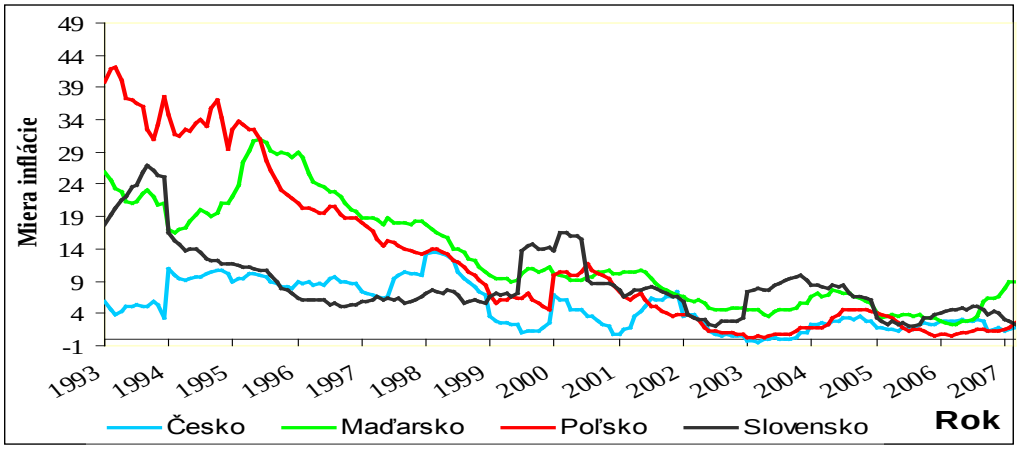
Prvým testom je test stopy matice $Trace_r$**:** test, že existuje najviac r kointegračných relácií, je test nulovej hypotézy H_0: Existuje najviac k kointegračných relácií, proti alternatívnej hypotéze H_1: Existuje (k + 1) a viac kointegračných relácií. Ak testovacia štatistika (v ďalšom texte označená $\mathbf{Trace_k}$) je väčšia ako tabelizovaná kritická hodnota (pozri napr. v [1, 4]), zamietame nulovú hypotézu H_0. Hodnosť matice **Π** je rovná r, ak prvá hypotéza H_0, ktorú nemôžeme zamietnuť je pre k = r. *Druhým testom je test maximálneho vlastného čísla* λ_{max}**:** test, že existuje práve r kointegračných relácií je test nulovej hypotézy H_0: Existuje najviac k kointegračných relácií, proti alternatívnej hypotéze H_1: Existuje práve (k + 1) kointegračných relácií. Ak je testovacia štatistika (v ďalšom texte označená $\boldsymbol{\lambda}_{\mathbf{max}}$) väčšia ako tabelizovaná kritická hodnota (pozri napr. v [1, 4]) zamietame nulovú hypotézu. Hodnosť matice **Π** je rovná r, ak prvá hypotéza H_0, ktorú nemôžeme zamietnuť je pre k = r-1.

Pri Johansenovej metóde hľadáme tie kombinácie prvkov náhodného vektora $\mathbf{Y}_t$, ktoré majú maximálnu parciálnu koreláciu so stacionárnymi premennými. Označme $\mathbf{M}_{ij}$ pre i, j = 0, 1, 2 súčiny momentových matíc, $S_{ij} = M_{ij} - M_{i1}M_{11}^{-1}M_{1j}$ pre i, j = 0, 2 sumy štvorcov reziduí. Vlastné čísla hľadáme z charakteristickej rovnice (pozri [1, 4]), $\left|\lambda \mathbf{S}_{22} - \mathbf{S}_{20}\mathbf{S}_{00}^{-1}\mathbf{S}_{02}\right| = 0$. Riešením sú vlastné čísla $\hat{\lambda}_1 > \cdots > \hat{\lambda}_r$ a vlastné vektory $\boldsymbol{\beta}_{\mathbf{1}, \dots,} \boldsymbol{\beta}_{\mathbf{r}}$. Ak máme r kointegračných vzťahov, potom ich získame ako $\boldsymbol{\beta}'_{\mathbf{1}}.\mathbf{Y}_{\mathbf{t}}, \dots, \boldsymbol{\beta}'_{\mathbf{r}}.\mathbf{Y}_{\mathbf{t}}$.

2.7. Odhad spoločného stochastického trendu, Gonzalo-Grangerová metóda (pozri [1, 2, 4]) .

Keď máme r kointegračných vzťahov v systéme m stochastických procesov, znamená to, že existuje spoločných (m - r) stochastických trendov. Jednou z metód, ktorá ich umožní ľahko získať je Gonzalo-Grangerová metóda. Pri Gonzalo-Grangerovej metóde hľadáme naopak tie kombinácie, ktoré majú minimálnu koreláciu. Vlastné čísla hľadáme z charakteristickej rovnice $\left|\lambda \mathbf{S}_{00} - \mathbf{S}_{02}\mathbf{S}_{22}^{-1}\mathbf{S}_{20}\right| = 0$ (bližšie pozri [1], [4]). Riešením sú tie isté vlastné čísla $\hat{\lambda}_1 > \cdots > \hat{\lambda}_m$, ale vlastné vektory $\hat{\mathbf{w}}_1, \cdots, \hat{\mathbf{w}}_m$ sú už iné. Gonzálo-Granger ukázali, že ak máme r kointegračných vzťahov, potom (m - r) stochastických trendov získame $\hat{w}'_{r+1}.Y_t, \dots, \hat{w}'_m.Y_t$.

3. PRAKTICKÁ ČASŤ PRÁCE.



Obr. 1. Vývoj miery inflácie v krajinách V4.

Analýza vzťahu medzi časovými radmi miery inflácie krajín V4 skúmaním kointegrácie.

Skúmať kointegráciu má zmysel len v prípade, že časové rady sú nestacionárne typu I(1), t. j. obsahujú stochastický trend. Prvým krokom je preto testovanie stacionarity jednotlivých časových radov pomocou rozšíreného Dickey – Fullerovho ADF testu.

Označme časové rady miery inflácie krajín V4 nasledujúcim spôsobom: $Y_{1,t}$ = Česko, $Y_{2,t}$ = Maďarsko, $Y_{3,t}$ = Poľsko, $Y_{4,t}$ = Slovensko. Pre všetky štyri časové rady som vypočítala hodnoty informačných kritérií AIC a BIC pre autoregresný model AR(p), $p \leq 5$, na základe ktorých som pre časové rady $Y_{1,t}$,$Y_{2,t}$,$Y_{3,t}$ zvolila optimálny rád p = 1 a pre $Y_{4,t}$ rád p = 2 (tabuľka P1 v prílohe).

Ďalším krokom je testovanie hypotézy H_0: Časový rad miery inflácie $Y_{i,t}$ obsahuje stochasticky trend, t.j. je typu I(1), proti hypotéze H_1: Časový rad je typu I(0) (pre i = 1, 2, 3, 4). Pre všetky štyri krajiny môžeme konštatovať, že $t_{1-\alpha/2} > T > t_{\alpha/2}$, z čoho vyplýva, že všetky štyri časové rady obsahujú stochastický trend. Má teda zmysel zisťovať, či aspoň niektoré z nich obsahujú spoločný stochastický trend, t.j. skúmať ich kointegraciu (tabuľka P2 v prílohe).

Pretože sú všetky skúmané časové rady typu I(1), ďalším krokom je výstavba EC modelu (3a). Pomocou informačných kritérii AIC a BIC (tabuľka P3 v prílohe) som určila optimálny rád p = 2 vektorového modelu VAR(p), ktorý som prepísala do tvaru modelu korekcie chyby VECM. Informácie o dlhodobých vzťahoch, možnej kointegrácii medzi prvkami vektora $\mathbf{Y}_t$ obsahuje matica $\mathbf{\Pi}$, ktorej hodnosť indikuje, či systém obsahuje určité kointegračné vzťahy.

Na určenie hodnosti matice $\mathbf{\Pi}$ som použila Johansenov test. Pri potvrdení existencie kointegračného vzťahu som spoločný stochastický trend určila použitím Gonzalo - Grangeovej metódy.

Johansenov test pre stopu matice: Pretože vypočítaná štatistika Trace $_{k-1}$ je pre k = 1 a 2 (r = k – 1 = 0 a 1) väčšia ako kritická hodnota, hypotézu H_0 zamietame. Pre k = 3 (r = k -1 = 2) je už vypočítaná štatistika $Trace_2$ menšia ako korešpondujúce kritické hodnoty (pozri tabuľku č.1.), preto H_0 nezamietame, tj. existujú najviac dva kointegračne vzťahy.

Druhým testom je test maximalneho vlastného čísla: H_0: Existuje najviac k - 1 kointegračných vektorov, proti H_1: Existuje práve k kointegračných vektorov. V tomto teste je už pre k = 2 vypočítaná štatistika $\lambda_{k,max}$ menšia ako korešpondujúce kritické hodnoty (pozri tabuľku č. 2), preto hypotézu H_0 nezamietame. Existuje teda práve jeden kointegračný vektor, čiže medzi jednotlivými časovými radmi miery inflácie **existuje práve jeden kointegračný vzťah.**

	Prvý test: $Trace_{k-1}$ - Kritické hodnoty			Vypočítanéštatistiky
k-1	α =0.10	α =0.05	α =0.01	$Trace_{k-1}$
0	7.52	9.24	12.97	**39.6333**
1	17.85	19.96	24.6	**24.9405**
2	**32.00**	**34.91**	**41.07**	**11.2136**
3	49.65	53.12	60.16	**4.29398**

Tabuľka č.1. Kritické hodnoty a vypočítané štatistiky pre test stopy matice.

	Druhý test: $\lambda_{k,max}$ - Kritické hodnoty			Vypočítané štatistiky
k	α =0.10	α =0.05	α =0.01	$\lambda_{k,max}$
k=1	7.52	9.24	12.97	14.6928
k=2	**13.75**	**15.67**	**20.2**	**13.7269**
k=3	19.77	22.00	26.81	6.91958
k=4	25.56	28.14	33.24	4.29398

Tabuľka č.2. Kritické hodnoty a vypočítané štatistiky, test maximalneho vlastného čísla.

Kointegračné vzťahy možno napísať vo všeobecnom tvare:
$\xi_k = \beta_{k,1}.\ Y_{1,t} + \beta_{k,2}\, Y_{2,t} + \beta_{k,3}\, Y_{3,t} + \beta_{k,4}\, Y_{4,t}$ kde $Y_{1,t}$ = Česko, $Y_{2,t}$ = Maďarsko, $Y_{3,t}$ = Poľsko, $Y_{4,t}$ = Slovensko sú časové rady miery inflácie, vektor $\mathbf{Y}_t$ je tvorený týmito štyrmi procesmi a kointegračný vektor $\boldsymbol{\beta}_k$ je k-ty vlastný vektor (pre k = 1, ..., r). V našom prípade je k = 1, teda kointegračný vektor je vlastný vektor $\boldsymbol{\beta}_1 = (\beta_{1,1}, \beta_{1,2}\, \beta_{1,3}\, \beta_{1,4})$.

k	Vlastné čísla λ_k	$\beta_{k,1}$	$\beta_{k,2}$	$\beta_{k,3}$	$\beta_{k,4}$
1	**0.0861969**	**-0.912385**	**0.233044**	**0.189516**	**-0.27808**
2	0.0807657	0.259466	-0.52785	0.532133	-0.609005
3	0.041563	0.253008	0.852219	-0.428444	-0.161694
4	0.0259995	-0.582731	-0.0072939	0.0256129	-0.812229

Tabuľka č. 3. Vlastné čísla a vektory (Johansenova metóda)

Z tabuľky č.3.vypočítame a vykreslíme stacionárnu linearnu kombináciu (obrázok O1 v prílohe): **Kointegračný vzťah : ξ = -0.912*$Y_{1,t}$ + 0.233*$Y_{2,t}$ + 0.1895*$Y_{3,t}$ -0.278*$Y_{4,t}$.**

Odtiaľ $Y_{1,t} = +\frac{0.23}{0.91} Y_{2,t} + \frac{0.19}{0.91} Y_{3,t} - \frac{0.28}{0.91} Y_{4,t} + \frac{\xi}{0{,}91} \Rightarrow$

$Y_{1,t}$ = 0.253*$Y_{2,t}$ + 0.21*$Y_{3,t}$ - 0.31*$Y_{4,t}$ + ε_t.
Dominantnou krajinou v kointegračnom vzťahu sa ukazuje Česko.

Testovanie stacionarity kointegračných relácií KPSS testom.

Pri skúmaní mnohorozmerných časových radov je najzaujímavejší prípad, keď kointegračný vektor vedie k stacionárnej lineárnej kombinácií, t.j. $\xi = \boldsymbol{\beta}^T.\mathbf{Y}_t$, ξ je typu I(0). Testujeme nulovú hypotézu H_0: lineárna kombinácia $\boldsymbol{\beta}^T.\mathbf{Y}_t$ je typu I(0), je stacionárna okolo konštanty Na hladine významnosti 0.01 a 0.05 vypočítaná štatistika η = 0.4 je menšia ako kritická hodnota $KH_{0.01}$ = 0.74 a $KH_{0.05}$ = 0.46 (Tabuľka P4 v prílohe), preto nezamietame hypotézu $H_0 \Rightarrow$ kointegračný vzťah ξ je typu I(0).

Výpočet spoločného stochastického trendu pre mieru inflácie krajín V4, Gonzalo-Grangerová metóda.

Vypočítame si ešte tri nezavislé stochastické trendy, ktoré získame ako lineárnu kombináciu $\hat{\mathbf{w}}_k^T . Y_t$, kde $\hat{w}_k = (w_{k,1}, w_{k,2}, w_{k,3}, w_{k,4})$ pre k=2, 3, 4; $\mathbf{Y}_t = (Y_{1,t}, Y_{2,t}, Y_{3,t}, Y_{4,t})$.

	Vlastné čísla	Vlastné vektory			
k	λ_k	$w_{k,1}$	$w_{k,2}$	$w_{k,3}$	$w_{k,4}$
1	0.0861969	0.771499	0.253857	-0.461877	0.356392
2	0.0807657	-0.465732	0.691452	-0.537484	-0.12688
3	0.041563	0.213446	0.905578	0.323618	-0.172163
4	0.0259995	0.11521	0.116341	-0.239239	-0.957056

Tabuľka č. 4. Vlastné čísla a vektor (*Gonzalo-Grangerová metóda*).

Prvý stochastický trend ST_1 (obrázok O2 v prílohe):

$ST_1 = \hat{w}_2^T \mathbf{Y}_t = -0.465732*Y_{1,t} + 0.691452*Y_{2,t} - 0.537484*Y_{3,t} - 0.12688*Y_{4,t} + \varepsilon_t$;

Druhý stochastický trend ST_2 (obrázok O3 v prílohe):

$ST_2 = \hat{w}_3^T \mathbf{Y}_t = 0.213446*Y_{1,t} + 0.905578*Y_{2,t} + 0.323618*Y_{3,t} - 0.172163*Y_{4,t} + \varepsilon_t$;

Tretí stochastický trend ST_3 (obrázok O4 v prílohe):

$ST_3 = \hat{w}_4^T \mathbf{Y}_t = 0.11521*Y_{1,t} + 0.116341*Y_{2,t} - 0.239239*Y_{3,t} - 0.957056*Y_{4,t} + \varepsilon_t$.

Analýza vzťahu medzi časovými radmi miery inflácie krajín Visegrádskej štvorky pomocou regresie.

Rovnica odhadnutého modelu (obrázok O5 v prílohe) je:

$Y_{1,t} = a*Y_{2,t} + b*Y_{3,t} - c*Y_{4,t} + d + \varepsilon_t$.

$Y_{1,t} = 0.239963*Y_{2,t} + 0.130126*Y_{3,t} - 0.257622*Y_{4,t} + \varepsilon_t$.

Keďže p-hodnota $<10^{-6}$ v ANOVA tabuľke (obrázok O5 v prílohe) je menšia ako 0.05 (0.01), hypotézu H_0 zamietame, model je štatisticky významný. Hodnota R- squared = 60.9643% t.j. model možno považovať za úspešný. Toto tvrdenie potvrdzuje aj významnosť parametrov v linearnom modeli, p-hodnota $<10^{-6}$ je menšia ako hladina významnosti 0.05 a 0.01. Parametre tohto regresného modelu sú porovnatelné s parametrami modelu odvodeného z kointegračného vzťahu, aj keď regresný model mierne podhodnocuje parametre.

Analýza vzťahu medzi časovými radmi miery inflácie krajín Visegrádskej štvorky po trojiciach.

Bol zopakovaný rovnaký typ analýzy pre všetky štyri kombinácie troch elementov subvektorov $\mathbf{Y_t}$. Pre všetky štyri triplety existuje práve jeden signifikantný kointegračný vzťah. Triplety vektorovo možno napísať nasledovne: bez Česka = $(Y_{2,t}, Y_{3,t}, Y_{4,t})$, bez Maďarska = $(Y_{1,t}, Y_{3,t}, Y_{4,t})$, bez Poľska = $(Y_{1,t}, Y_{2,t}, Y_{4,t})$, bez Slovenska = $(Y_{1,t}, Y_{2,t}, Y_{3,t})$.

4. ZÁVER.

Skúmať kointegráciu má zmysel len v prípade, že časové rady sú nestacionárne typu I(1), t. j. obsahujú stochastický trend. Prvým krokom bolo testovanie stacionarity jednotlivých časových radov pomocou rozšíreného Dickey – Fullerovho testu . Všetky štyri časové rady boli typu I(1), t.j. obsahovali stochastický trend. Malo teda zmysel zisťovať, či aspoň niektoré z nich obsahujú spoločný stochastický trend, t.j. skúmať ich kointegraciu. Použitím Johansenovho testu sa potvrdila existencia práve jedného kointegračného vzťahu.

Následne som Gonzalo - Grangerovou metódou vypočítala tri stochastické trendy. Kointegračný vzťah som na záver porovnala s prislúchajúcim regresným vzťahom.

Model pre štvoricu krajín V4, vektor $\mathbf{Y_t} = (Y_{1,t}, Y_{2,t}, Y_{3,t}, Y_{4,t})$.
Kointegračný vzťah: $Y_{1,t} = 0.25*Y_{2,t} + 0.21*Y_{3,t} - 0.31*Y_{4,t} + \varepsilon_t$.
Regresný vzťah: $Y_{1,t} = 0.24*Y_{2,t} + 0.13*Y_{3,t} - 0.26*Y_{4,t} + \varepsilon_t$, hodnota R- squared = 61% .
Model pre triplet bez Slovenska, vektor $(Y_{1,t}, Y_{2,t}, Y_{3,t})$.
Kointegračný vzťah: $Y_{1,t} = 0.325*Y_{2,t} - 0.095*Y_{3,t} + \varepsilon_t$.
Regresný vzťah: $Y_{1,t} = 0.339*Y_{2,t} - 0.003*Y_{3,t} + \varepsilon_t$, hodnota R- squared = 54% .
Model pre triplet bez Maďarska, vektor $(Y_{1,t}, Y_{3,t}, Y_{4,t})$.
Kointegračný vzťah : $Y_{4,t} = -0.823*Y_{1,t} + 0.640*Y_{3,t} + \varepsilon_t$.
Regresný vzťah: $Y_{4,t} = -0.694*Y_{1,t} + 0.430*Y_{3,t} + \varepsilon_t$, hodnota R- squared = 59% .
Model pre triplet bez Česka, vektor $(Y_{2,t}, Y_{3,t}, Y_{4,t})$.
Kointegračný vzťah: $Y_{3,t} = 0.914*Y_{2,t} + 0.891*Y_{4,t} + \varepsilon_t$.
Regresný vzťah: $Y_{3,t} = 1.066*Y_{2,t} + 0.822*Y_{4,t} + \varepsilon_t$, hodnota R- squared = 86%.
Model pre triplet bez Poľska, vektor $(Y_{1,t}, Y_{2,t}, Y_{4,t})$.
Kointegračný vzťah: $Y_{1,t} = 0.351*Y_{2,t} - 0.002*Y_{4,t} + \varepsilon_t$.
Regresný vzťah: $Y_{1,t} = 0.378*Y_{2,t} - 0.015*Y_{4,t} + \varepsilon_t$, hodnota R- squared = 62%.

Pri vysokých hodnotách miery R-squared sa parametre regresného modelu približovali k parametrom kointegračného modelu, parametre oboch modelov boli takmer rovnaké. Blízkosť týchto modelov je zrejme ovplyvnená hodnotou R- squared. Predpokladám, že totožnosť oboch modelov pravdepodobne nastane pre hodnotou R- squared blízku 100%. V skúmaných modeloch v tejto práci sa empiricky preukázala vzájomná prepojenosť medzi touto mierou, poprípade korelačným koeficientom a blízkosťou parametrov kointegračného a regresného vzťahu.

5. PRÍLOHA.

	Y_4=Slovensko		Y_2 =Maďarsko		Y_3 =Poľsko		Y_1 =Česko	
	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
1	0.778595	0.798145	**0.835186**	**0.854737**	**1.93473**	**1.95428**	**0.149615**	**0.169165**
2	**0.745434**	**0.784534**	0.840405	0.879506	1.93527	1.97437	0.16147	0.200571
3	0.749362	0.808013	0.852134	0.910785	1.94585	2.0045	0.167322	0.225973

Tabuľka P1. Hodnoty AIC, BIC kritéria pre AR(p) modelu, optimálny rád modelu.

KH pre dĺžku časového radu m = 250				Štatistika T pre koeficient ρ (pri Y_{t-1})			
$KH_{0.025}$	$KH_{0.05}$	$KH_{0.95}$	$KH_{0.975}$	Slovensko	Poľsko	Maďarsko	Česko
-3.14	-2.88	-0.06	0.24	-0.012043	-0.0051624	-0.0215008	-0.007010

Tabuľka P2. Výsledky Dickey – Fullerovho testu - model s konštantou bez trendu.

AIC	1.97017	**1.78251**	1.8902	1.95534	2.06397	2.15532
BIC	**2.2726**	2.38736	2.79747	3.16504	3.5761	3.96987

Tabuľka P3. Hodnoty AIC, BIC kritérií, optimálny rad modelu VAR(p) pre p=1, 2, ..., 6.

$KH_{0,01}$	$KH_{0,05}$	$KH_{0,1}$	štatistika η
0.74	0.46	0.35	0.4

Parameter	Estimate	Standard Error	T Statistic	P-Value
madarsko	0,239963	0,0490107	4,89613	0,0000
polsko	0,130126	0,0395233	3,29238	0,0012
slovensko	-0.257622	0.049683	-5.18531	0.0000

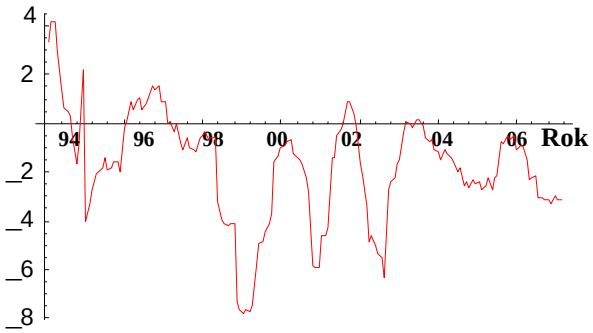
Analysis of Variance

Source	Sum of Squares	Df	Mean Square	F-Ratio	P-Value
Model	1338.96	3	446.32	83,29	0,0000
Residual	857,342	160	5,35839		
Total (Corr.)	2196,3	163			

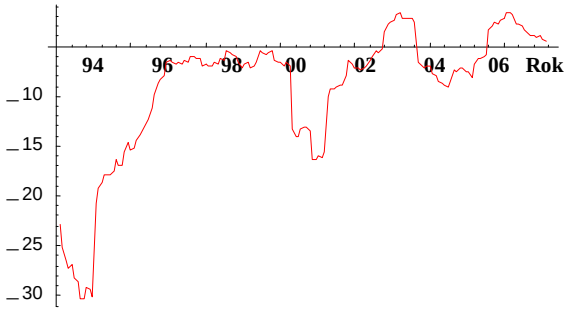
R-squared = 60.9643 percent

Obrázok O5. Regresná analýza medzi časovými radmi miery inflácie krajín V4.

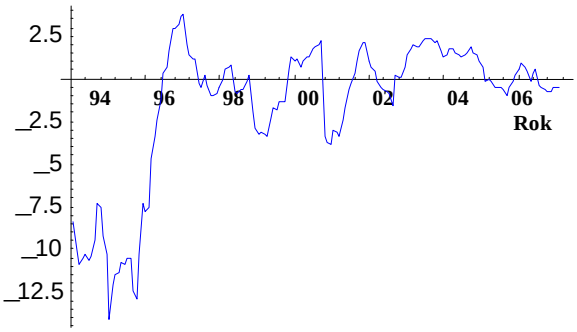
Tabuľka P4. Kritické hodnoty η_{krit} , štatistika η - KPSS test



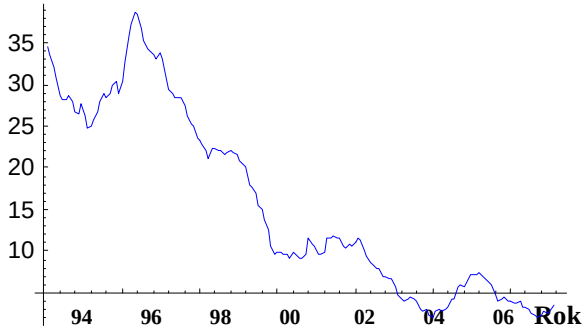
Obr. O1. Kointegračný vzťah ξ.



Obr. O2. Spoločný stochastický trend ST_1.



Obr. O3. Spoločný stochastický trend ST_2.



Obr. O4. Spoločný stochastický trend ST_3.

6. LITERATÚRA.

[1] Arlt, J. (1999), *Moderní metódy modelování ekonomických časových řad*, GRADA Publishing

[2] Granger, C. W. J., Teräsvirta, T. (1993), *Modelling nonlinear economic relationships*, Oxford University Press.

[3] Granger, C., and P. Newbold (1974), *Spurious Regression in Econometrics*, Journal of Econometrics, 2, 111-120

[4] Franses, P. H. (1998), *Time series models for business and economic forecasting*, Cambridge University Press

[5] Kwiatkowski, D., P.C.B. Phillips, P. Schmidt, Y. Shin, (1992), *Testing the Null Hypothesis of stationarity against Alternative of a Unit Root*, Journal of Econometrics, 73

[6] Dickey, D. A., W. A. Fuller, (1979), *Distribitions of the Estimators for Autoregressive Time Series with a Unit Root*, Journal of the American Statistical Association, 74.

ROZDĚLENÍ FÁZOVÉHO TYPU A JEHO APLIKACE VE SPOLEHLIVOSTI

GEJZA DOHNAL

ABSTRAKT. *Při analýze stochastických modelů narážíme často na narůstající složitost podmíněných pravděpodobnostních rozdělení, která je největší překážkou pro nalezení explicitního řešení. Dokonce i v případě i relativně velmi jednoduchých stochastických modelů nejsme mnohdy schopni získat analytické řešení. Tento problém lze odstranit použitím "rozdělení fázového typu", neboli PH-rozdělení. Třídu PH-rozdělení zavedl Neuts v roce 1975 a od té doby se stala velice užitečnou pro stochastické modelování. V tomto příspěvku se budeme zabývat modelem stárnoucího opravitelného systému. Aproximací rozdělení dob mezi poruchami PH rozdělením jsme schopni najít stacionární rozdělení systému a spočítat řadu spolehlivostních charakteristik.*

Klíčová slova: *Rozdělení fázového typu, markovský proces, spolehlivost, opravitelný systém, pohotovost, intenzita poruch, stacionérní rozdělení.*

1. ROZDĚLENÍ FÁZOVÉHO TYPU

Uvažujme markovský proces s $m < +\infty$ přechodnými stavy (představujícími jednotlivé "fáze" při vývoji v čase) a jedním absorpčním stavem. Počáteční rozdělení označme (α, α_{m+1}), kde α je řádkový m-dimensionální vektor. Matici intenzit přechodu tohoto procesu označme

$$\begin{pmatrix} \mathbf{T} & \mathbf{T}^0 \\ \mathbf{0}' & 0 \end{pmatrix},$$

kde $\mathbf{0}$ je sloupcový nulový vektor, $\mathbf{T}$ je regulární matice řádu m, obsahující infinitezimální generátory markovova procesu odpovídající jeho přechodovým stavům (generátor fázového typu) a platí, že $-\mathbf{T}\,\mathbf{e} = \mathbf{T}^0 \geq \mathbf{0}$. V dalším textu budeme ještě používat označení $\mathbf{I}$ pro jednotkovou matici, $\mathbf{O}$ matici samých nul a $\mathbf{e}$ sloupcový vektor samých jedniček (všechny tyto matice a vektory nechť mají odpovídající rozměry, vycházející z kontextu).

Definice 1. Je-li W náhodná veličina, reprezentující dobu do absorpce ve výše popsaném konečném markovově řetězci, potom její rozdělení pravděpodobnosti F nazveme rozdělením fázového typu (PH-rozdělení) s reprezentací $(\alpha, \mathbf{T})$ a v dalším jej budeme označovat jako $PH(m, \alpha, \mathbf{T})$.

Distribuční funkce rozdělení $PH(m, \alpha, \mathbf{T})$ je definována vztahem

$$F(w) = \begin{cases} 1 - \alpha \exp(\mathbf{T}w)\mathbf{e}, & w \geq 0, \\ 0, & w < 0, \end{cases}.$$

$F(w)$ je spojitá, pokud $\alpha\mathbf{e} = 1$. Hustota rozdělení $PH(m, \alpha, \mathbf{T})$ má tvar

$$f(w) = \begin{cases} \alpha \exp(\mathbf{T}w)\mathbf{T}^0\mathbf{e}, & w \geq 0, \\ 0, & w < 0, \end{cases}.$$

Je-li T regulární, má náhodná veličina W všechny momenty konečné a

$$m_k = EW^k = (-1)^k k! \alpha \mathbf{T}^{-k}\mathbf{e}, \qquad k \in N.$$

Nejjednodušším příkladem PH-rozdělení jsou směsi a konvoluce exponenciálních rozdělení (speciálně Erlangovo rozdělení). Obecněji, třída PH-rozdělení zahrnuje rozdělení libovolného sériově/paralelního uspořádání prvků s exponenciální dobou do poruchy, případně i se zpětnou vazbou. Třída PH-rozdělení je podtřídou takzvaných Coxových rozdělení, to jest všech rozdělení na $\langle 0, \infty)$, jejichž laplaceova transformace m a tvar racionální funkce (Cox, 1953). Ukazuje se, že třída PH-rozdělení je hustá (pro m nabývající libovolnou konečnou hodnotu) v množině rozdělení pravděpodobností na $[0, \infty)$ a tedy v důsledku toho lze jakékoli rozdělení na $\langle 0, \infty)$ aproximovat libovolně přesně rozdělením fázového typu.

Jednou z metod nalezení aproximujícího PH rozdělení je EM-algoritmus. Ten byl původně navržen pro iterativní výpočet maximálně věrohodného odhadu z pozorovaných dat, která jsou pouze částí rozsáhlejšího experimentu. Zobecněním EM algoritmu na nekonečné množství dat - kompletní distribuci, dostaneme iterativní algoritmus pro aproximaci spojitého rozdělení. Označíme-li f původní, aproximovanou hustotu, pak jeho použitím nalezneme rozdělení s hustotou h, které má minimální tzv. Kullback-Leiblerovu vzdálenost

$$I(f, h) = \int \ln \frac{f(x)}{h(x)} \mu dx.$$

Podrobný popis aplikace EM algoritmu s příklady lze nalézt v [3].

2. Stárnoucí opravitelné systémy

V této části se budeme zabývat opravitelným zařízením, jehož opravy nejsou obnovou (po opravě není zařízení v takovém stavu, jako když bylo nové). To lze modelovat zavedením předpokladu, že rozdělení jednotlivých dob provozu je ovlivňováno počtem předchozích oprav. Doby oprav budeme modelovat jako nezávislé, stejně rozdělené náhodné veličiny s PH-rozdělením. Uvažujeme tři typy poruch: náhodné (z vnějších příčin), které mohou být opravitelné i neopravitelné a poruchy

vzniklé opotřebením (vnitřní příčiny), které jsou vždy pouze neopravitelné. Na počátku je instalováno nové funkční zařízení. Nastane-li porucha, je zařízení opraveno a po opravě se vrací do funkčního stavu, jehož doba trvání má stejné rozdělení jako pro nové zařízení, pouze se mění měřítko času podle počtu předchozích oprav, což odpovídá geometrickému procesu (čas běží "rychleji"). Po každé neopravitelné poruše je zařízení obnoveno (bude ve stavu jako nové zařízení).

Definice 2. Řekneme, že posloupnost náhodných veličin $\{X_n, n = 1, 2, \dots\}$ tvoří geometrický proces, jestliže existuje číslo $a > 0$ takové, že $\{a^{n-1}X_n, n = 1, 2, \dots\}$ tvoří proces obnovy. a se nazývá *kvocient* procesu. Je-li $a \geq 1$, je geometrický proces nerostoucí, pokud je $0 < a \leq 1$, je geometrický proces neklesající. Pro $a = 1$ dostáváme proces obnovy.

Nyní zkonstruujeme markovský proces modelující činnost takovéhoto zařízení, se stavy odpovídajícími jednotlivým fázím buď funkčního zařízení, nebo fázím opravy.

Předpoklad 1. Výskyt náhodných poruch v čase tvoří Poissonův proces s intenzitou λ. Každá z těchto poruch je opravitelná s pravděpodobností p nebo neopravitelná s pravděpodobností $q = 1 - p$. Zařízení se také může porouchat díky opotřebení a tato porucha je vždy neopravitelná. Všechny poruchy nastávají nezávisle na sobě.

Předpoklad 2. Označme X_n dobu života zařízení po jeho $(n-1)$ opravách, $n = 1, 2, \dots$. Předpokládáme, že každé X_n se řídí PH-rozdělením s maticí přechodů mezi provozními stavy $a^{n-1}\mathbf{T}$ a vektorem počátečních pravděpodobností závisejícím na předchozí historii zařízení (bude se lišit v jednotlivých modelech).

Dobu opravy zařízení po n-té poruše označíme Y_n. Předpokládáme, že posloupnost $\{Y_n, n = 1, 2, \dots\}$ tvoří proces obnovy a rozdělení každého Y_n je PH-rozdělení s reprezentací $(\beta, \mathbf{S})$.

Matice $\mathbf{T}, \mathbf{S}$ a vektory α, β mají odpovídající rozměry tak, aby uvažované maticové výrazy měly smysl.

Délky trvání oprav mají všechny stejné PH-rozdělení $G(x)$ s reprezentací $(\beta, \mathbf{S})$. Po dokončení i-té opravy se čas zrychlí a^i-krát, $a \geq 1$. V důsledku toho střední doba mezi poruchami zařízení klesá spolu s rostoucím počtem oprav. Tento pokles, způsobený neúplnými opravami, je reprezentován geometrickým procesem s kvocientem $a \geq 1$.

Předpoklad 3. Posloupnosti $\{X_n, n = 1, 2, \dots\}$ a $\{Y_n, n = 1, 2, \dots\}$ jsou vzájemně stochasticky nezávislé.

Předpoklad 4. Pokud bylo zařízení již opraveno N-krát, po další poruše bude vyměněno za stejné nové, ať je tato porucha jakéhokoli typu.

Stavy zařízení lze rozdělit do dvou skupin: stavy, kdy je zařízení v provozu W, a stavy kdy zařízení nepracuje (opravuje se) F. Stav i bude označovat situaci, kdy zařízení je funkční a bylo na něm provedeno i oprav, $i = 0, 1, 2, \ldots, N$. Stavem i_R budeme označovat situaci, kdy je zařízení v opravě po $i-1$ dokončených opravách, $i_R = 1, 2, \ldots, N_R$. Potom je $W = \{O, 1, 2, \ldots, N\}$ a $F = \{1_R, 2_R, \ldots, N_R\}$. W a F jsou takzvané "makrostavy" a matici generátorů lze konstruovat nad fázemi těchto makrostavů jako blokovou matici. Nechť matice $\mathbf{S}$ a $\mathbf{T}$ mají rozměry $n \times n$ a $m \times m$.

Pokud je zařízení funkční, je v jedné z m funkčních fází. Pokud nepracuje, je opravováno a nachází se v jedné z n fází opravy. Provozní stav i sestává ze všech fází $1, 2, \ldots, m$. Stav opravy i_R sestává ze všech dvojic (i, j), kde $1 \leq i \leq m, 1 \leq j \leq n$. První složka indikuje fázi funkčního stavu, při níž došlo k poruše, druhá složka indikuje fázi opravy.

Je-li zařízení ve funkčním stavu i $(i = 0, 1, 2, \ldots, N-1)$ a nastane neopravitelná porucha, nastává přechod z provozního stavu i do stavu 0. Matice intenzit přechodu je dána vztahem $a^i\mathbf{T}^0\alpha + \lambda q\mathbf{e}\alpha$. Přechod z provozního stavu N do 0 je možný pro jakýkoli druh poruchy a je reprezentován hodnotou $a^N\mathbf{T}^0\alpha + \lambda\mathbf{e}\alpha$.

Je-li zařízení ve funkčním stavu i $(i = 0, 1, 2, \ldots, N)$ a nastane opravitelná porucha, znamená to přechod z provozního stavu i do stavu $(i+1)_R$. Při takovýchto přechodech z libovolné funkční fáze přechází zařízení do fáze opravy podle vektoru počátečního rozložení β. Matice intenzit přechodu je dána vztahem $\lambda p\mathbf{e}\beta$.

Pokud je zařízení ve stavu opravy i_R $(i = 1, 2, \ldots, N)$, dojde po ukončení opravy k přechodu do provozního stavu i podle vektoru počátečního rozložení. Matice intenzit přechodu je dána vztahem $\mathbf{S}^0\alpha$.

Diagonální bloky matice intenzit přechodu, které odpovídají přechodům $i \to i$ $(i = 1, 2, \ldots, N)$ mají tvar $a^i\mathbf{T} - \lambda\mathbf{I}$, pro přechody $i_R \to i_R$ $(i = 1, 2, \ldots, N)$ potom pouze $\mathbf{S}$.

Ostatní přechody nemohou nastat a odpovídající bloky jsou nulové matice. Celá matice intenzit přechodu Q má potom tvar

$$\begin{pmatrix} \mathbf{T} - \lambda\mathbf{I} + \mathbf{T}^0\alpha + \lambda q\mathbf{e}\alpha & \lambda p\mathbf{e}\beta & \cdots & \cdots & \\ \mathbf{S} & \mathbf{S}^0 & \cdots & & \\ a\mathbf{T}^0\alpha + \lambda q\mathbf{e}\alpha & & a\mathbf{T} - \lambda\mathbf{I} & \cdots & \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ a^N\mathbf{T}^0\alpha + \lambda\mathbf{e}\alpha & \cdots & \cdots & \cdots & a^N\mathbf{T} - \lambda\mathbf{I} \end{pmatrix}.$$

3. Stacionární rozdělení.

Představme si chování zařízení v ustáleném stavu, kdy lze definovat různé míry provozu. V tomto stavu jsou pravděpodobnosti toho, že systém bude v jednotlivých stavech, nezávislé na čase. Nejprve najdeme

stacionární rozdělení stavů systému ve tvaru

$$\pi = \{\pi(0), \pi(1_R), \pi(1), \pi(2_R), \ldots, \pi(N)\},$$

vyjádřené pomocí bloků matice Q. V tomto zápise označuje $\pi(i)$ pravděpodobnost, že systém se nachází ve funkčním stavu i, $\pi(i_R)$ je pravděpodobnost stavu opravy i_R. Každý z těchto dílčích vektorů zahrnuje i pravděpodobnosti jednotlivých fází v každém stavu. Vektor π je definován rovnicí $\pi\mathbf{Q} = \mathbf{0}$ a normalizační podmínkou $\pi\mathbf{e} = 1$.

K řešení použijeme Laplace-Stieltjesovu transformaci $\Phi(s)$ distribuční funkce $F(t)$ PH-rozdělení s reprezentací $(\alpha, \mathbf{T})$:

$$\Phi(s) = \int_0^\infty \mathrm{e}^{-sx}\mathrm{d}F(x) = \alpha(s\mathbf{I} - \mathbf{T})^{-1}\mathbf{T}^0$$

Označme

$$\psi_i(\lambda; a) = \prod_{k=0}^{i}(1 - \Phi(a^{-k}\lambda)), \qquad \varepsilon_i = p^i\psi_i(\lambda; a).$$

Stacionární rozdělení lze potom zapsat ve tvaru

$$\begin{aligned}
\pi(0) &= K\lambda^{-1}(1 - \Phi(\lambda)) = K\lambda^{-1}\varepsilon_0 \\
\pi(i)\mathbf{e} &= K\lambda^{-1}p^i\psi_i(\lambda; a) = K\lambda^{-1}\varepsilon_i \qquad i = 1, 2, \ldots, N, \\
\pi(i_R)\mathbf{e} &= K\mu_R p^i\psi_{i-1}(\lambda; a) = K\mu_R p\ \varepsilon_{i-1} \qquad i = 1, 2, \ldots, N.
\end{aligned}$$

kde $\mu_R = -\beta\mathbf{S}^{-1}\mathbf{e}$ je střední doba opravy. Konstantu K spočteme z normalizační podmínky $\pi\mathbf{e} = 1$:

$$K^{-1} = \lambda^{-1}\sum_{i=0}^{N}\varepsilon_i + \lambda p\mu_R\sum_{i=0}^{N-1}\varepsilon_i.$$

Konstantu K lze interpretovat jako střední počet neopravitelných poruch. Podobně i Laplace-Stieltjesova transformaci první doby provozu $\Phi(\lambda)$ lze interpretovat jako pravděpodobnost, že nový system se porouchá díky vnitřním příčinám (opotřebení) dříve, než vlivem externích příčin (náhodné poruchy).

Také další výrazy je třeba interpretovat z praktického hlediska. Tak například, pravděpodobnost, že nový systém se dostane do počátečního stavu $\pi(0) = K\lambda^{-1}(1 - \Phi(\lambda))$ je součinem tří faktorů: středního počtu neopravitelných poruch K, střední doby mezi náhodnými poruchami $1/\lambda$ pokud je zařízení v provozu a pravděpodobnosti, že se náhodná porucha objeví dříve, než porucha z opotřebení, je-li zařízení nové.

Výraz $\psi_i(\lambda; a)$ reprezentuje pravděpodobnost náhodných poruch v posloupnosti stavů systému až do stavu i, před opotřebením. Potom můžeme interpretovat pravděpodobnosti $\pi(i)$ a $\pi(i_R)$ takto: pravděpodobnost $\pi(i)\mathbf{e}$, že systém se nalézá ve stavu i zahrnuje faktor p^i, což jest pravděpodobnost výskytu i opravitelných poruch, $i = 1, 2, \ldots, N$. Pravděpodobnost $\pi(i_R)\mathbf{e}$ že systém se nalézá ve stavu i_R zahrnuje střední doby opravy μ_R. Výrazy ε_i lze interpretovat podobně.

Pohotovost (*availibility*) $A(t)$ je pravděpodobnost, že systém bude funkční v čase t. Uvažujeme-li stacionární stav, potom (stacionární) pohotovost $A = \lim_{t\to\infty} A(t)$. Tato veličina je interpretována jako podíl času, po který je zařízení v provozu. Podíl času, kdy je zařízení nefunkční je potom $\bar{A} = 1 - A$. V našem modelu je to

$$A = \sum_{i=0}^{N} \pi(i)\mathbf{e} = K\lambda^{-1} \sum_{i=0}^{N} \varepsilon_i, \qquad \bar{A} = \sum_{i=1}^{N} \pi(i_R)\mathbf{e} = Kp\mu_R \sum_{i=0}^{N-1} \varepsilon_i$$

Jak lze vidět, stacionární pohotovost závisí na rozdělení doby provozu, počtu oprav a střední době opravy, přičemž rozdělení doby opravy samotné není pro toto vyjádření významné.

Intenzita výskytu poruch (ROCOF) je definována následujícím způsobem: Nechť $N(t)$ je počet poruch do doby t a $M(t) = E(N(t))$ je střední počet poruch za dobu t. *Intenzitou výskytu poruch* potom nazveme funkci $m(t) = (d/dt)M(t)$. Matematické vyjádření ROCOF pro markovské procesy odvodil Lam (1997) v termínech funkcí intenzit přechodu q_{ij}

$$m(t) = \sum_{i\in W, j\in F} p_i(t) q_{ij},$$

přičemž $p_i(t)$ je pravděpodobnost, že systém se nachází ve stavu i v čase t. V ustáleném stavu je $m = \lim_{t\to\infty} m(t)$ a dostáváme stacionární ROCOF. Tato míra je interpretována jako střední počet poruch za jednotku času. Protože uvažujeme tři zásadně odlišné typy poruch, spočteme ROCOF pro každý typ poruch zvlášť.

Opravitelné náhodné poruchy:

$$m_1 = \lambda p \sum_{i=0}^{N-1} \pi(i)\mathbf{e} = Kp \sum_{i=0}^{N-1} \varepsilon_i$$

Neopravitelné náhodné poruchy:

$$m_2 = \lambda q \sum_{i=0}^{N-1} \pi(i)\mathbf{e} + \lambda\pi(N)\mathbf{e} = K\left(q \sum_{i=0}^{N-1} \varepsilon_i + \lambda\varepsilon_N\right)$$

Neopravitelné poruchy z opotřebení:

$$m_3 = \sum_{i=0}^{N} a^i \pi(i)\mathbf{T}^0 = K\left(\Phi(\lambda) + p \sum_{i=1}^{N} \varepsilon_{i-1}\Phi(a^{-i}\lambda)\right).$$

Je zřejmé, že intenzita výskytu náhodných poruch (opravitelných i neopravitelných) je $m_1 + m_2 = \lambda A$. Celková intenzita výskytu libovolného typu poruch v systému je dána vztahem

$$m = m_1 + m_2 + m_3 = \lambda A + \sum_{i=0}^{N} a^i \pi(i)\mathbf{T}^0.$$

Z tohoto vztahu vyplývá, že celková intenzita výskytu poruch je součtem intenzity výskytu poruch, po nichž systém stárne a λ-násobku pohotovosti.

Normalizační konstanta $K = m_2 + m_3$ je součtem intenzit výskytu neopravitelných poruch a proto K^{-1} reprezentuje střední dobu mezi obnovami zařízení.

Porovnáním výrazů pro podíl času, kdy je zařízení nefunkční $\bar{A}$ a ROCOF pro opravitelné náhodné poruchy m_1 dostáváme vztah $\bar{A} = \mu_R m_1$, což znamená, že ve stacionárním stavu je podíl doby opravy roven střední době opravy vynásobené intenzitou výskytu opravitelných poruch (středním počtem opravitelných poruch za jednotku času). Tento vztah lze vyjádřit také takto:

$$\sum_{i=1}^{N} \pi(i_R)\mathbf{e} = \mu_R \lambda p \sum_{i=0}^{N-1} \pi(i)\mathbf{e}.$$

Využitím normalizační podmínky lze získat další výraz pro pohotovost

$$A = \sum_{i=0}^{N} \pi(i)\mathbf{e} = \frac{1 + \lambda p \mu_R \pi(N)\mathbf{e}}{1 + \lambda p \mu_R}.$$

Podobně jako pohotovost ve stacionárním stavu, i ROCOF pro různé typy poruch závisí na rozložení doby provozu, počtu oprav a střední době opravy. Tvar rozdělení doby opravy přitom je nevýznamný.

Tento model lze aplikovat i pro modelování růstu spolehlivosti. Pokud spolehlivost roste po každé opravě, potom koeficient a lze interpretovat jako parametr zlepšování a musí být $a < 1$. Položíme-li $N = \infty$ a $a = 1$, dostáváme proces obnovy s PH-rozdělením doby setrvání v jednotlivých stavech.

Reference

[1] Neuts M. F.: *Matrix geometric solutions in stochastic models.* Baltimore: The Johns Hopkins University Press, 1981.

[2] Neuts M. F., Pérez-Ocón R., Torres-Castro I.: *Repairable models with operating and repair times governed by phase type distributions.* Adv. Appl. Prob. **32**, 468-479, (2000).

[3] Dohnal G., Meca M.: *Aproximace systémů hromadné obsluhy PH-rozdělením.* Sborník konference ROBUST '2002, JČMF Praha, 2002.

Centrum pro jakost a spolehlivost výroby, Fakulta strojní ČVUT v Praze,, Karlovo náměstí 13, 121 35 Praha 2

E-mail address: dohnal@cqr.cz

Conditional states on D-posets

Drobná Eva[1], Chovanec Ferdinand[1],
Kôpka František[2], Nánásiová Oľga[3] *

[1] *Department of Natural Sciences, Academy of Armed Forces*
P. O. Box 45, 031 01 Liptovský Mikuláš, Slovak Republic
e-mail: drobna@aoslm.sk; chovanec@aoslm.sk

[2] *Department of Engineering Fundamentals, Faculty of Electrical Engineering*
University of Žilina, Workplace Liptovský Mikuláš, Slovak Republic
e-mail: kopka@lm.utc.sk

[3] *Department of Mathematics and Descriptive Geometry, Faculty of Civil Engineering*
Slovak Technical University
Radlinského 11, 813 68 Bratislava, Slovak Republic
e-mail: olga@vox.svf.stuba.sk

Abstract

In this paper a new approach to a conditional probability on D-posets is studied.

1 Introduction

Conditional probability plays a basic role in the classical probability theory. Some of the most important areas of the theory such as martingales, stochastic processes rely heavily of this concept. Conditional probabilities on a classical measurable space are studied in several different ways, but result in equivalent theories. The classical probability theory does not decsribe the causality model.

The situation changes when non-standard spaces are considered. For example, it is well known that the set of random events in quantum mechanics experiments is a more general structure than Boolean algebra. In the quantum logic approach the set of random events is assumed to be a quantum logic L. Such model can be found not only in the quantum theory, but also in economics, biology etc.

*This work has been partially supported by the Slovak Academy of Sciences via the project Center of Excellence - Physics of Information, by the grant I/2/2005 and by the Slovak Research and Development Agency under the contract No. APVV-0071-06.

In this paper we will study a conditional state on a D-poset using Renyi's approach (or Bayesian principle). This approach helps us to define such independence of events that admits the situation completely different from that one known in the classical probability theory. Namely, if an event a is independent of an event b, then the event b can be dependent on the event a (problem of causality). We can find some results in [2]-[4] for a quantum logic and an orthomodular lattice. However, a D-poset is more general structure than an orthomodular lattice.

2 Basic notions

To make the going through this paper easier the authors give the short guide to an axiomatic foundation for the D-poset theory in this section.

Let $\mathcal{P}$ be a bounded partially ordered set with the least element $0_{\mathcal{P}}$ and the greatest one $1_{\mathcal{P}}$. Let $\ominus$ be a partial binary difference operation on $\mathcal{P}$ such that there exists $b \ominus a$ in $\mathcal{P}$ if and only if $a \leq b$ and the following axioms hold.

(D1) $a \ominus 0_{\mathcal{P}} = a \quad$ for any $\quad a \in \mathcal{P}$.

(D2) $a \leq b \leq c \quad$ implies $\quad c \ominus b \leq c \ominus a \quad$ and $\quad (c \ominus a) \ominus (c \ominus b) = b \ominus a$.

The structure $(\mathcal{P}, \leq, \ominus, 0_{\mathcal{P}}, 1_{\mathcal{P}})$ is called *a difference poset* (*a D-poset*). For the simplicity of the notation, we write $\mathcal{P}$ instead of $(\mathcal{P}, \leq, \ominus, 0_{\mathcal{P}}, 1_{\mathcal{P}})$.

A lattice ordered D-poset is called a *D-lattice.* In the case of D-lattices it is possible to extend the partial difference operation to the total one as follows

$$b - a := b \ominus (a \wedge b).$$

By *a σ-complete* D-poset we mean a D-poset $\mathcal{P}$ such that for any countable sequence $\{a_n\}_{n=1}^{\infty}$ of elements of $\mathcal{P}$ the least upper bound $\bigvee\limits_{n=1}^{\infty} a_n$ and the greatest lower bound $\bigwedge\limits_{n=1}^{\infty} a_n$ exist in $\mathcal{P}$.

A non-zero element a from a D-poset $\mathcal{P}$ is called *an atom* if the inequality $b \leq a$ entails either $b = 0_{\mathcal{P}}$ or $b = a$. A D-poset $\mathcal{P}$ is said to be *atomic* if for any non-zero element $b \in \mathcal{P}$ there exists an atom $a \in \mathcal{P}$ such that $a \leq b$.

For any element a in a D-poset, the element $1_{\mathcal{P}} \ominus a$ is called the *orthosupplement* of a and is denoted by $a^{\perp}$. The unary operation $\perp\colon a \mapsto a^{\perp}$ is an involution ($(a^{\perp})^{\perp} = a$) and order reversing ($a \leq b$ implies $b^{\perp} \leq a^{\perp}$).

A *sum of orthogonal elements,* denoted by $\oplus$, is a dual partial binary operation to a difference defined by the formula

$$a \oplus b := (a^{\perp} \ominus b)^{\perp} \text{ for } \ a, b \in \mathcal{P}, b \leq a^{\perp}.$$

If $\mathcal{P}$ is a D-lattice, $a, b \in \mathcal{P}$ such that $a \leq b^{\perp}$ and $a \wedge b = 0_{\mathcal{P}}$, then $a \oplus b = a \vee b$.

Let $F = \{a_1, \ldots, a_n\}$ be a finite sequence in a D-poset $\mathcal{P}$. We define

$$a_1 \oplus \ldots \oplus a_n = (a_1 \oplus \ldots \oplus a_{n-1}) \oplus a_n,$$

for any $n \geq 3$, supposing that $a_1 \oplus \ldots \oplus a_{n-1}$ and $(a_1 \oplus \ldots \oplus a_{n-1}) \oplus a_n$ exist in $\mathcal{P}$. We say that a finite system $F = \{a_1, a_2, ..., a_n\}$ of a D-poset $\mathcal{P}$ is $\oplus-$*orthogonal,* if $a_1 \oplus a_2 \oplus ... \oplus a_n$ exists in $\mathcal{P}$ and then we write

$$a_1 \oplus a_2 \oplus ... \oplus a_n = \bigoplus_{i=1}^{n} a_i.$$

An arbitrary system G of $\mathcal{P}$ is $\oplus-$*orthogonal* if every finite subsystem of G is $\oplus-$orthogonal.

For every element $a \in \mathcal{P}$, we define $0a = 0_{\mathcal{P}}$, and $(n+1)a = na \oplus a$ if all involved elements exist. The greatest n such that na exists is called the *isotropic index*, denoted $\iota(a)$, of a. If na exists for every integer n then $\iota(a) = \infty$.

A D-poset $\mathcal{P}$ is *Archimedean* if the statement $na \in \mathcal{P}$ for any $n \geq 1$ implies that $a = 0$.

Elements a and b from a D-poset $\mathcal{P}$ are *compatible* $(a \leftrightarrow b)$, if there exist $c, d \in \mathcal{P}$ such that $c \leq a \leq d$, $c \leq b \leq d$ and $d \ominus a = b \ominus c$.

If $\mathcal{P}$ is a D-lattice, then $a \leftrightarrow b$ if and only if $(a \vee b) \ominus a = b \ominus (a \wedge b)$.

A finite set of atoms $\{a_1, a_2, \ldots, a_n\}$ of a D-poset is called an *atomic partition of the unit*, if

$$\bigoplus_{i=1}^{n} \iota(a_i)a_i = 1_{\mathcal{P}}.$$

If $\mathcal{P}$ is an atomic Archimedean D-poset with a finite set of its atoms, then every element $x \in \mathcal{P}$ is possible to express in the form

$$x = \bigoplus_{i=1}^{n} k_i a_i,$$

where $\{a_1, a_2, \ldots, a_n\}$ is an atomic partition of the unit and $0 \leq k_i \leq \iota(a_i)$.

F. Kôpka in [5] studied the compatibility in D-posets and defined a *Boolean D-poset*.

A poset $\mathcal{P}$ with the least element $0_{\mathcal{P}}$ and the greatest element $1_{\mathcal{P}}$ is said to be a *Boolean D-poset* if there exists a binary operation " $-$ " on $\mathcal{P}$ satisfying the following conditions.

(BD1) $a - 0_{\mathcal{P}} = a$ for any $a \in \mathcal{P}$.

(BD2) $a - (a - b) = b - (b - a)$ for every $a, b \in \mathcal{P}$.

(BD3) $a, b \in \mathcal{P}$, $a \leq b$ implies $c - b \leq c - a$ for any $c \in \mathcal{P}$.

(BD4) $(a - b) - c = (a - c) - b$ for every $a, b, c \in \mathcal{P}$.

From the algebraic point of view, a Boolean D-poset is a D-lattice of pairwise compatible elements and vice versa, therefore Boolean D-posets are algebraically equivalent to MV-algebras.

3 Filters in D-posets

Definition 3.1 *A non-empty subset $\mathcal{F}$ of a D-poset $\mathcal{P}$ is called a filter in $\mathcal{P}$, if the following axioms hold.*

(F1) $a \in \mathcal{F}$, $b \in \mathcal{P}$, $a \leq b$ *implies* $b \in \mathcal{F}$.

(F2) $a \in \mathcal{F}$, $b \in \mathcal{P}$, $b \leq a$ *and* $(a \ominus b)^{\perp} \in \mathcal{F}$ *implies* $b \in \mathcal{F}$.

The axiom (F2) is equivalent to the following one

(F2*) $a \in \mathcal{F}$, $b \in \mathcal{F}$, $b^{\perp} \leq a$ implies $a \ominus b^{\perp} \in \mathcal{F}$.

Let us remark that the element $a \ominus b^{\perp}$ is denoted by $a \odot b$ in the MV-algebras theory.

A filter $\mathcal{F}$ in a D-poset $\mathcal{P}$ is *proper* if $0_{\mathcal{P}} \notin \mathcal{F}$.

Lemma 3.2 *Let $\mathcal{F}$ be a proper filter in a D-poset $\mathcal{P}$. Then $a \in \mathcal{F}$ implies $a^{\perp} \notin \mathcal{F}$.*

PROOF: Suppose that $a \in \mathcal{F}$ and $a^{\perp} \in \mathcal{F}$. From (F2) we have $(a^{\perp} \ominus 0_{\mathcal{P}})^{\perp} = a \in \mathcal{F}$, which gives $0_{\mathcal{P}} \in \mathcal{F}$.

Lemma 3.3 *Let $\mathcal{F}$ be a proper filter in a D-lattice $\mathcal{P}$. If $a, b \in \mathcal{F}$ and $a \leftrightarrow b$ then $a \wedge b \in \mathcal{F}$.*

PROOF: From the assumptions we have $a, b \in \mathcal{F}$, $a \wedge b \leq a$, so

$$(a \ominus (a \wedge b))^{\perp} = ((a \vee b) \ominus b)^{\perp} \geq b,$$

which gives that $(a \ominus (a \wedge b))^{\perp} \in \mathcal{F}$ and pursuant to (F2) of the Definition 3.1 we get that $a \wedge b \in \mathcal{F}$.

The equality $\iota(a) = 1$ is the necessary condition for a nonzero element a to belong into an proper filter. Indeed, if $\iota(a) > 1$, it is equivalent to $a \leq a^{\perp}$, so $a^{\perp} \in \mathcal{F}$, which conflicts with Lemma 3.2.

Lemma 3.4 *Let a be an atom in a D-poset such that $\iota(a) = n, 1 \leq n < \infty$. Then the element ka does not belong to any proper filter for every k, $0 \leq k < n$.*

PROOF: If $n = 1$ then $k = 0$, so $0a = 0_{\mathcal{P}}$ and the element $0_{\mathcal{P}}$ does not belong to any proper filter.

Let $n > 1$. Then $ka \leq (ka)^{\perp}$ for every $k \leq \frac{n}{2}$.

Suppose that $k > \frac{n}{2}$ and $ka \in \mathcal{F}$, where $\mathcal{F}$ is a proper filter. Then $ka \leq a^{\perp}$, which implies that $a^{\perp} \in \mathcal{F}$.

Because $(k-1)a \leq ka$ and $(ka \ominus (k-1)a)^{\perp} = a^{\perp} \in \mathcal{F}$, hence $(k-1)a \in \mathcal{F}$. Similarly $(k-2)a \leq (k-1)a$ and $((k-1)a \ominus (k-2)a)^{\perp} = a^{\perp} \in \mathcal{F}$, so $(k-2)a \in \mathcal{F}$. This procedure we repeat m-times, when $k - m \leq \frac{n}{2}$.

Lemma 3.5 *Let $\mathcal{P}$ be a D-lattice. Let $a \in \mathcal{P}$ be an atom such that $\iota(a) = n$, $1 \leq n < \infty$. Then the interval $[na, 1_{\mathcal{P}}] = \{x \in \mathcal{P} : na \leq x \leq 1_{\mathcal{P}}\}$ is a proper filter in $\mathcal{P}$.*

PROOF: Let $x \in [na, 1_{\mathcal{P}}]$. If $y \in \mathcal{P}$ and $y \geq x$, then evidently $y \in [na, 1_{\mathcal{P}}]$.

Suppose that $x, y \in [na, 1_{\mathcal{P}}], y^{\perp} \leq x$. From the assumption $na \leq x$ it follows that $na \leftrightarrow x$ and $x^{\perp} \wedge na = 0_{\mathcal{P}}$. The inequality $y^{\perp} \leq (na)^{\perp}$ gives that $x - (na)^{\perp} \leq x - y^{\perp}$, therefore,

$$x \ominus y^{\perp} = x - y^{\perp} \geq x - (na)^{\perp} = x \ominus \left(x \wedge (na)^{\perp}\right) = \left((na) \vee x^{\perp}\right) \ominus x^{\perp}$$

$$= (na) \ominus \left((na) \wedge x^{\perp}\right) = na,$$

which gives that $x \ominus y^{\perp} \in [na, 1_{\mathcal{P}}]$.

Let us note that the interval $[na, 1_{\mathcal{P}}]$ is the maximal proper filter in a Boolean D-poset (an MV-algebra), but it is not true in a general D-poset.

4 A conditional state on a D-poset

In this part we introduce the notions of a state and a conditional state and their basic properties.

Definition 4.1 *A map $m : \mathcal{P} \to [0, 1]$ such that*

(i) $m(1_{\mathcal{P}}) = 1$.

(ii) *If $a \leq b$ then $m(b \ominus a) = m(b) - m(a)$*

is called a state on a D-poset $\mathcal{P}$.

Definition 4.2 *Let $\mathcal{P}$ be a D-poset and $\mathcal{P}_0$ is a nonempty subset of $\mathcal{P}$. A mapping $f : \mathcal{P} \times \mathcal{P}_0 \to [0, 1]$ is called a conditional state on $\mathcal{P}$ if the following axioms are fulfilled for every $c \in \mathcal{P}_0$.*

(CS1) $f(c, c) = 1$.

(CS2) $f(1_{\mathcal{P}}, c) = 1$.

(CS3) *If $a, b \in \mathcal{P}$, $a \leq b$ then $f(b \ominus a, c) = f(b, c) - f(a, c)$.*

(CS4) *If $b, c \ominus b \in \mathcal{P}_0$ then $f(a, c) = f(a, b)f(b, c) + f(a, c \ominus b)f(c \ominus b, c)$ for every $a \in \mathcal{P}$.*

(CS5) *If $b \in \mathcal{P}_0$ and $c \ominus b \notin \mathcal{P}_0$ then $f(a, c) = f(a, b)f(b, c)$ for every $a \in \mathcal{P}$.*

The set $\mathcal{P}_0$ is called a conditional system in $\mathcal{P}$.

Remark 4.3 (i) *From* (CS2) *and* (CS3) *it follows that $f(., c)$ is a state on $\mathcal{P}$ for any $c \in \mathcal{P}_0$.*

(ii) *The axiom* (CS3) *is equivalent to the other one*

(CS3*) *If $a, b \in \mathcal{P}$, $a \leq b^{\perp}$ then $f(a \oplus b, c) = f(a, c) + f(b, c)$ for every $c \in \mathcal{P}_0$.*

(iii) *If $\mathcal{P}_0 = \{1_\mathcal{P}\}$, then the set of all conditional states is identical to the set of all states on $\mathcal{P}$. It suffices to put $f(a, 1_\mathcal{P}) = m(a)$ for every $a \in \mathcal{P}$ and a state m.*

Proposition 4.4 *Let $a \in \mathcal{P}$ and $c \in \mathcal{P}_0$ such that $c \leq a$. Then the following assertions are true.*

(i) $f(a, c) = 1$.

(ii) $f(a \ominus c, c) = 0$.

PROOF: (i) From the Definition 4.2 we have $0 \leq f(a, c) \leq 1$. From (CS3) we obtain

$$0 \leq f(a \ominus c, c) = f(a, c) - f(c, c) = f(a, c) - 1,$$

and so

$$1 \leq f(a, c) \leq 1.$$

(ii)

$$f(a \ominus c, c) = f(a, c) - f(c, c) = 1 - 1 = 0.$$

Corollary 4.5 *$f(c^\perp, c) = 0$ for every $c \in \mathcal{P}_0$.*

Proposition 4.6 *Let $\mathcal{P}$ be an Archimedean D-poset, $a \in \mathcal{P}$ and $c \in \mathcal{P}_0$. Then*

$$f(na, c) = nf(a, c)$$

for every $n \in \{0, 1, ..., \iota(a)\}$.

PROOF: It is evident that the assertion is true for $n = 0$.

Suppose that the assertion holds for $0 \leq k < \iota(a)$. Let $n = k + 1$. Then $ka = (k + 1)a \ominus a$ and by (CS3) of Definition 4.2 we have

$$f(ka, c) = f((k + 1)a \ominus a, c) = f((k + 1)a, c) - f(a, c),$$

and hence

$$f((k + 1)a, c) = f(ka, c) + f(a, c) = kf(a, c) + f(a, c) = (k + 1)f(a, c).$$

Proposition 4.7 *Let $b, c \in \mathcal{P}_0$, $b \leq c$. If $c \ominus b \in \mathcal{P}_0$, then the following assertions are equivalent for every $a \in \mathcal{P}$.*

(i) $f(a, b) = f(a, c)$.

(ii) $f(a, b) = f(a, c \ominus b)$.

PROOF: (i) $\Rightarrow$ (ii) From (CS4) of Definition 4.2 we have

$$\begin{aligned} f(a, c\ominus b)f(c\ominus b, c) &= f(a,c) - f(a,b)f(b,c) = f(a,b)(1-f(b,c)) \\ &= f(a,b)(f(c,c) - f(b,c)) = f(a,b)f(c\ominus b, c), \end{aligned}$$

and hence $f(a,b) = f(a, c\ominus b)$.

(ii) $\Rightarrow$ (i) Similarly

$$\begin{aligned} f(a,c) - f(a,b)f(b,c) &= f(a, c\ominus b)f(c\ominus b, c) = f(a,b)(f(c,c) - f(b,c)) \\ &= f(a,b)(1-f(b,c)) = f(a,b) - f(a,b)f(b,c), \end{aligned}$$

therefore $f(a,b) = f(a,c)$.

Proposition 4.8 *If $x \in \mathcal{P}$ and $x \leq x^\perp$ then $x \notin \mathcal{P}_0$.*

PROOF: Let $x \in \mathcal{P}_0$ and $x \leq x^\perp$. Then according to (i) of Proposition 4.4 we have $f(x^\perp, x) = 1$, but in accordance with Corollary 4.5 we get $f(x^\perp, x) = 0$.

Consequently $x \notin \mathcal{P}_0$.

Proposition 4.9 *Let a be an atom in a D-poset $\mathcal{P}$ such that $\iota(a) < \infty$. Then $ka \notin \mathcal{P}_0$ for any $k, 0 \leq k < \iota(a)$.*

PROOF: If $k = 0$ then $0a = 0_\mathcal{P} \notin \mathcal{P}_0$.

Let us assume that $ka \in \mathcal{P}_0$, where $1 \leq k < \iota(a)$. Then $f(ka, ka) = 1$ and also $f(\iota(a)a, ka) = 1$, hence $f(\iota(a)a \ominus ka, ka) = 0$. Since $a \leq \iota(a)a \ominus ka$, we have $f(a, ka) = 0$ and consequently

$$f(ka, ka) = \underbrace{f(a,ka) + f(a,ka) + \ldots + f(a,ka)}_{k-times} = 0,$$

which contradicts our assumption.

Proposition 4.10 *Let $\mathcal{P}$ be an atomic Archimedean D-poset and let $w = \bigoplus_{i=1}^{n} k_i a_i$, where $\{a_1, a_2, \ldots, a_n\}$ is an atomic partition of the unit and $0 \leq k_i < \iota(a_i)$. Then $w \notin \mathcal{P}_0$.*

PROOF: Suppose that $w \in \mathcal{P}_0$. Then $f(w,w) = 1$, $f(1_\mathcal{P}, w) = 1$ and $f(1_\mathcal{P} \ominus w, w) = 0$.

$$1_\mathcal{P} \ominus w = \bigoplus_{i=1}^{n} \iota(a_i)a_i \ominus \bigoplus_{i=1}^{n} k_i a_i =$$

$$= \bigoplus_{i=1}^{n} (\iota(a_i)a_i \ominus k_i a_i) = \bigoplus_{i=1}^{n} (\iota(a_i) - k_i)a_i \geq a_i$$

for every $i = 1, 2, ..., n$. Therefore $f(a_i, w) = 0$ and also $f(k_i a_i, w) = 0$ for every $i = 1, 2, ..., n$ and so $f(w,w) = f(\bigoplus_{i=1}^{n} k_i a_i, w) = \sum_{i=1}^{n} f(k_i a_i, w) = 0$, which contradicts the assumption $f(w,w) = 1$.

Corollary 4.11 *Let $f : \mathcal{P} \times \mathcal{P}_0 \to [0,1]$ be a conditional state on an atomic D-poset $\mathcal{P}$ with a finite set of its atoms. Then the conditional system $\mathcal{P}_0$ is a subset of the union of all maximal proper filters in $\mathcal{P}$.*

References

[1] Chovanec, F., Kôpka, F.: *D-posets*, Appendix A, pp. 278-311, In: Riečan, B., Neubrunn, T.: *"Integral, Measure, and Ordering"*, Kluwer Acad. Publ. Dordrecht, Ister Science, Bratislava, 1997.

[2] Nánásiová O.: *A note on the independent events on a quantum logic*, Busefal **76**, (1998), pp. 53-57.

[3] Nánásiová O.: *Principle conditioning*, Int. Journ. of Theor. Phys., vol. 43 (2004), pp. 1757-1768.

[4] Nánásiová O.: *Map for simultaneous measurements for a quantum logic*, Int. Journ. of Theor. Phys., vol. 42 (2003), pp. 1889-1903.

[5] F. Kôpka, *On compatibility in D-posets*, International Journal of Theoretical Physics, vol. 34 (1995), 1525-1531.

Aproximácia kvantilov necentrálneho t-rozdelenia

Ivan Garaj

FCHPT STU

Ústav informatizácie, automatizácie a matematiky

Radlinského 9, 812 37 Bratislava

ivan.garaj@stuba.sk

Abstract

An approximation formula for a percentage point of the non-central t-distribution is derived in this paper. It was deduced from the Cornish-Fisher expansion for the statistics based on a linear combination of a normal random variable and a chi-random variable. Numerical result is also presented.

***Key words**: Percentage point; Non-central t-distribution; Skewness; Kurtosis; Cornish-Fisher expansion; Gamma function.*

1. Úvod

Nech X je normálne rozdelená náhodná premenná so strednou hodnotou $\mu = \delta$ a smerodajnou odchýlkou $\sigma = 1$ a je nezávislá od náhodnej premennej $f\,S^2$, ktorá má χ^2- rozdelenie s $f = n - 1$ stupňami voľnosti. Náhodná premenná $T_{f,\delta} = X/S$ má [11] necentrálne t-rozdelenie s parametrom necentrality δ. Náhodná premenná S (výberová smerodajná odchýlka) má asymptoticky normálne rozdelenie, pričom pre $\sigma = 1$ platia pre jej strednú hodnotu a disperziu nasledujúce vzťahy [1]:

$$E(S) = b_f = \sqrt{\tfrac{2}{f}}\,\frac{\Gamma\left(\frac{f+1}{2}\right)}{\Gamma\left(\frac{f}{2}\right)},\quad E\left(S^2\right)=1,\quad E\left(S^3\right)=\left(1+\tfrac{1}{f}\right)b_f,\quad E\left(S^4\right)=1+\tfrac{2}{f},\quad D(S)=1-b_f^2 \qquad (1)$$

kde $\Gamma(f) = \int_0^\infty x^{f-1}\,e^{-x}\,dx$ je gama funkcia $(f > 0)$.

2. Aproximácia kvantilov necentrálneho t-rozdelenia využitím Cornishovho-Fisherovho rozvoja

Nech $t_\alpha \equiv t_\alpha(f,\delta)$ je 100α percentný kvantil necentrálneho t-rozdelenia $(0<\alpha<1)$. Zo symetrie [11] vyplýva, že $t_\alpha(f,\delta) = -t_{1-\alpha}(f,-\delta)$. Označme $Z = X - \delta$. Potom platia nasledujúce vzťahy:

$$P\left(T_{f,\delta} \le t_\alpha\right) = \alpha = P\left(\frac{X}{S} \le t_\alpha\right) = P\left(\frac{Z+\delta}{S} \le t_\alpha\right) = P\left(Z+\delta \le t_\alpha\, S\right) = P\left(Z - t_\alpha\, S \le -\delta\right) =$$
$$= P\left(\frac{Z - t_\alpha\,(S - b_f)}{\sqrt{1+t_\alpha^2\,(1-b_f^2)}} \le \frac{t_\alpha\, b_f - \delta}{\sqrt{1+t_\alpha^2\,(1-b_f^2)}}\right) = \alpha \qquad (2)$$

Označme náhodnú premennú

$$W = \frac{Z - t_\alpha\,(S - b_f)}{\sqrt{1+t_\alpha^2\,(1-b_f^2)}} \qquad (3)$$

Pretože $E(W)=0$ a $D(W)=1$ možno na aproximáciu kvantilov necentrálneho t-rozdelenia použiť Cornishov-Fisherov rozvoj [2]

$$\frac{t_\alpha\, b_f - \delta}{\sqrt{1+t_\alpha^2\,(1-b_f^2)}} = u_\alpha + \frac{u_\alpha^2-1}{6}\,k_{3,f}(W) + \frac{u_\alpha^3-3\,u_\alpha}{24}\,k_{4,f}(W) - \frac{2\,u_\alpha^3-5\,u_\alpha}{36}\,k_{3,f}^2(W) -$$
$$- \frac{u_\alpha^4-5\,u_\alpha^2+2}{24}\,k_{3,f}(W)\,k_{4,f}(W) - \frac{3\,u_\alpha^5-24\,u_\alpha^3+29\,u_\alpha}{384}\,k_{4,f}^2(W) + \frac{12\,u_\alpha^4-53\,u_\alpha^2+17}{324}\,k_{3,f}^3(W) +$$
$$+ \frac{14\,u_\alpha^5-103\,u_\alpha^3+107\,u_\alpha}{288}\,k_{3,f}^2(W)\,k_{4,f}(W) - \frac{252\,u_\alpha^5-1\,688\,u_\alpha^3-1\,511\,u_\alpha}{7\,776}\,k_{3,f}^4(W) + \ldots \qquad (4)$$

kde u_α je kvantil normovaného normálneho rozdelenia $N(0,1)$, $k_{3,f}(W) = \mu_{3,f}(W)$ je normovaný koeficient šikmosti, $k_{4,f}(W) = \mu_{4,f}(W) - 3$ je normovaný koeficient špicatosti, $\mu_{3,f}(W)$ je tretí a $\mu_{4,f}(W)$ štvrtý centrálny moment pričom druhý centrálny moment $\mu_{2,f}(W) = D(W) = 1$. Potom platí [1]

$$k_{3,f}(W) = \frac{E\,[\{Z - t_\alpha(S - b_f)\}^3]}{\sqrt{[\,1+t_\alpha^2(1-b_f^2)\,]^3}} \qquad (5)$$

$$k_{4,f}(W) = \frac{E\,[\{Z - t_\alpha(S - b_f)\}^4] - 3\left([\,1+t_\alpha^2(1-b_f^2)\,]^2\right)}{[\,1+t_\alpha^2(1-b_f^2)\,]^2} \qquad (6)$$

Náhodné premenné Z a S sú nezávislé a $E(Z)=0$, $E(Z^2)=1$, $E(Z^3)=0$, $E(Z^4)=3$ a vzhľadom k (1) platia vzťahy:

$$E[(S-b_f)]=0 \tag{7}$$

$$E[(S-b_f)^2]=D(S)=E(S^2)-2b_f\,E(S)+b_f^2=1-b_f^2 \tag{8}$$

$$E[(S-b_f)^3]=E(S^3)-3b_f\,E(S^2)+3b_f^2\,E(S)-b_f^3=b_f\,[2(b_f^2-1)+\frac{1}{f}] \tag{9}$$

$$E[(S-b_f)^4]=E(S^4)-4b_f\,E(S^3)+6b_f^2\,E(S^2)-4b_f^3\,E(S)+b_f^4=$$

$$=\frac{2}{f}(1-2b_f^2)+(1-b_f^2)(1+3b_f^2) \tag{10}$$

$$E[\{Z-t_\alpha(S-b_f)\}^2]=1+t_\alpha^2\,(1-b_f^2) \tag{11}$$

$$E[\{Z-t_\alpha(S-b_f)\}^3]=-t_\alpha^3\,E[(S-b_f)^3]=t_\alpha^3\,b_f\,[2(1-b_f^2)-\frac{1}{f}] \tag{12}$$

$$E[\{Z-t_\alpha(S-b_f)\}^4]=E(Z^4)+6t_\alpha^2\,E[(S-b_f)^2]+t_\alpha^4\,E[(S-b_f)^4]=$$

$$=3+6t_\alpha^2\,(1-b_f^2)+t_\alpha^4\,[\frac{2}{f}\,(1-2b_f^2)+(1-b_f^2)(1+3b_f^2)] \tag{13}$$

Po dosadení (7) až (13) do (5) a (6) dostaneme

$$k_{3,f}(W)=\frac{t_\alpha^3}{[1+t_\alpha^2\,(1-b_f^2)]^{3/2}}\,b_f\,[2(1-b_f^2)-\frac{1}{f}]=\frac{t_\alpha^3}{[1+t_\alpha^2\,(1-b_f^2)]^{3/2}}\,r_{3,f} \tag{14}$$

kde

$$r_{3,f}=b_f\,[2(1-b_f^2)-\frac{1}{f}] \tag{15}$$

$$k_{4,f}(W)=\frac{t_\alpha^4}{[1+t_\alpha^2\,(1-b_f^2)]^2}\,[\frac{2}{f}\,(1-2b_f^2)+2(1-b_f^2)(3b_f^2-1)]=\frac{t_\alpha^4}{[1+t_\alpha^2\,(1-b_f^2)]^2}\,r_{4,f} \tag{16}$$

kde

$$r_{4,f}=\frac{2}{f}(1-2b_f^2)+2(1-b_f^2)(3b_f^2-1) \tag{17}$$

V [3] možno nájsť rozvoj funkcie

$$b_f=1-\frac{1}{4f}+\frac{1}{32f^2}+\frac{5}{128f^3}-\frac{21}{2\,048f^4}-\frac{399}{8\,192f^5}+\frac{869}{65\,536f^6}+\frac{39\,325}{262\,144f^7}-O\left(\frac{1}{f^8}\right) \tag{18}$$

Pomocou rozvoja (18) možno získať nasledujúce tri rýchle rozvoje:

$$1-b_f^2=\frac{1}{2f}-\frac{1}{8f^2}-\frac{1}{16f^3}+\frac{5}{128f^4}+\frac{23}{256f^5}-\frac{53}{1\,024f^6}-\frac{593}{2\,048f^7}+O\left(\frac{1}{f^8}\right) \tag{19}$$

$$r_{3,f}=-\frac{1}{4f^2}-\frac{1}{16f^3}+\frac{13}{128f^4}+\frac{75}{512f^5}-\frac{1\,215}{8\,192f^6}-\frac{17\,403}{32\,768f^7}+O\left(\frac{1}{f^8}\right) \tag{20}$$

$$r_{4,f}=\frac{3}{16f^4}+\frac{3}{16f^5}-\frac{45}{128f^6}-\frac{57}{64f^7}+O\left(\frac{1}{f^8}\right) \tag{21}$$

Rozvoj (4) bude v tvare

$$\frac{t_\alpha\, b_f\, -\delta}{\sqrt{1+t_\alpha^2\,(1-\,b_f^2\,)}} = u_\alpha\, + \frac{(u_\alpha^2-1)\;t_\alpha^3}{6\,[1+t_\alpha^2\,(1-b_f^2)]^{\,3/2}}\; r_{3,f} + \frac{(u_\alpha^3-3\,u_\alpha)\;t_\alpha^4}{24[1+t_\alpha^2\,(1-b_f^2)]^{\,2}}\; r_{4,f}\; - \frac{(2\,u_\alpha^3-5\,u_\alpha)\;t_\alpha^6}{36[1+t_\alpha^2\,(1-b_f^2)]^{\,3}}\; r_{3,f}^2\; -$$

$$- \frac{(u_\alpha^4-5\,u_\alpha^2+2)\;t_\alpha^7}{24[1+t_\alpha^2\,(1-b_f^2)]^{\,7/2}}\; r_{3,f}\; r_{4,f}\; - \frac{(3\,u_\alpha^5-24\,u_\alpha^3+29\;u_\alpha)\;t_\alpha^8}{384\,[1+t_\alpha^2\,(1-b_f^2)]^{\,4}}\, r_{4,f}^2\; + \frac{(12\,u_\alpha^4-53\,u_\alpha^2+17)\;t_\alpha^9}{324[1+t_\alpha^2\,(1-b_f^2)]^{\,9/2}}\; r_{3,f}^3\; +$$

$$+ \frac{(14\,u_\alpha^5-103\,u_\alpha^3+107\,u_\alpha)\;t_\alpha^{10}}{288\;[1+t_\alpha^2\,(1-b_f^2)]^{\,5}}\, r_{3,f}^2\; r_{4,f}\; - \frac{(252\,u_\alpha^5-1\,688\;u_\alpha^3-1\,511\;u_\alpha)\;t_\alpha^{12}}{7\;776[1+t_\alpha^2\,(1-b_f^2)]^{\,6}}\; r_{3,f}^4\; + \ldots \qquad (22)$$

Vynechajme na pravej strane vzťahu (22) okrem prvého člena u_α všetky ostatné členy a vypočítajme t_α. Dostaneme historicky najstaršiu aproximáciu kvantilov necentrálneho t-rozdelenia v literatúre [4] známu pod názvom Jennettova-Welchova

$$t_\alpha \approx \frac{\delta\; b_f\; + u_\alpha\; \sqrt{b_f^2\; + (1-b_f^2)(\delta^2\; - u_\alpha^2)}}{b_f^2\; - u_\alpha^2\,(1-b_f^2)} \qquad (23)$$

Veľmi presnú aproximáciu kvantilov necentrálneho t-rozdelenia získal pomocou rozvoja (22) Akahira [6], keď do výpočtu zahrnul prvé dva členy tohto rozvoja a prvé dva členy rozvoja (20). Kvantily t_α sa potom vypočítajú pre $f = n-1$ numerickým riešením rovnice

$$\frac{\mathrm{t_\alpha\, b_f\, -\delta}}{\sqrt{1+\mathrm{t_\alpha^2\left(1-b_f^2\right)}}} \approx \mathrm{u_\alpha} - \frac{\mathrm{t_\alpha^3\;(u_\alpha^2-1)}}{24\,\sqrt{\left[1+\mathrm{t_\alpha^2\;\left(1-b_f^2\right)}\right]^3}}\left(\frac{1}{\mathrm{f^2}}+\frac{1}{4\,\mathrm{f^3}}\right) \qquad (24)$$

V [7] možno nájsť porovnanie Akahirovej aproximácie s presným riešením, pričom rovnica (24) je v tvare algebraickej rovnice 6. stupňa.

3. Záver

Kvantily necentrálneho t-rozdelenia majú v matematickej štatistike všestranné využitie. Najznámejší je problém výpočtu silofunkcie t-testu. V [11] možno nájsť rozsiahle tabuľky jednostranných tolerančných činiteľov k normálneho rozdelenia, ktorých výpočet úzko súvisí s exaktným výpočtom kvantilov necentrálneho t-rozdelenia vzťahom

$$k = -\frac{t_\alpha\left(n\;-1,\;-u_p\;\sqrt{n}\right)}{\sqrt{n}} = \frac{t_{1-\alpha}\left(n\;-1,\;u_p\;\sqrt{n}\right)}{\sqrt{n}} \qquad (25)$$

kde $\delta = u_p\;\sqrt{n}$ je parameter necentrality a u_p $(0<p<1)$ kvantil normovaného normálneho rozdelenia $N(0,1)$. Problém je však v tom, že programový systém *Mathematica* [5] má síce zabudovanú kvantilovú funkciu $t_\alpha\,(n-1,\delta)$, no pre veľké

$\delta = u_p \sqrt{n} \geq 80$ trvá výpočet aj 30 hodín. Dobrá aproximacia má preto okrem praktického významu aj ekonomický význam. V tabuľke č. 1 je urobené porovnanie presného výpočtu tolerančných činiteľov k podľa vzťahu (25) s presnosťou na 8 desatinných miest, Akahirovej aproximácie (24) a rozvoja (22). Z tabuľky č. 1 vidno, že aproximácia (22) dáva presnejšie výsledky, ako Akahirova aproximácia (24).

Podobnou problematikou sa zaoberajú práce [8], [9] a [10].

Tento článok vznikol s podporou grantového projektu VEGA č. 1/3182/06 "Zlepšovanie kvality produkcie strojárskych výrobkov pomocou moderných štatistických metód".

Literatúra

[1] LIKEŠ, J., MACHEK, J. *Matematická statistika*. Praha: SNTL, 1983. 180 s. (in Czech).

[2] CORNISH, E. A., FISHER, R. A. Moments and cumulants in the specification of distributions. In *Revue de l´Institut International de Statistique, 5, 1937, p. 307-320.*

[3] GARAJ, I.. Číslo π a rozvoje funkcií b_f a $\ln(b_f)$. In *FORUM STATISTICUM SLOVACUM*, 5/2006, p. 60-64.

[4] JENNETT, W. T., WELCH, B. L. The control of proportion defective as judged by a single quality characteristic varying on a continuous scale. In *Journal of the Royal Statistical Society B 6, 1939, p. 80-88*.

[5] WOLFRAM, S. *The Mathematica Book*. 3rd ed. Wolfram Media/Cambridge University Press, 1996. 1403 p. ISBN 0-521-58889-8.

[6] AKAHIRA, M. A higher order approximation to a percentage point of the non-central t-distribution. In *Communications in Statistics, Part B: Simulation and Computation, 1995, 24(3), p. 595-605.*

[7] GARAJ, I. Porovnanie Akahirovej aproximácie s exaktným výpočtom jednostranných tolerančných činiteľov normálneho rozdelenia. In *FORUM STATISTICUM SLOVACUM*, 2/2005, p .19-24.

[8] JANIGA, I., MIKLÓŠ, R. Statistical Tolerance Intervals for a Normal Distribution. In *Measurement Science Review*. ISSN 13, 2001, vol. 1, no. 1, p. 29-32.

[9] GARAJ, I., JANIGA I. *Dvojstranné tolerančné medze pre neznámu strednú hodnotu a rozptyl normálneho rozdelenia*. Bratislava: Vydavateľstvo STU, 2002. 147 s. ISBN 80-227-1779-7.

[10] GARAJ, I., JANIGA I. *Dvojstranné tolerančné medze normálnych rozdelení s neznámou strednou hodnotou a spoločným rozptylom. Two Sided Tolerance Limits of Normal Distributions with Unknown Means and Unknown Common Variability.* Bratislava, Vydavateľstvo STU, 2004, 218 s. ISBN 80-227-2019-4.

[11] GARAJ, I., JANIGA I. Jednostr*anné tolerančné medze normálneho rozdelenia s neznámou strednou hodnotou a rozptylom. One Sided Tolerance Limits of Normal Distributions with Unknown Mean and Variability.* Bratislava, Vydavateľstvo STU, 2005, 214 s. ISBN 80-22-2218-9.

Tabuľka č. 1. Porovnanie tolerančných činiteľov k

1-α = 0,95						
p	*n*	Presne (25)	A=(24)	G=(22)	| P-A |	| P-G |
0,75	100	0,86963145	0,86963196	0,86963172	0,00000051	0,00000027
	150	0,83157811	0,83157831	0,83157821	0,00000020	0,00000010
	200	0,80943700	0,80943710	0,80943704	0,00000010	0,00000004
	250	0,79454469	0,79454476	0,79454472	0,00000007	0,00000003
	300	0,78365997	0,78366002	0,78365999	0,00000005	0,00000002
	400	0,76853069	0,76853072	0,76853070	0,00000003	0,00000001
	500	0,75830279	0,75830281	0,75830280	0,00000002	0,00000001
0,90	100	1,52674875	1,52675274	1,52675072	0,00000399	0,00000197
	150	1,47778886	1,47779062	1,47778962	0,00000176	0,00000076
	200	1,44955118	1,44955219	1,44955158	0,00000101	0,00000040
	250	1,43066286	1,43066352	1,43066310	0,00000066	0,00000024
	300	1,41691112	1,41691159	1,41691128	0,00000047	0,00000016
	400	1,39787260	1,39787288	1,39787269	0,00000028	0,00000009
	500	1,38505219	1,38505238	1,38505224	0,00000019	0,00000005
0,95	100	1,92653886	1,92654634	1,92654256	0,00000748	0,00000370
	150	1,86983886	1,86984229	1,86984033	0,00000343	0,00000147
	200	1,83723565	1,83723765	1,83723642	0,00000200	0,00000077
	250	1,81546848	1,81546981	1,81546895	0,00000133	0,00000047
	300	1,79964194	1,79964290	1,79964225	0,00000096	0,00000031
	400	1,77776079	1,77776137	1,77776096	0,00000058	0,00000017
	500	1,76304592	1,76304631	1,76304602	0,00000039	0,00000010
0,975	100	2,27580587	2,27581682	2,27581136	0,00001095	0,00000549
	150	2,21192703	2,21193214	2,21192922	0,00000511	0,00000219
	200	2,17526187	2,17526489	2,17526302	0,00000302	0,00000115
	250	2,15081037	2,15081240	2,15081107	0,00000203	0,00000070
	300	2,13304620	2,13304767	2,13304668	0,00000147	0,00000048
	400	2,10850605	2,10850694	2,10850630	0,00000089	0,00000025
	500	2,09201614	2,09201675	2,09201630	0,00000061	0,00000016
0,99	100	2,68395786	2,68397302	2,68396558	0,00001516	0,00000772
	150	2,61135090	2,61135809	2,61135400	0,00000719	0,00000310
	200	2,56973712	2,56974142	2,56973876	0,00000430	0,00000164
	250	2,54201098	2,54201388	2,54201199	0,00000290	0,00000101
	300	2,52188081	2,52188292	2,52188148	0,00000211	0,00000067
	400	2,49409048	2,49409177	2,49409084	0,00000129	0,00000036
	500	2,47542870	2,47542958	2,47542891	0,00000088	0,00000021
0,995	100	2,96281363	2,96283170	2,96282292	0,00001807	0,00000929
	150	2,88409016	2,88409879	2,88409389	0,00000863	0,00000373
	200	2,83900406	2,83900925	2,83900605	0,00000519	0,00000199
	250	2,80897830	2,80898182	2,80897952	0,00000352	0,00000122
	300	2,78718560	2,78718817	2,78718642	0,00000257	0,00000082
	400	2,75711005	2,75711163	2,75711049	0,00000158	0,00000044
	500	2,73692025	2,73692134	2,73692052	0,00000109	0,00000027
0,999	100	3,53948435	3,53950836	3,53949691	0,00002401	0,00001256
	150	3,44783327	3,44784488	3,44783835	0,00001161	0,00000508
	200	3,39540040	3,39540743	3,39540310	0,00000703	0,00000270
	250	3,36050567	3,36051047	3,36050734	0,00000480	0,00000167
	300	3,33519119	3,33519472	3,33519232	0,00000353	0,00000113
	400	3,30027230	3,30027448	3,30027291	0,00000218	0,00000061
	500	3,27684234	3,27684385	3,27684272	0,00000151	0,00000038
0,9999	100	4,24651815	4,24654928	4,24653471	0,00003113	0,00001656
	150	4,13866789	4,13868308	4,13867460	0,00001519	0,00000671
	200	4,07701927	4,07702853	4,07702286	0,00000926	0,00000359
	250	4,03601285	4,03601920	4,03601507	0,00000635	0,00000222
	300	4,00627557	4,00628025	4,00627707	0,00000468	0,00000150
	400	3,96527118	3,96527409	3,96527199	0,00000291	0,00000081
	500	3,93776805	3,93777007	3,93776856	0,00000202	0,00000051

Analytic representation of the chord length distribution density of the right circular cone

Wilfried Gille
Martin-Luther-University Halle-Wittenberg, Institute of Physics, Hoher Weg 8,
D-06120 Halle
gille@physik.uni-halle.de

Abstract
Chord length distributions are defined for plane and spatial geometric figures (circle, triangle, cube, ellipsoid ...). The chord length r is a random variable the distribution density of which is denoted by $A(r)$ (isotropic uniform random chords). The function A reflects size and shape of a specific figure. Physical apparatuses allow an automatic recording of A for automatic shape recognition. An estimation of size and shape of an unknown object results by comparing experimental chord length data with theoretical models. However, this *step of comparing theory and experiment* mainly depends on the existence of a large spectrum of model-functions $A(r, a, b, c, \ldots)$ of geometric shapes described by certain parameters a,b,... For the cone an analytic representation of $A(r, R, H)$ via parametric integrals is discussed. The intrinsic details (interval splitting and analytic representation), are explained. *Mathematica* has been applied to handle the result.

1. Introduction

Chord length distributions (CLDs) describe size and shape of geometric figures. Frequently, an object (for example a typical microparticle) can be approximated by a suited, selected geometric figure. In the following, the CLDs of right circular cones of radius R and height H are considered for isotropic uniform random chords (so-called IUR chords) [13].
CLDs are applied in many fields, like materials science [10], astronomy, astrophysics and in archaeology. The distribution law of random chords, belonging to a geometric object, can be obtained by use of scattering methods. The field of data evaluation of small-angle scattering (SAS) experiments involves the theory of CLDs [11,12]. The basic idea is to compare experimental CLDs with theoretical CLD model calculations. For a convex particle (and the cone belongs to this group of models), there exists the basic connection between the SAS correlation function (CF) $\gamma(r)$ (similar to the set covariance) and $A(r)$,

$$A(r) = \bar{l} \cdot \gamma''(r), \ (0 \le r \le L), \tag{1}$$

where L is the largest particle diameter. Eq. (1) has been frequently used to determine the CLD of a certain particle shape, see [3,7,11]. For a single homogeneous particle, the CF can be obtained by determining the volume overlapping between the original particle and its r-translate,

$$\gamma(r) = \frac{\overline{V(r, orientation)}}{V_0}, \ (0 \le r < L), \tag{2}$$

where V_0 is the particle volume, and the numerator-average is the mean overlapping volume for all directions of the r-translate relative to the original particle. The

CF is a strictly monotonously decreasing function, disappearing for larger r-values, $L < r < \infty$. In the following sections, Eq. (2) is operated with in a first step. Then, the function $A(r)$ results by differentiating $\gamma(r)$. The first CLD-moment can be traced back to V_0 and to the whole surface area of the particle S_0 via $\bar{l} = 4V_0/S_0$. Eqs. (1-2) are no approximations and have been applied to reach analytic results for CLDs in many cases; - for ellipsoids, tetrahedrons and hemispheres, see [3,8,9]. The following sections deal with the overlapping volume between the original cone volume and its r-translate, see Fig. 1, in more detail. This problem has been studied several times [5,6], leading to an initial representation of the function $A(r)$. The expenditure for such a calculation is immense. Computer algebra programs are indispensable [7,14].

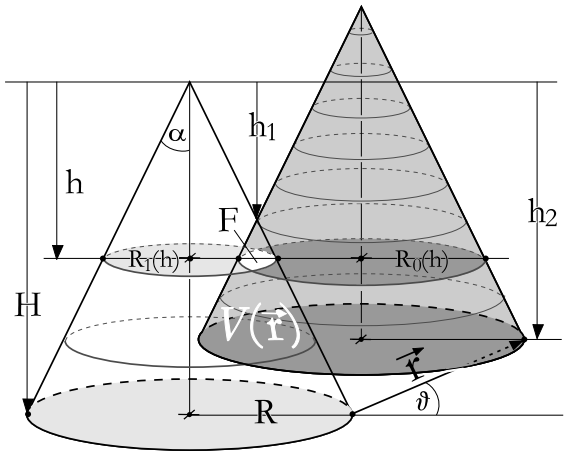


Figure 1: Geometric situation for determining the intersection volume (overlapping volume) $V(r, \vartheta, R, H)$ between the original cone and its r-translate. The symbol F denotes the intersection area of two circles of diameters $R_1 = v \cdot h$ and $R_0 = v[h + r \cdot \sin(\vartheta)]$ as depending on h, v and ϑ. To describe the configuration, $h_{1,2}$ have been introduced.

2. The cone

As a simplifying substitution, $v = R/H$ is introduced. The circular area of diameter $L = 2R$ and the rod of length $L = H$are limiting cases involved in these studies. The cones have the surface area S_c and volume V_c and a half vertex angle α. The direction angle ϑ, varying in the interval $0 \leq \vartheta \leq \pi/2$, defines the direction of the r-translated cone, Fig. 1. Because of the rotational symmetry, no additional polar angle is required. The integration variable for determining the volume $V(r)$ from is h (integration limits from h_2 to h_1). The parameters h, h_1 and h_2 are the distances from the top of the original cone up to planes parallel to the basic plane. The radii R_0 and R_1 of the circles in a plane, parallel to the basic plane, depend on h. The symbol F, $F(r, \vartheta, h)$, denotes the overlapping area of these circles in terms of ϑ, h and r, whereas h is connected with all these variables. The maximum chord length of a cone can either be $2R$ or $\sqrt{H^2 + R^2}$. Thus, there exist two main cases for defining the functions $\gamma(r)$ and $A(r)$, whereas each of these main cases splits into further sub-cases. Four fundamental cone-types must be distinguished. A vivid description of the integration limits, $\vartheta_{1,2,T}$, is explained in Fig. 2. The mean sums of all overlapping volumes for $r = const$ result from the volume cutting into slates

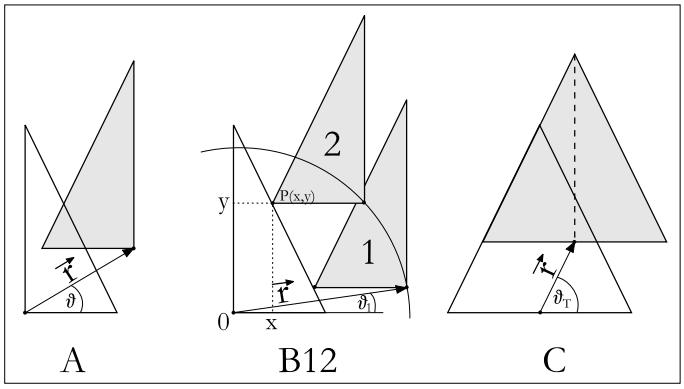


Figure 2: Analysis of overlapping cases between two cones: A: In the basic overlapping case, a configuration as in Fig. 1, there exist limiting angles $\vartheta_{1,2}$. B12: The limiting cases of touching cones define the limiting angles $\vartheta_{1,2}$. There can exist one point $P(x, y)$ or two. C: In the case of total overlapping, $\vartheta_T < \vartheta$, the analysis of $V(r)$ is trivial.

of height dh, arranged parallel to the base of the cone,

$$V_c \cdot \gamma(r) = \int_{\vartheta(r)=...}^{\vartheta(r)=...} \Big(\int_{h=h_1(r,\vartheta)}^{h=h_2(r,\vartheta)} F(r,\vartheta,h)dh \Big) \cos(\vartheta)d\vartheta = \int_{\vartheta(r)=...}^{\vartheta(r)=...} V(r,\vartheta)\cos(\vartheta)d\vartheta, \tag{3}$$

where the factor $\cos(\vartheta)$ controls the averaging process (because all the positions of the cone-ghost in space have the same probability and no direction is privileged, see [13]). Parametric integrals, Eq. (3), involve general integration limits, which are not yet specified to any actual case $[r, R, H]$. The h-integration starts from the top towards the base. The exact specification of all the upper and lower integration limits for ϑ is performed based on the shift of the position of point $P(x, y)$, (see Fig. 2, B12 and the total case C),

$$\vartheta_{1,2} = \arccos\Big(\frac{R}{H^2+R^2} \cdot \Big[\frac{2H^2}{r} \pm \frac{\sqrt{r^2(H^2+R^2) - 4H^2R^2}}{r} \Big] \Big), \tag{4}$$

$$\vartheta_T = \arcsin\Big(\frac{H}{\sqrt{R^2+H^2}} \Big) = \arcsin\Big(\frac{1}{\sqrt{1+v^2}} \Big), \quad \vartheta_E = \arcsin\Big(\frac{H}{r} \Big). \tag{5}$$

The condition of real-valued $\vartheta_{1,2}$, $0 \le r^2(H^2+R^2) - 4H^2R^2$, is reflected as the additionally defined length variable t, $t = 2HR/\sqrt{R^2+H^2}$, see Eqs. (19- 21) below. Furthermore, the specification of the lengths $h_{1,2}$ is

$$h_1 = \frac{r}{2} \cdot \Big[\frac{\cos(\vartheta)}{v} - \sin(\vartheta) \Big], \quad h_2 = H - r \cdot \sin(\vartheta), \tag{6}$$

and

$$h_{max} = h_2 - h_1 = H - \frac{r}{2} \cdot \Big[\frac{\cos(\vartheta)}{v} + \sin(\vartheta) \Big] \tag{7}$$

results (Fig. 1). However, the volume of the intersection region cannot be represented for all the ϑ-values in the trivial form $V(r) = F(r, \vartheta, h_2) \cdot (h_2 - h_1)/3$, as the overlapping volume involves a non-coned surface. The case $h_2 < h_1$ must be excluded. A case $h_2 = h_1$ can exist. The CF results from Eq. (3) from a combination

of five integrals. The notification used, V_{symbol}, is connected with the integration limits for the direction angle ϑ given by Eq. (4-5),

$$V_{01}(r) = \int_0^{\vartheta_1} V(r,\vartheta)\cdot\cos(\vartheta)d\vartheta, \tag{8}$$

$$V_{2T}(r) = \int_{\vartheta_2}^{\vartheta_T} V(r,\vartheta)\cdot\cos(\vartheta)d\vartheta, \tag{9}$$

$$V_{0T}(r) = \int_0^{\vartheta_T} V(r,\vartheta)\cdot\cos(\vartheta)d\vartheta, \tag{10}$$

$$V_{TE}(r) = \int_{\vartheta_T}^{\vartheta_E} V_T(r,\vartheta)\cdot\cos(\vartheta)d\vartheta, \tag{11}$$

$$V_{T\pi}(r) = \int_{\vartheta_T}^{\pi/2} V_T(r,\vartheta)\cdot\cos(\vartheta)d\vartheta. \tag{12}$$

The terms V and V_T, involved in these integrals, can be traced back to a sum $F(r) = F_0(r) + F_1(r)$, see [5]. These terms are

$$\begin{aligned} F_0(r) = &\quad v^2[h + r\cdot\sin(\vartheta)]^2\cdot\arccos\left[\tfrac{a_0+hb}{v[h+r\cdot\sin(\vartheta)]}\right] \\ - &\quad (a_0+hb)\cdot\sqrt{v^2[h+r\cdot\sin(\vartheta)]^2-(a_0+hb)^2}, \end{aligned} \tag{13}$$

$$F_1(r) = v^2\cdot h^2\cdot\arccos\Big[\frac{a_1 - h\cdot b}{vh}\Big] - (a_1 - h\cdot b)\cdot\sqrt{v^2h^2-(a_1-bh)^2}. \tag{14}$$

Eqs. (13, 14) apply substitutions a_0, a_1 and b, defined by

$$a_{0,1} = \frac{r\cdot[\cos^2(\vartheta)\pm v^2\sin^2(\vartheta)]}{2\cos(\vartheta)}, \quad b = \frac{v^2\cdot\sin(\vartheta)}{\cos(\vartheta)}. \tag{15}$$

The basic overlapping volumes in Eqs. (8-12),

$$V(r,\vartheta) = \int_{h_1(r,\vartheta)}^{h_2(r,\vartheta)} [F_0(r,h,\vartheta) + F_1(r,h,\vartheta)]dh, \tag{16}$$

$$V_T(r,\vartheta) = \frac{1}{3}\pi v^2\cdot[H - r\cdot\sin(\vartheta)]^3, \tag{17}$$

result. Eq. (17) represents the case of "total overlapping". Altogether, Eqs. (9,10), together with Eqs. (16,17), fix the cone CF, by summing up over all the ϑ-intervals as depending on $r = const$. Here, an r-interval splitting is indispensable.

2.1. Interval splitting for r

In most of the volume terms Eqs. (8-12), the integration limits depend on r. It is indispensable to introduce a sequence of r-intervals for determining the overlapping integrals Eqs. (8-12). At this, the shape of the cone defined by the parameters R and H influences all calculation steps. Altogether, *four cone-types* must be distinguished: The type 1.1, the *low cone*, is characterised by $0 \le H \le R$, whereas $L = 2R$. The same L exists for the type 1.2, where $R \le H \le \sqrt{3}R$, the *balanced cone*. Furthermore, the types 2.1, *well-balanced cone*, $\sqrt{3}R \le H \le 2R$, and 2.2, *steep cone*, $2R \le H < \infty$, possess the property $L = \sqrt{H^2+R^2}$.
Finally, introducing a special r-value, the length $t = 2R/\sqrt{1+v^2}$, the product

$V_c \cdot \gamma(r)$ results (for the four cases 1.1, 1.2 and 2.1, 2.2) by puzzling the following integral terms together:

$$case\ 1.1:\ V_c \cdot \gamma(r) = \begin{cases} V_{0T} + V_{T\pi}, & 0 \le r \le H \\ V_{0T} + V_{TE}, & H \le r \le \sqrt{H^2 + R^2} \\ V_{01}, & \sqrt{H^2 + R^2} \le r \le 2R \end{cases} \tag{18}$$

$$case\ 1.2:\ V_c \cdot \gamma(r) = \begin{cases} V_{0T} + V_{T\pi}, & 0 \le r \le H \\ V_{0T} + V_{TE}, & H \le r \le t \\ V_{01} + V_{2T} + V_{TE}, & t \le r \le \sqrt{H^2 + R^2} \\ V_0 1, & \sqrt{H^2 + R^2} \le r \le 2R \end{cases} \tag{19}$$

$$case\ 2.1:\ V_c \cdot \gamma(r) = \begin{cases} V_{0T} + V_{T\pi}, & 0 \le r \le t \\ V_{01} + V_{2T} + V_{T\pi}, & t \le r \le H \\ V_{01} + V_{2T} + V_{TE}, & H \le r \le 2R \\ V_{2T} + V_{TE}, & 2R \le r \le \sqrt{H^2 + R^2} \end{cases} \tag{20}$$

$$case\ 2.2:\ V_c \cdot \gamma(r) = \begin{cases} V_{0T} + V_{T\pi}, & 0 \le r \le t \\ V_{01} + V_{2T} + V_{T\pi}, & t \le r \le 2R \\ V_{2T} + V_{T\pi}, & 2R \le r \le H \\ V_{2T} + V_{TE}, & H \le r \le \sqrt{H^2 + R^2} \end{cases} \tag{21}$$

Outside these r- intervals, the function $\gamma(r)$ disappears. Numerical checks with a random number generator for the variables R and H show that the equations given fulfil

$$V_c = \int_0^L 4\pi r^2 \cdot \gamma(r) dr, \tag{22}$$

see [11,12]. The final step of the project is to calculate $4V_c/S_c \cdot \gamma''(r)$, Eq. (1).

2.2. From the CF to the CLD

Clearly, Eq. (1) could be handled numerically. On the other hand, a twice differentiation of the sum of the parametric integrals can be performed. The terms simplify, as expected from experiences from other geometric figures, see [3,7]. The differentiation of the terms $V''_{T\pi}(r)$ and $V''_{TE}(r)$, simple to handle, yields

$$V''_{T\pi}(r) = \frac{2\pi H v^2}{3} \cdot \Big[1 - \frac{1}{(1+v^2)^{3/2}}\Big] - \frac{\pi r v^2}{2} \cdot \Big[1 - \frac{1}{(1+v^2)^2}\Big] \tag{23}$$

and

$$V''_{TE}(r) = \frac{\pi H^4 v^2}{6r^3} + \frac{\pi r v^2}{2(1+v^2)^2} - \frac{2\pi H v^2}{3(1+v^2)^{3/2}}. \tag{24}$$

An elaborate calculation leads to parametric integrals for the three terms $V''_1(r)$, $V''_{2T}(r)$ and $V''_{0T}(r)$. Thus, the whole CLD is fixed by a sum of maximally five terms,

$$A(r) = \frac{4}{S_c}[c_{01}V''_{01}(r) + c_{2T}V''_{2T}(r) + c_{0T}V''_{0T}(r) + c_{T\pi}V''_{T\pi}(r) + c_{TE}V''_{TE}(r)]. \tag{25}$$

The values of the five coefficients are 0 or 1, as depending on r, see Eqs. (18-21). From a long calculation, the parametric integrals V''_{01}, V''_{2T} and V''_{0T},

$$V_{01}''(r) = \int_0^{x_1(r)} I_1''(r,x)dx,\ V_{2T}''(r) = \int_{x_2(r)}^{x_T} I_1''(r,x)dx,\ V_{0T}''(r) = \int_0^{x_T} I_1''(r,x)dx, \tag{26}$$

result. Eqs. (26) apply several substitutions. The integration limits are defined by

$$x_T = \frac{1}{\sqrt{1+v^2}}, \quad x_{1,2} = \frac{2Hv^2 \mp \sqrt{r^2(1+v^2) - 4H^2v^2}}{r(1+v^2)}. \tag{27}$$

For I_1'' a representation

$$\begin{aligned} I_1''(r,x) = & \ \frac{r}{2v}\left(1 - \frac{x^2}{{x_T}^2}\right)^{3/2} \cdot \ln\left[\frac{v[2H/r - x] + (1/x_T)\cdot\sqrt{(x_1-x)(x_2-x)}}{\sqrt{1-x^2}}\right] \\ + & \ 2v^2x^2 \cdot (H - rx) \cdot \arccos\left[\frac{vx(rx-2H) + (r/v)\cdot(1-x^2)}{2(H-rx)\cdot\sqrt{1-x^2}}\right] \\ + & \ \frac{r\cdot x}{x_T} \cdot \sqrt{1 - \frac{x^2}{{x_T}^2}} \cdot \sqrt{(x_1 - x)(x_2 - x)} \end{aligned} \tag{28}$$

follows. Several synonymous versions of Eq. (28) are possible. However, a rigorous simplification of all these terms and expressions has not been reached yet.

3. The CLD of flat, balanced, well-balanced and steep cones

The CLDs are represented by Plot3D, see [14]. As the dynamics of the CLD

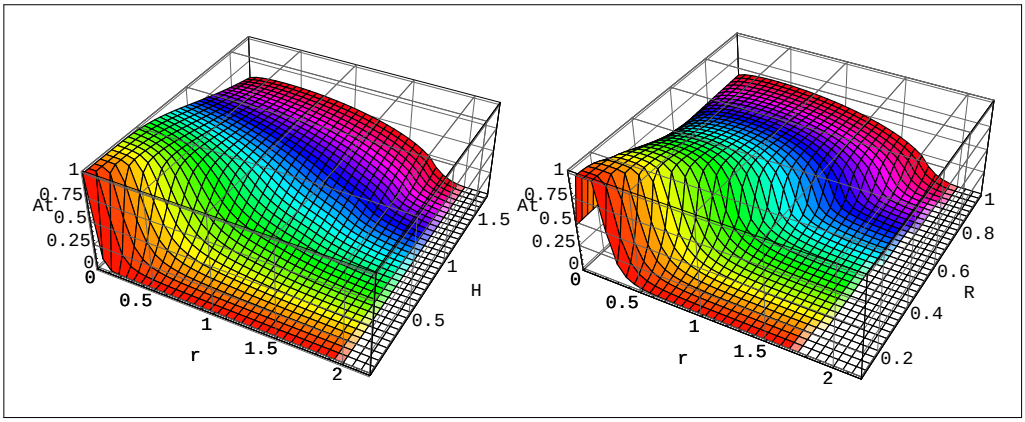


Figure 3: The chord length distribution density of cones. A transformation $At(r,R,H) = \tanh[A(r,R,H)]$ has been used. Left-hand side: *F*lat and *b*alanced cones, $R = 1$, $L = 2$, $0 < H \leq \sqrt{3}$. Right-hand side: *W*ell- balanced and *s*teep cones, $H = const = \sqrt{3}$, $H < L \leq 2$, $0 < R \leq 1$.

is relatively high, any complex CLD representation requires a transformation of $A(r,R,H)$, see Fig. 3. The transforming function $f = \tanh(x)$, possessing the property $0 \leq f < 1$ for non-negative x. For small arguments x, $f \approx x$ holds. For larger arguments the CLD is somewhat distorted. By use of the transformation

$$I(h) = \frac{4\pi}{V \cdot \bar{l}} \int_0^L A(r,R,H) \cdot \frac{2 - 2\cos(hr) - hr\sin(hr)}{h^4} dr, \tag{29}$$

see [3], the SAS intensity $I(h) = I(h,R,H)$ (scattering vector amount h) results. For example, the function $I(h)$ can be recorded by neutron scattering experiments.

5. Summary and conclusions

The cone is an elementary geometric figure. However, besides [5,6] the author does not know any paper involving a more compact analytic formula for this CLD. The current paper involves a transformed 3D plot of the CLD, see Fig. 3. The

expenditure for tracing back the parametric integrals to standard functions of mathematical physics is greater than expected. From a more general point of view, this fact has been explained by Ciccariello [1,2]. Mainly, there exist three reasons for these restrictions:
1. There seems to be no way for avoiding r-interval splitting and considering different basic types of cones [5,6].
2. Furthermore, the author does not know of any general algorithm for receiving CLDs by use of computer algebra programs (at least these programs are too weak for automatically detecting an interval splitting).
3. Even the restriction to a numerical analysis of the parametric integrals is not simple.
The results obtained are reproducible and can be obtained operating with all *Mathematica* versions. It is a worthwhile attempt to summarise the three parametric integrals in question, Eqs. (26), by a more suited way, see [4], followed by simplification, in order to reach a better representation.

References

1. Ciccariello, S.: Acta Cryst. **A49**(1993), pp. 750-758.
2. Ciccariello, S.: J. Math. Phys. **36**(1995), pp. 219-246.
3. Feigin L. A. & Svergun, D. I.: *Structure Analysis by Small-Angle X-ray and Neutron Scattering*. New York: Plenum Press, 1987.
4. Filipescu, D., Trandafir, R. & Zorilescu D. : *Probabilitati Geometrice Si Applicatii*, (in Roumanian language), Cluj-Napoca: Editura Dacia, 1981.
5. Gille, W.: Mathematical & Computer Modelling, **30**(1999) pp. 107-130.
6. Gille, W.: Computers & Mathematics with Applications, **40**(2000) pp. 1027-1035.
7. Gille, W.: *Mathematica Programs for Chord Length Distributions of Selected Geometric Figures,Zbornik radova*, PrimMath[2001], *Mathematica u znanosti, tehnologiji i obrazovanju*, Zagreb, 27-28.9.2001.
8. Gille, W.: *Das Konzept der Sehnenlaengenverteilung*, Habilitationsarbeit, Martin-Luther-University Halle-Wittenberg, Fachbereich Physik, Halle, 1995.
9. Gille, W.: Adv. Appl. Prob. (SGSA) **34**(2002) pp. 11-20.
10. Hermann, H.: *Stochastic Models of Heterogeneous Materials*, Materials Science Forum, Volume 78. Zuerich: Trans Tech Publications, 1991.
11. Porod, G.: Kolloid-Z. **124**(1951) pp. 83-111; & **125**(1952) pp. 51-57 & **125**(1952) pp. 108-122.
12. Porod G.: In: O. Glatter & O. Kratky (Eds.); Small-Angle X-ray scattering, Section I. *The principles of diffraction*, General Theory, page 34. Academic press: London, 1982.
13. Stoyan D., W.- S. Kendall & J. Mecke: *Stochastic Geometry and Its Applications*. New York: John Wiley, 1987.
14. Wolfram Research (2005). Inc. *Mathematica* vers. 5.2, Champaign, Illinois.

Štatistická optimalizácia drsnosti povrchu pri sústružení ocele ISO 683/1-87 pomocou DoE

Miroslav Gombár, Alena Vagaská, Sergej Hloch

Fakulta výrobných technológií

Technická univerzita v Košiciach

Bayerova 1

080 01 Prešov

vagaska.alena@fvt.sk, mirek@unipo.sk, hloch.sergej@fvt.sk

Abstrakt

Príspevok sa zaoberá štatistickou optimalizáciou drsnosti povrchu pri pozdĺžnom sústružení ocele ISO 683/1-87 gradientnou metódou s využitím výsledkov úplného faktorového experimentu 2^3. Ako nezávislé premenné boli brané do úvahy rezné podmienky: rezná rýchlosť v_c, posuv f a hĺbka rezu a_p.

Kľúčové slová: štatistická optimalizácia, plánovaný experiment, regresná analýza

Úvod

Častým základným cieľom pri hodnotení sprievodných javov rezného procesu je hľadanie optimálnych podmienok pri obrábaní [4], [7], [8]. V najširšom slova zmysle optimalizácia predstavuje získanie najlepších výsledkov v určitých podmienkach. Matematická formulácia optimalizácie sa vo všeobecnosti zakladá na hľadaní maxima (a pri zmene znamienka tiež minima) funkcie:

$$q = f(x_0, x_1, x_2, \dots, x_N, y_0, y_1, y_2, \dots, y_N) \tag{1}$$

kde $x_0, x_1, \dots, x_N, y_1, y_2, \dots, y_N$ je konečný počet reálnych premenných. Funkcia q sa nazýva kritériom optimalizácie (kriteriálna funkcia, účelová funkcia, cieľová funkcia, optimalizačný funkcionál) a charakterizuje kvantitatívne získané výsledky [4]. Riadené, resp. kontrolované premenné $x_0, x_1, \dots, x_N$ charakterizujú podmienky, na ktoré je možné vplývať ľubovoľným spôsobom s cieľom získania najlepších podmienok podľa vybraného kritéria. Neriadené premenné $y_1, y_2, \dots, y_N$ charakterizujú podmienky, na ktoré nie je možné vplývať a sú určované nezávislými činiteľmi. Väzbu oboch druhov premenných s kritériom optimalizácie q určuje predpis funkcie (1). Teória regresnej analýzy je založená na výpočte extrémov funkcií a v základnej úlohe predpokladá, že veličina y (meraná veličina) stochastickým spôsobom závisí na premenných x_j, $j = 1, 2, 3, \dots, k$ ($k -$ je počet nezávisle premenných). Ak sú známe hodnoty y_j,

$j = 1, 2, 3, \ldots, N$ (N – je počet nezávisle meraní) spolu s odpovedajúcimi x_{ij}, potom úlohou regresie je odhad parametrov $b_0, b_1, \ldots, b_k$ vo funkcii (2):

$$\hat{y} = f(x_0, x_1, x_2, ..., x_N, b_0, b_1, b_2, ..., b_k) \quad (2)$$

Parametre $b_0, b_1, \ldots, b_k$ sa určujú tak, aby minimalizovali funkciu odchýlky (3):

$$u_i = \hat{y}_j - f(x_0, x_1, x_2, ..., x_N, b_0, b_1, b_2, ..., b_k) \quad (3)$$

kde $j = 1, 2, 3, \ldots, N$ a $i = 1, 2, 3, \ldots, k$. Ak funkcia odchýlok je vyjadrená v tvare (4):

$$f_0(u_1, u_2, ..., u_N) = u_1^2 + u_2^2 + \ldots + u_N^2 \quad (4)$$

vtedy ide o metódu najmenších štvorcov. Najjednoduchším prípadom je, ak rovnica (2) je určená lineárnym, resp. linearizovaným tvarom podľa vzťahu (5):

$$y = b_0 . x_0 + \sum_{j=1}^{k} b_j . x_j + \varepsilon \quad (5)$$

kde ε je náhodná chyba s nulovou strednou hodnotou , resp. pre N meraných výsledkov pri k premenných v maticovom tvare (6):

$$\boldsymbol{y} = \boldsymbol{X}.\boldsymbol{b} + \boldsymbol{\varepsilon} \quad (6)$$

kde $\boldsymbol{y} = [y_1, y_2, \ldots, y_N]^T$ je vektor meraných výsledkov, $\boldsymbol{X}$ – matica podmienok merania, $\boldsymbol{\varepsilon}$ – náhodný vektor s nulovou strednou hodnotou. Vektor koeficientov viacnásobnej lineárnej resp. linearizovanej regresie $\boldsymbol{b} = [b_0, b_1, \ldots, b_k]^T$ sa určuje maticovou analýzou regresných závislostí. Lineárna, resp. linearizovaná závislosť (5) nemá lokálne extrémy pre svoje maximum, resp. minimum a teda nie je vhodná pre popis oblastí ležiacich v blízkosti optima. Pomocou (5) resp. (6) je možné určiť smer rastu, alebo poklesu regresnej funkcie. Pohyb po viacrozmernom priestore $E_k[x_0, x_1, \ldots, x_k]$ je diskrétny a vyjadruje sa pomocou tzv. optimalizačných krokov. Týmto postupom pri použitej stratégii experimentovania možno získať v každom štádiu hodnovernejšie informácie o optimálnom priebehu vyšetrovaného procesu či javu [4].

Podmienky experimentu

Úlohou plánovania experimentov je naplniť model reálnych situácií vplyvu rezných podmienok na morfológiu obrobeného povrchu pri pozdĺžnom sústružení ocele ISO 683/1-87 konkrétnymi číslami. V praxi nepôsobia tieto procesné parametre vždy len aditívnym spôsobom, ale spoločne vo vzájomnej interakcii. Analyzovať tento model umožňujú faktorové experimenty, pri ktorých sa zrealizujú pokusy pre všetky kombinácie úrovní uvažovaných faktorov. Z hľadiska optimalizácie rezných podmienok vzhľadom na hodnotu najväčšej výšky nerovnosti profilu drsnosti Rz a strednej aritmetickej odchýlky profilu drsnosti Ra sú v tab.1. uvedené Kódované podmienky pokusov a v tab.2 podmienky pri ktorých sa experiment realizoval.

Tab.1 Kódovanie podmienok pokusov.

Poradie pokusov	Kódované podmienky pokusov			Aktuálne podmienky pokusov		
	x_1	x_2	x_3	v_c m.min^{-1}	f mm	a_p mm
1.	-1	-1	-1	8,792	0,1	0,1
2.	+1	-1	-1	351,68	0,1	0,1
3.	-1	+1	-1	8,792	0,5	0,1
4.	+1	+1	-1	351,68	0,5	0,1
5.	-1	-1	+1	8,792	0,1	3,0

6.	+1	-1	+1	351,68	0,1	3,0
7.	-1	+1	+1	8,792	0,5	3,0
8.	+1	+1	+1	351,68	0,5	3,0

Tab.2 Podmienky vykonania experimentu.

Kód experimentu		**Rz.v_c,f,a_p – 12 050.1**			
Použitý obrábací stroj :		SU 40			
Použitý rezný nástroj		**Držiak**	**Rezná platnička**	**Rezný materiál**	**r_ε [mm]**
		MWLNR	KNUX 190 405 EL	P20 podľa ISO	0.50
Rezné podmienky		**v_c [m.min^{-1}]**	**a_p [mm]**	**f [mm]**	
		7.892 – 351.680	0.10 – 3.00	0.100 – 0.500	
Nastavenie nástroja voči osi obrobku		h = 0 [mm]	**Chladenie**	nie	
Obrábaný materiál		ISO 683/1-87			
Použité meracie prístroje		Mitutoyo Surftest SJ – 301 na meranie parametrov drsnosti povrchu			
Presnosť výpočtu	E = 1/1000	**Zvolená hladina významnosti**		α = 0.05	
Počet prípadov	N= 8	**Počet opakovaných meraní pre každý prípad**		m = 5	

Aplikácia štatistickej optimalizácie pre podmienky realizovaného výskumu

Štatistická optimalizácia rezných podmienok vo vzťahu k najväčšej výške nerovnosti profilu drsnosti Rz a strednej aritmetickej odchýlky profilu drsnosti Ra, bola realizovaná na základe vykonaného plánovaného experimentu 2^3, pri použití reznej rýchlosti v_c, posuvu f a hĺbky rezu a_p ako nezávisle premenných – faktorov a hodnôt Rz a Ra ako parametrov optimalizácie, za účelom stanovenia optimálnych hodnôt rezných podmienok vo vzťahu k dosiahnutiu optimálnej hodnoty skúmaných parametrov drsnosti povrchu. Pre štatistickú optimalizáciu boli využité výsledky meraní plánovaného experimentu a mocninových regresných modelov v logaritmických súradniciach a regresný model je v normovanom tvare. Teda základný tvar regresnej funkcie v logaritmickej súradnicovej sústave pre štatistickú optimalizáciu má tvar (7):

$$Rz, Ra = b_0 + b_1 x_1(v_c) + b_2 x_2(f) + b_3 x_3(a_p) \qquad (7)$$

Výsledky regresnej analýzy pre výpočet regresných koeficientov rovnice (7) sú v zjednodušenej podobe uvedené v tab.3. Vzhľadom na to, že vzťah (7) je uvažovaný pre logaritmickú transformáciu (kódovaná funkcia), pre veľkosť vstupných premenných pri optimalizácii výstupných parametrov sú uvažované logaritmické hodnoty vstupných veličín, je potrebné počítať logaritmický stred definovaný vzťahom (8) a logaritmický variačný interval pôsobiacich faktorov podľa vzťahu (9):

$$\log z_{jo} = \frac{\log z_{j\max} + \log z_{j\min}}{2} \;,\; \log \Delta z_j = \frac{\log z_{j\max} - \log z_{j\min}}{2} \qquad (8,9)$$

Tab.3 Podmienky vykonania experimentu.

Veličina	$\hat{R}z = b_0 + b_1 x_1(v_c) + b_2 x_2(f) + b_3 x_3(a_p)$	$\hat{R}a = b_0 + b_1 x_1(v_c) + b_2 x_2(f) + b_3 x_3(a_p)$
Odhad regresných koeficientov		
b_0	1.242139 ± 0.325081	0.322289 ± 0.263193
b_1	-0.000929 ± 0.000809	-0.000758 ± 0.000655
b_2	1.375359 ± 0.693209	1.881198 ± 0.561239
b_3	0.008627 ± 0.095615	0.020402 ± 0.077412

Štatistická významnosť regresných koeficientov		
$k_{kritické}$	2.571000	2.571000
t_0	9.823831	3.148269
b_0	signifikantný	signifikantný
t_1	2.956325	2.975048
b_1	signifikantný	signifikantný
t_2	5.100985	8.617652
b_2	signifikantný	signifikantný
t_3	0.229289	0.677585
b_3	štatisticky nevýznamný	štatisticky nevýznamný

Tab.4 Určenie bázových koeficientov a optimalizačných krokov.

Vstupná veličina x_j		x_1 (v_c)	x_2 (f)	Parameter optimalizácie
Koeficient regresie b_j		-0.000929	1.375359	Rz
		-0.000758	1.881198	Ra
Logaritmický stred log z_{jo}		1.745118	-0.650515	Rz, Ra
Logaritmický interval $log\ \Delta z_j$		0.801029	0.349485	Rz, Ra
$b_j . log\ \Delta z_j$		-0.000744	0.480667	Rz
		-0.000607	0.657450	Ra
Bázový koeficient b_b		-	1.375359	Rz
		-	1.881198	Ra
$\lambda o = \frac{\mu}{\lvert b_b \rvert}$	μ = 0.1 → λo = 0.072708 μ = 1 → λo = 0.727082			Rz
	μ = 0.1 → λo = 0.053158 μ = 1 → λo = 0.531576			Ra

Pri minimalizácii výstupných parametrov Ra a Rz sa uvažoval teda vplyv rezných podmienok, a túto optimalizáciu je teda nutné chápať iba vo vzťahu k rezným podmienkam, bez uvažovania vplyvu ostatných podmienok a procesov ktoré pôsobia pri vytváraní obrobeného povrchu obrábaním. Na základe uvedených vzťahov maximálna hodnota súčinu $b_j\, log\, \Delta z_j$ určuje bázový koeficient b_b, s ktorým sa uvažovalo pri výpočte optimalizačného kroku. Optimalizačný krok bol zvolený s ohľadom na interval $(0 < \mu \leq 1)$ tak, aby výpočtové operácie boli čo najjednoduchšie. Preto sa uvážili nasledovné prípady: $\mu = b_b$, $\lambda_o = 1$, $\mu = 1$, $\lambda_o = 1/b_b$

Výsledky spracovania údajov potrebných pre optimalizáciu sú uvedené v tab.4. Hodnoty skúmaných parametrov Rz a Ra získané výpočtovým modelom s krokom $\lambda_0 = 0{,}072708$ pre Rz a $\lambda_0 = 0{,}053158$ pre Ra; $\lambda_0 = 0{,}072708$ resp. $\lambda_0 = 0{,}053158$ sú uvedené v tab.5 pre parameter Rz a v tab.6 pre parameter Ra. Hodnoty skúmaných parametrov sa v tomto kroku optimalizácie sledujú až po dosiahnutie tzv. kvázistacionárnej oblasti, pričom z tab.5 a tab.6 je zrejmé, že minimalizácia výstupného parametra nastane len pri zmenšení posuvu na otáčku.

Tab.5 Výpočtové hodnoty parametrov optimalizácie pre parameter Rz

Krok			λo = 0.072708				λo = 0.727082
Číslo kroku			l = 0	l=1	l=2	l=3	l = 1
	f [mm]	x_2	-0.650515	-0.685463	-0.720412	-0.755360	-0.999999
		z_2	0.223607	0.206318	0.190366	0.175647	0.100000
Parameter optimalizácie	Vypočítaná hodnota	Rz	2.987257	2.849722	2.712186	2.574651	1.611899

Tab.6 Výpočtové hodnoty parametrov optimalizácie pre parameter *Ra*

Krok			λo = 0.053158				λo = 0.531576
Číslo kroku			l = 0	l=1	l=2	l=3	l = 1
	f [mm]	x_2	-0.650515	-0.685464	-0.720413	-0.755361	-1.000004
		z_2	0.223607	0.206318	0.190365	0.175646	0.099999
Parameter optimalizácie	Vypočítaná hodnota	*Ra*	2.067407	1.879286	1.691165	1,503044	0.186188

Na základe uvedených výsledkov získaných pri optimalizácii linearizovaného modelu sa preukázal najväčší vplyv na parameter optimalizácie *Rz* a *Ra* z rezných podmienok posuv. Dosiahnutá kvázistacionárna oblasť sa aproximovala polynómom druhého stupňa, kde vo všeobecnosti ako nezávislé premenné vystupujú ti faktory, ktoré majú podstatný vplyv na zmenu parametra optimalizácie žiadaným smerom. Pre konkrétny prípad je to posuv a aproximačný polynóm má tvar (10):

$$\hat{R}a, \hat{R}z = b_0 + b_1(f) + b_{11}x_2(f)^2 \tag{10}$$

Výsledky štatistického spracovania pre dané podmienky merania vzťahom (10) sú v tab.7. Po vylúčení štatisticky nevýznamného koeficientu b_{11} má rovnica pre kvázistacionárnu oblasť strednej aritmetickej odchýlky profilu drsnosti tvar (11):

$$\log \hat{R}a = 0.781722 + 0.309764x_1(f) \tag{11}$$

Tento vzťah, ktorý je v kódovanom tvare prevedieme do prirodzenej mierky pomocou prevodového vzťahu (12):

$$x_j = \frac{2(\log z_j - \log z_{j\max})}{\log z_{j\max} - \log z_{j\min}} + 1 \tag{12}$$

Tab.7 Regresná analýza optimalizačného polynómu.

Veličina	$\hat{R}z$	$\hat{R}a$
Odhad regresných koeficientov		
b_0	1.500369	0.781722
b_1	0.505125	0.309764
b_{11}	0.228057	0.146772
Štatistická významnosť regresných koeficientov		
$\mathbf{k_{kritické}}$	2.571000	2.571000
t_0	16.499109	10.067062
b_0	signifikantný	signifikantný
t_1	5.554709	3.989166
b_1	signifikantný	signifikantný
t_{11}	2.50487	1.890146
b_{11}	štatisticky nevýznamný	štatisticky nevýznamný

V prirodzenej mierke prostredníctvom prevodového vzťahu (12) pre minimálne a maximálne hodnoty posuvov (13)

$$\log \hat{R}a = b_0 + b_1\left(\frac{2(\log f - \log f_{\max})}{\log f_{\max} - \log f_{\min}} + 1\right) = b_0 + b_1\left(\frac{2 * \log f - \log f_{\max} - \log f_{\min}}{\log f_{\max} - \log f_{\min}}\right) \tag{13}$$

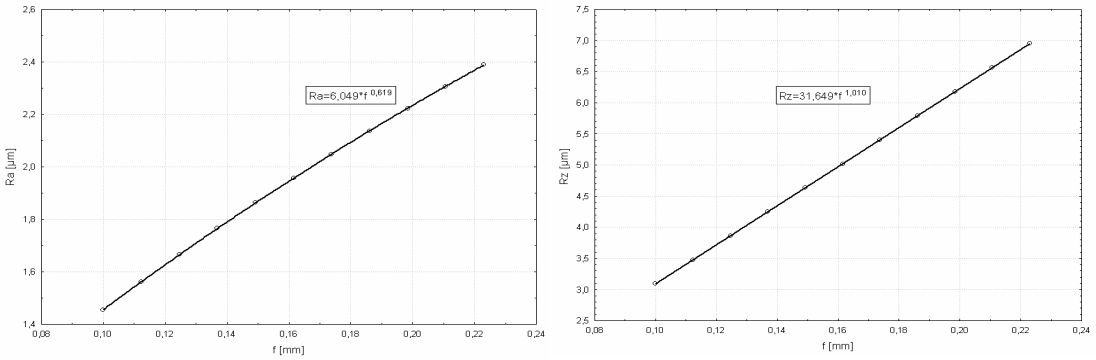
dostávame rovnicu pre kvázistacionárnu oblasť strednej aritmetickej odchýlky profilu drsnosti (14):

$$\hat{R}a = 6{,}049 f^{0{,}619} \tag{14}$$

A rovnakým spôsobom dostávame rovnicu pre kvázistacionárnu oblasť najväčšej výšky nerovnosti profilu drsnosti (15):

$$\hat{R}z = 31{,}649\, f^{1{,}010} \tag{15}$$

Grafické znázornenie priebehu optimalizačných funkcií (14) a (15) pre kvázistacionárnu oblasť je na *Obr.1* pre parameter optimalizácie Ra a na *Obr.2* pre parameter optimalizácie Rz .



Obr. 1 Závislosť Ra *a* f *pre kvázistacionárnu oblasť. Obr. 2 Závislosť* Rz *a* f *pre kvázistacionárnu oblasť.*

Záver

Na základe uvedených výpočtov možno konštatovať, že najväčší vplyv na skúmané parametre v Rz a Ra má posuv, ktorý v kvázistacionárnej oblasti, t.j. oblasti v blízkosti optima pre skúmanú oceľ ISO 683/1-87 pri daných rezných podmienkach je definovaný vzťahmi $\hat{R}a = 6{,}049\, f^{0{,}619}$ pre strednú aritmetickú odchýlku profilu drsnosti a $\hat{R}z = 31{,}649\, f^{1{,}010}$ pre najväčšiu výšku nerovnosti profilu drsnosti. Na základe výsledkov možno ďalej uviesť, že štatisticky nevýznamný vplyv na skúmané parametre profilu drsnosti ma hĺbka rezu, najvýraznejší sa prejavil vplyv posuvu, pričom rezná rýchlosť je nevýznamná v oblasti vypočítaného optima pričom jej vplyv na skúmané parametre v rozsahu použitých hodnôt je však preukázateľný.

Použitá literatúra

[1] BÁTORA, B., VASILKO, K. *Obrobené povrchy, Technologická dedičnosť, funkčnosť*. Trenčín : Trenčianska univerzita, 2000, 183 s.

[2] BÉKES, J., ANDONOV, I. *Analýza a syntéza strojárskych objektov a procesov*.1.vyd., Bratislava : ALFA, 1986, 376 s.

[3] BUMBÁLEK, B., ODVODY, V., OŠŤADAL, B.: *Drsnost povrchu*. Praha : SNTL, 1989, 330 s.

[4] KAŽIMÍR, I., BEŇO, J. *Teória obrábania – Návody na cvičenia*. Bratislava : ALFA, 1989, 280 s.

[5] MELOUN, M., MILITKÝ, J. *Kompendium statistického zpracování dat*. Praha : ACADEMIA, 2001, 766 s., ISBN : 80-200-1008-4.

[6] ZVÁRA, K. *Regresní analýza*. Praha : ACADEMIA, 1989, 245 s., ISBN 80-200-0125-5.

[7] MONKA, P. Analytické vzťahy pre maximálnu výšku nerovnosti povrchu pri sústružení, *In: Nové smery vo výrobnom inžinierstve 2002*, FVT TU v Košiciach, Prešov, ISBN 80 - 70099 - 828 – 8.

[8] MONKA, P. Teoretické vzťahy najvyššej nerovnosti profilu, *Výrobné inžinierstvo, 2-3*, 2003, II. Ročník, pp. 20-21, FVT TU v Košiciach, Prešov, ISSN 1335-7972-01.

Poďakovanie

Článok vznikol za podpory projektu VEGA č. 1/4157/07.

Conditioning, independence and S-additivity

Zuzana Havranová
Dept. of Mathematics
Slovak University of Technology
Radlinského 11, 813 68 Bratislava
e-mail: zuzana@math.sk

Abstract

This paper is devoted to studying of the Cartesian product of S-additive measures. It is shown under which conditions this Cartesian product is again S-additive. In the last part of the paper conditional measures and independence of S-additive measures are investigated.

In [3] we have shown how we can construct a Cartesian product of Łukasiewicz filters (Łukasiewicz filters are, in fact, special fuzzy measures). However, using the construction from [3], we have no guarantee that, when starting with S-additive filters (so-called (T_L, S)-filters), the resulting filter will inherit the S-additivity. In this paper we will construct two-dimensional S-additive measures for t-conorms S which are transformations of the Łukasiewicz t-conorm. Using these two-dimensional S-additive measures and some conjunctor, we will define conditional measures and study the independence of S-probability spaces with respect to the conditional measure. The S-additivity of conditional measures was already studied by P. Benvenuti and R. Mesiar in [1]. Here, we present another viewpoint at the conditioning and we show under which conditions an S-probability space is independent from another one.

Recall that a $\mathbf{t-conorm}$ $S : [0,1]^2 \to [0,1]$ is a commutative and associative operator having 0 as neutral element and 1 as its annihilator (see e.g., [4], [7]).
There are three basic types of t-conorm:
∘ nilpotent, i.e., such that there exist some $x, y < 1$ for which $S(x, y) = 1$
∘ cancelative t-conorm, i.e., $S(x, y_1) = S(x, y_2) \Leftrightarrow y_1 = y_2$ for each $x < 1$
∘ maximum t-conorm.

A **copula** (see e.g., [2], [4]) $C : [0,1]^2 \to [0,1]$ is an isotone operator having 1 as neutral element and fulfilling

$$C(x_2, y_2) + C(x_1, y_1) \geq C(x_1, y_2) + C(x_2, y_1).$$

A connection between copulas, two-dimensional distribution functions and dependence of random variables the reader can find in [8].
A **conjunctor** (see e.g., [4]) $\Gamma : [0,1]^2 \to [0,1]$ is an isotone operator having 1 as neutral element. The most important notion for our considerations will be the S-additivity:

Definition 1 *Let (X, σ, μ) be a measurable space with a fuzzy measure μ ($\mu(X) = 1$). Let S be a t-conorm. we say that μ is S-additive if the following holds:*

$$\mu(A \cup B) = S(\mu(A), \mu(B)), \qquad \text{if } A \cap B = \emptyset.$$

In this paper we will put $X = [0, 1]$. The concept of S-additivity was already studied by several authors, see e.g., [1], [4], [5].

1 S-additivity

If we have two measurable spaces, $([0, 1], \mathcal{B}_1, \mu_1)$, $([0, 1], \mathcal{B}_2, \mu_2)$, equipped with S-additive measures and we want to get their Cartesian product, then we can use an approach similar to that when constructing the Cartesian product of two probability spaces. Namely, in the latter case we use copulas. So, we are going to introduce so-called S-copulas.

Definition 2 *A two-dimensional operator $C_S : [0, 1]^2 \to [0, 1]$ is said to be an S-copula if and only if it is monotone in both coordinates and if the following are satisfied:*

C1 $\forall x \in [0, 1]$ $C_S(0, x) = C_S(x, 0) = 0$

C2 $\forall x \in [0, 1]$ $C_S(1, x) = C_S(x, 1) = x$

C3 $\forall x_1 \leq x_2 \in [0, 1], \forall y_1 \leq y_2 \in [0, 1]$ *yield*

$$C_S(x_2, y_2) \geq S\left(C_S(x_1, y_2), C_S(x_2, y_1) -_S C_S(x_1, y_1)\right) \tag{1}$$

where $a -_S b = \sup\{t \in [0, 1]; S(t, b) < a\}$ with the convention $\sup \emptyset = 0$

An important role in our consideration will play so-called S-probability spaces:

Definition 3 *Let $\mathcal{B}$ be the σ-algebra of Borel subsets of $[0, 1]$ and S be some continuous t-conorm. We say that $F : [0, 1] \to [0, 1]$ is a distribution function if $F(0) = 0$, $F(1) = 1$ and if it is non-decreasing. We say that an S-additive measure $\mu : \mathcal{B} \to [0, 1]$ is an S-probability, generated by F, if $\mu([0, x]) = F(x)$ for each $x \in [0, 1]$ and if for each interval $]x_1, x_2] \subseteq [0, 1]$ the following holds*

$$\mu(]x_1, x_2]) = F(x_2) -_S F(x_1)$$

The space $([0, 1], \mathcal{B}, \mu)$ will be called the S-probability space.

Nilpotent t-conorms

Lemma 1 ([4]) *A continuous t-conorm S is nilpotent if it is of the form $S(x, y) = f^{-1}(\min\{1, f(x) + f(y)\})$, where $f : [0, 1] \to [0, 1]$ is a left-continuous strictly increasing function with $f(0) = 0$, $f(1) = 1$.*

If f is strictly increasing and left-continuous, then for $a \geq b$, we have

$$a -_S b = f^{-1}(f(a) - f(b)). \tag{2}$$

Theorem 1 *Let $f : [0,1] \to [0,1]$ be a strictly increasing left-continuous function with $f(0) = 0$, $f(1) = 1$. Further, let the S be a nilpotent t-conorm, given by*

$$S(x,y) = f^{-1}(\min\{1, f(x) + f(y)\}). \tag{3}$$

Then a two-dimensional function operator $C_S : [0,1]^2 \to [0,1]$, defined by

$$C_S(x,y) = f^{-1}\left(C(f(x), f(y))\right) \tag{4}$$

is an S-copula, where C can be an arbitrary copula.

From now on, we will deal with a fixed t-conorm S, given by formula (3), where $f : [0,1] \to [0,1]$ will be a given continuous, strictly increasing function with $f(0) = 0$ and $f(1) = 1$.

Let us have two S-probability spaces, $([0,1], \mathcal{B}, \mu_1)$, $([0,1], \mathcal{B}, \mu_2)$, as these were introduced in Definition 3. Once we have S-copulas, when constructing a Cartesian product of the spaces $([0,1], \mathcal{B}, \mu_1)$ and $([0,1], \mathcal{B}, \mu_2)$, we may construct a Cartesian product of two S-probability spaces. By choosing some S-copula C_S we define a two-dimensional S-distribution function $F : [0,1]^2 \to [0,1]$ by

$$F(s,t) = C_S\left(F_1(s), F_2(t)\right) \tag{5}$$

Thereafter, in a usual way we can define an S-probability $\nu : \sigma\left(\mathcal{B} \times \mathcal{B}\right) \to [0,1]$, where $\sigma\left(\mathcal{B} \times \mathcal{B}\right)$ is the least σ-algebra containing all Cartesian products $B_1 \times B_2$, $B_1 \in \mathcal{B}$, $B_2 \in \mathcal{B}$.

Definition 4 *Let $([0,1], \mathcal{B}, \mu)$ be some S-probability space. The measure μ will be called S-uniform if and only if each interval $[a,b] \subseteq [0,1]$ yields $\mu([a,b]) = \varphi(b-a)$, where $\varphi : [0,1] \to [0,1]$ is some non-decreasing function with $\varphi(0) = 0$, $\varphi(1) = 1$.*

Comment 1 As Definition 4 says, an S-uniform measure of an interval depends only on the length of this interval. I.e., if the length of $[a,b]$ is $\frac{1}{n}$, we get

$$\begin{aligned} 1 &= S(\underbrace{\mu([a,b]), \mu([a,b]), \ldots, \mu([a,b])}_{n\times}) = f^{-1}(n \cdot f(\mu([a,b])))\\ \Rightarrow \qquad & \mu([a,b]) = f^{-1}\left(n^{-1}\right) \end{aligned} \tag{6}$$

Theorem 2 *Let $([0,1], \mathcal{B}, \mu)$ be the S-probability space, equipped with the S-uniform measure. Denote by $F : [0,1] \to [0,1]$ the corresponding S-distribution function and $\Pi_S(a,b) = f^{-1}(f(a) \cdot f(b))$. Let $\nu : \mathcal{B}^2 \to [0,1]$ be a measure, induced by the two-dimensional S-distribution function, G, defined by*

$$G(a,b) = \Pi_S(F(a), F(b))$$

Then ν is a two-dimensional S-uniform measure.

Proof. We have to prove the following

$$\nu([0,a] \times [0,b]) = \nu([c, a+c] \times [d, b+d])$$

where, in general,

$$\nu([c,a]\times[d,b]) = G(a,b) -_S S(G(c,d), G(a,b) -_S G(b,d))$$

At the left-hand-side we get

$$\nu([0,a]\times[0,b]) = f^{-1}(f(\mu([0,a]))\cdot f(\mu([0,b])))$$

at the right-hand-side we get

$$\begin{aligned}\nu([c,a+c]\times[d,b+d]) &= f^{-1}\left(f(G(a+c,b+d)) + f(G(c,d))\right.\\ &\quad \left. -f(G(c,b+d)) - f(G(a+c,d))\right) =\\ &= f^{-1}\left(f(F(a+c))\cdot f(F(b+d)) - f(F(c))\cdot f(F(b+d))\right.\\ &\quad \left. -\left(f(F(a+c))\cdot f(F(d)) - f(F(c))\cdot f(F(d))\right)\right)\end{aligned}$$

By formula 2 and since μ is S-uniform, we have

$$\begin{aligned}f(F(a+c)) - f(F(c)) &= f(F(a+c) -_S F(c)) = f(\mu([c,a+c])) = f(\mu([0,a]))\\ f(F(b+d)) - f(F(d)) &= f(F(b+d) -_S F(d)) = f(\mu([d,b+d])) = f(\mu([0,b]))\end{aligned}$$

and hence

$$\begin{aligned}\nu([c,a+c]\times[d,b+d]) &= f^{-1}\left(f(\mu([0,a]))\cdot f(F(b+d)) - f(\mu([0,a]))\cdot f(F(d))\right)\\ &= f^{-1}\left(f(\mu([0,a]))\cdot f(\mu([0,b]))\right)\end{aligned}$$

and the proof is done. □

Example 1 Let $f(x) = \sqrt{x}$. Then

$$W_S(a,b) = \left(\max\left\{0, \sqrt{a}+\sqrt{b}-1\right\}\right)^2,$$

the transformed product, Π_S, is given by

$$\Pi_S(a,b) = \left(\sqrt{a}\cdot\sqrt{b}\right)^2 = a\cdot b$$

the S-uniform measure μ is given by

$$\mu\left(\left[0,\frac{1}{n}\right]\right) = \frac{1}{n^2}$$

Cancelative t-conorms

Lemma 2 ([4]) *A continuous t-conorm S is cancelative if it is of the form*

$$S(x,y) = f^{-1}(f(x)\cdot f(y)),$$

where $f : [0,1] \to [0,1]$ is a left-continuous strictly increasing function with $f(0) = 0$, $f(1) = 1$.

If S is cancelative, then for $a \geq b$, we have

$$a -_S b = \begin{cases} f^{-1}\left(\frac{f(a)-f(b)}{1-f(b)}\right), & \text{if } b < 1 \\ 0, & \text{it } b = 1 \end{cases} \tag{7}$$

Lemma 3 *Let S be a cancelative t-conorm. The least S-copula is the drastic product T_D.*

Proof. The boundary conditions **C1** and **C2** are fulfilled, since T_D is a t-norm. We show that condition **C3** is also fulfilled. It is enough to consider the case when $x_2 = y_2 = 1$ and $x_1 < 1$, $y_1 < 1$. In that case we have

$$T_D(x_2, y_2) = 1,\ T_D(x_2, y_1) = y_1,\ T_D(x_1, y_2) = x_1,\ T_D(x_1, y_1) = 0$$

and hence

$$S\left(T_D(x_2, y_1) -_S T_D(x_1, y_1), T_D(x_1, y_2)\right) = S(y_1, x_1) < 1 = T_D(x_2, y_2)$$

□

Theorem 3 *There exist discrete one-dimensional uniform S-probability, but there is no uniform two-dimensional S-probability.*

Proof. It is enough to realize that if $x_0 \in]0, 1[$ be an arbitrary value then

$$\lim_{n\to\infty} S(\underbrace{x_0, \ldots, x_0}_{n\times}) = 1$$

Hence we get the discrete one-dimensional uniform S-probability. On the other hand, for no $x_0 \in]0, 1[$ there exists any $s \in]0, 1[$ such that

$$\lim_{n\to\infty} S(\underbrace{s, \ldots, s}_{n\times}) = 1.$$

. □

Maximum t-conorm

For maximum, S_M, and $b \geq a \in [0, 1]$ we get

$$b -_{S_M} a = \begin{cases} b, & \text{if } b > a \\ 0, & \text{if } b = a \end{cases}$$

Each S_M-probability, generated by a distribution function F, is of the form

$$\mu(]x_1, x_2]) = \begin{cases} F(x_2), & \text{if } F(x_1) < F(x_2) \\ 0, & \text{if } F(x_1) = F(x_2) \end{cases} \tag{8}$$

Theorem 4 *Each conjunctor is an S_M-copula.*

Proof. We have to prove condition **C3** for an arbitrary conjunctor Γ and each quadruple $x_1 \leq x_2$, $y_1 \leq y_2$.

$$S_M\left(\Gamma(x_2, y_1) -_{S_M} \Gamma(x_1, y_1), \Gamma(x_1, y_2)\right) \leq \max(\Gamma(x_2, y_1), \Gamma(x_1, y_2)) \leq \Gamma(x_2, y_2)$$

since conjunctors are nondecreasing in both variables. □

2 Conditioning and independence

The motivation for this section is the classical definition of the conditional probability (see, e.g. [6]):

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \tag{9}$$

assuming $P(B) \neq 0$, where A, B are from a σ-algebra $\mathcal{S}$ of P-measurable sets. The formula (9) can be rewritten into the form

$$P(A|B) = \sup\{x \in [0,1]; x \cdot P(B) \leq P(A \cap B)\}. \tag{10}$$

I.e., the conditional probability might be viewed at as a residual operator with respect to the product. Now, we generalize formula (10) in such a way that instead of a probability we will use some S-probability μ and instead of the product we may use, in fact, any conjunctor.

Definition 5 *Let $([0,1], \mathcal{B}, \mu_1)$ and $([0,1], \mathcal{B}, \mu_2)$ be S-probability spaces. We say that $\nu : \sigma(\mathcal{B} \times \mathcal{B}) \to [0,1]$ has marginal S-probabilities μ_1 and μ_2 if there exists some S-copula $\tilde{C}$ such that ν can be constructed using the corresponding one-dimensional S-distributions, F_1, F_2 and the S-copula $\tilde{C}$.*

Let $([0,1], \mathcal{B}, \mu_1)$ and $([0,1], \mathcal{B}, \mu_2)$ be S-probability spaces and $\nu : \sigma(\mathcal{B} \times \mathcal{B}) \to [0,1]$ be a two-dimensional S-probability with its marginal S-probabilities μ_1 and μ_2. Let us denote

$$\mathcal{B}_0 = \{B \in \mathcal{B}; \mu_2(B) \neq 0\}.$$

Then, for an arbitrary conjunctor Γ we denote by $\mu_\Gamma(\cdot|\cdot) : \mathcal{B} \times \mathcal{B}_0$ the following conditional measure:

$$\mu_\Gamma(A|B) = \sup\{x \in [0,1]; \Gamma(\mu_2(B), x) < \nu(A \times B)\} \tag{11}$$

Once having defined the conditional measure, we may consider the independence of the S-probability spaces $(X_1, \mathcal{B}_1, \mu_1)$ and $(X_2, \mathcal{B}_2, \mu_2)$ in the following way:

Definition 6 *Let us have S-probability spaces $(X_1, \mathcal{B}_1, \mu_1)$ and $(X_2, \mathcal{B}_2, \mu_2)$ and a two dimensional S-probability space $(X_1 \times X_2, \sigma(\mathcal{B}_1 \times \mathcal{B}_2), \nu)$, having its marginal S-probabilities μ_1 and μ_2. Let $\mu_\Gamma(\cdot|\cdot) : \mathcal{B}_1 \times \mathcal{B}_0$ be the conditional measure for some conjunctor Γ. We say that $(X_1, \mathcal{B}_1, \mu_1)$ is independent of $(X_2, \mathcal{B}_2, \mu_2)$ if $\mu_\Gamma(A|B) = \mu_1(A)$ for each $A \in \mathcal{B}_1$ and each $B \in \mathcal{B}_0$.*

Obviously, if $(X_1, \mathcal{B}_1, \mu_1)$ is independent of $(X_2, \mathcal{B}_2, \mu_2)$ with respect to μ_Γ, then the conjunctor Γ coincides with that one used to define the corresponding two-dimensional measure ν. This is the reason why we will speak of the independence with respect to a conjunctor Γ in the following theorems.

Theorem 5 *Let S be a nilpotent t-conorm, $([0,1], \mathcal{B}, \mu_1)$ and $([0,1], \mathcal{B}, \mu_2)$ be S-probability spaces having continuous S-probabilities μ_1 and μ_2, respectively. The S-probability space $([0,1], \mathcal{B}, \mu_1)$ is independent of $([0,1], \mathcal{B}, \mu_2)$ with respect to the conjunctor (S-copula) $\Gamma = \Pi_S$.*

Proof. As a corollary to Theorem 2 we get

$$\nu(A \times B) = \Pi_S(\mu_1(A), \mu_2(B)) = f^{-1}(f(\mu_1(A)) \cdot f(\mu_2(B)))$$

and formula (11) and the monotonicity and continuity of Π_S as a two-place function, imply

$$\mu_{\Pi_S}(A|B) = \mu_1(A)$$

On the other hand, if $\Gamma \neq \Pi_S$, we can find intervals $A = [a_1, a_2] \in \mathcal{B}_1$ and $B = [b_1, b_2] \in \mathcal{B}_0$ such that

$$\begin{aligned} \Gamma(\mu_1(A), \mu_2(B)) \quad &\neq \quad \nu(A \times B) = \Gamma(F_1(a_2), F_2(b_2)) -_S \Gamma(F_1(a_2), F_2(b_1)) \\ &-_S \quad (\Gamma(F_1(a_1), F_2(b_2)) -_S \Gamma(F_1(a_1), F_2(b_2))) \end{aligned}$$

what was to be proved. □

Lemma 4 *The conditional measure, defined using the S-copula Π_S, is of the form*

$$\mu(A|B) = f^{-1}\left(\frac{f(\mu_1(A))}{f(\mu_2(B))}\right)$$

Theorem 6 *Let S be a cancelative t-conorm, $([0,1], \mathcal{B}, \mu_1)$ and $([0,1], \mathcal{B}, \mu_2)$ be S-probability spaces. Then (except of trivial cases, i.e., if one of the S-probabilities is a Dirac measure) regardless of the choise of the conjunctor Γ, the S-probability space $([0,1], \mathcal{B}, \mu_1)$ does depend on $([0,1], \mathcal{B}, \mu_2)$ with respect to Γ.*

Proof. If at least one of the S-probability spaces is uniformly distributed, it follows by Theorem 3. If at least one of the S-probabilities is continuous, then we are able to reduce this case to that of the uniform distribution, since we can find a sequence of pairwise disjoint sets $\{A_i\}_i$, all having the same S-probability (remind that S is cancelative). If the S-probabilities are discrete, then either there are points x_1, x_2 such that $\mu_1(\{x_1\}) = 1$ and $\mu_2(\{x_2\}) = 1$. In this case it is easy to see that the spaces are not independent.

The other possibility is that the S-probabilities are discrete, but infinite. In this case the values of μ_1 and μ_2 give divergent sequences, but $\nu(A, \cdot)$ and $\nu(\cdot, B)$ are convergent sequences for each A, B, whose measures are not equal to 1. This consideration shows that $([0,1], \mathcal{B}, \mu_1)$ depends on $([0,1], \mathcal{B}, \mu_2)$ regardless of the conjunctor Γ. □

Theorem 7 *Let $S = S_M$. Then, let Γ be a conjunctor such that $\Gamma(x, \cdot)$ and $\Gamma(\cdot, y)$ are strictly increasing. S_M-probability space $([0,1], \mathcal{B}, \mu_1)$ is independent of $([0,1], \mathcal{B}, \mu_2)$ with respect to the conjunctor Γ.*

Proof. Since Γ is strictly increasing in both coordinates, it is cancelative. Take some $A \in \mathcal{B}$ with $\mu_1(A) > 0$. Then by formula (8) we get

$$\Gamma(\mu_\Gamma(A|B), \mu_2(B)) = \nu(A, B) = \Gamma(\mu_1(A), \mu_2(B))$$

By the cancelativity of Γ we get $\mu_\Gamma(A|B) = \mu_1(A)$. In case $\mu_1(A) = 0$, the identity $\mu_\Gamma(A|B) = \mu_1(A)$ is trivial. □

Acknowledgement. This work was supported by the VEGA grant agency, grant number 1/4026/07.

References

[1] Benvenuti, P., Mesiar, R., Pseudo-additive measures and triangular-norm-based conditioning, *Annals of Mathematics and Artificial Inteligence* **35** (2002), 63-69.

[2] Calvo, T., Kolesárová, A., Komorníková, M., Mesiar, R., A review of aggregation operators, University Press, Alcalá de Henares, Spain, 2001.

[3] Kalina, M., Łukasiewicz filters and their Cartesian products, in: Proceedings of EUSFLAT 2005 conference, Barcelona, Spain 1301-1306.

[4] Klement, E.P., Mesiar, R., Pap, E.: Triangular norms, *Trends in Logic, Studia Logica Library*, volume 8, Kluwer Acad. Publishers, Dordrecht, 2000.

[5] Pap, E.: Null-additive measures, Kluwer Acad. Publishers Dordrecht, Boston, London and Isterscience,Bratislava, 1995.

[6] Rényi, A.: Wahrsheinlichtkeitsrechnung mit einem Anhang über Informationstheorie, VEB deutscher Verlag der Wissenschaften, Berlin, 1962.

[7] Schweizer, B., Sklar, A.: Probabilistic metric spaces, North-Holland, New York, Amsterdam, Oxford, 1983.

[8] Volauf, P.: O asociácií náhodných veličín a kopuliach, in: Forum Statisticum Slovacum 3 (2005), 91-98.

Fuzzy preference structures and triangular norms

Dana Hliněná
Dept. of Mathematics
FEEC Brno University of Technology
Technická 8, 616 00 Brno, Czech Rep.
hlinena@feec.vutbr.cz
Pavol Král'
Dept. of Quantitative Methods and Informatics
Faculty of Economics Matej Bel University
Tajovského 10, 975 90 Banská Bystrica, Slovakia
pavol.kral@umb.sk

Abstract
In multicriterial optimization, one of the tasks is training with input data and finding the right utility function. The first step in determining the right utility function is finding an appropriate ordering of criteria from given expert evaluation. In our contribution we discuss and compare three possible approaches to this problem. All three approaches are illustrated on a practical example.
Keywords: Multicriteria optimization, Preference structures, Fuzzy preference structures, Similarity measures, Complementarity, Redundance.

1. Introduction to preference structures and fuzzy preference structures

A preference structure is a basic concept of preference modeling. In a classical preference structure (PS), a decision-maker makes three decisions for any pair (a, b) from the set $\boldsymbol{A}$ of all alternatives. A preference structure (PS) on a set $\boldsymbol{A}$ is a triplet (P, I, J) of binary relations on $\boldsymbol{A}$ such that

(ps1) I is reflexive, P and J are irreflexive.

(ps2) P is asymmetric, I and J are symmetric.

(ps3) $P \cap I = P \cap J = I \cap J = \emptyset$.

(ps4) $P \cup I \cup J \cup P^t = \mathbf{A} \times \mathbf{A}$ where $P^t(x, y) = P(y, x)$.

A preference structure can be characterized by the reflexive relation $R = P \cup I$ called the large preference relation. The relation R can be interpreted as

$$(a, b) \in R \Leftrightarrow a \text{ is prefered to } b \text{ or } a \text{ and } b \text{ are indifferent.}$$

Decision-makers are often uncertain even inconsistent in their judgements. The restriction on two-valued relations has been an important drawback in their practical use. A natural demand led researchers to the introduction of a fuzzy preference structure (FPS). The original idea of using numbers between zero and one to describe the strength of links between two alternatives goes back to Menger. The introduction of fuzzy relations allowed to express degrees of preference, indifference and incomparability. Of course, the attempts to simply replace the notion used in the definition of (PS) by their fuzzy equivalents have brought some problems. To define (FPS) it is necessary to consider some fuzzy connectives. We shall consider a continuous De Morgan triplet (T, S, N) consisting of a continuous t-norm T, continuous t-conorm S and a strong negator N such that $T(x, y) = N(S(N(x), N(y)))$. The main problem lies in the fact that the completeness condition (ps4) can be written in many forms, e.g.:

$$co(P \cup P^t) = I \cup J, P = co(P^t \cup I \cup J), P \cup I = co(P^t \cup J).$$

Let (T,S,N) be a De Morgan a triplet. A fuzzy preference structure (FPS) on a set $\boldsymbol{A}$ is a triplet (P, I, J) of binary fuzzy relations on $\boldsymbol{A}$ such that:

(f1) I is reflexive, P and J are irreflexive. $I(a, a) = 1$, $P(a, a) = J(a, a) = 0$

(f2) P is T-asymmetric, I and J are symmetric. $T(P(a, b), P(b, a)) = 0$

(f3) $T(P, I) = T(P, J) = T(I, J) = 0$ for all pairs of alternatives

(f4) $(\forall (a, b) \in A^2) S(P, P^t, I, J) = 1$ or $N(S(P, I)) = S(P^t, J)$ or another completeness conditions.

Note that it was proved [2, 9] that reasonable constructions of fuzzy preference structure (FPS) should use a nilpotent t-norm only. Since any nilpotent t-norm (t-conorm) is isomorphic to the Lukasiewicz t-norm (t-conorm), it is enough to restrict our attention to the De Morgan triplet $(T_L, S_L, 1 - x)$.

2. Concepts of similarity

In multicriterial optimization, one of the tasks is training with input data and finding the right utility function. The first step in determining the right utility function is finding an appropriate ordering of criteria from given expert evaluation. In this step we have to solve the problem with similarity of sets, especially similarity of relations. The main aim of our contribution is to find a suitable way of comparing fuzzy preference relations and to illustrate our approaches on practical example.

Example 1 *Let us consider six students and their grades in mathematics, physics, informatics, physical education, English, music lessons. Students are evaluated by a test score (mathematics, physics, informatics) and by a qualitative score which is given by teacher (physical education, English, music lessons). Students' names, grades and expert's global score are given in Table 1. We can assign the numerical score to each criterion (Table 2). We can also construct the following fuzzy preference relations:* FP_M, FP_P, FP_I, FP_{PE}, FP_E, FP_{ML}, FP_{GS}, *which are in a correspondence with our criteria – mathematics, physics, informatics, physical education, English, music lessons and the global score.*

	mathematics	physics	informatics	physical education	English	music lessons	expert evaluation	rank in the class
Paul	99	95	100	C	A	C	good	2
James	90	90	95	B	B	C	good	1
Eve	80	75	70	A	B	B	average	1
Jane	75	65	85	C	C	A	average	2
Patty	70	60	60	B	D	A	bad	1
John	60	65	80	A	E	D	bad	2

Table 1: Grades of the different students

	mathematics	physics	informatics	physical education	English	music lessons	global score
Paul	1	1	1	0,5	1	0,5	1
James	1	1	1	0,75	0,75	0,5	0,875
Eve	0,75	0,625	0,5	1	0,75	0,75	0,75
Jane	0,625	0,375	0,875	0,5	0,5	1	0,625
Patty	0,5	0,25	0,25	0,75	0,25	1	0,375
John	0,25	0,375	0,75	1	0	0,25	0,25

Table 2: Numerical scores on criteria

For example, the value of fuzzy preference in mathematics (FP_M) *for Paul and James is computed from Table 2 as* $FP_M(P,J) = \max\{x_M^P - x_M^J, 0\}$, *where* x_M^P *and* x_M^J *are Paul's and James's mathematics score in Table 2, etc. This is one possible way of fuzzification a preference relation. Our task is to find a suitable way of comparing these fuzzy preference relations.*

Our first approach is based on the formula

$$X \succ Y \iff \frac{\sum\limits_{i,j} |\{fp_{x_{ij}}\} \cap \{fp_{gs_{ij}}\}|}{\sum\limits_{i,j} |\{fp_{x_{ij}}\} \triangle \{fp_{gs_{ij}}\}|} > \frac{\sum\limits_{i,j} |\{fp_{y_{ij}}\} \cap \{fp_{gs_{ij}}\}|}{\sum\limits_{i,j} |\{fp_{y_{ij}}\} \triangle \{fp_{gs_{ij}}\}|}, \tag{1}$$

where X, Y are our criteria (mathematics, physics, English, informatics, music lessons, physical education), m is a number of alternatives, $fp_{x_{ij}}, fp_{y_{ij}}$ are values of fuzzy preference structures of criterion X, Y, $fp_{gs_{ij}}$ are values of fuzzy preference relation which is based on expert's global score and $i, j \in \{1, 2 \dots m\}$. Note that our intersection $\cap$ and symmetric difference $\triangle$ are ordinary. The basic idea is evident. The greater similarity between FP_X and FP_{GS} means the greater importance of a criterion. Using the Formula 1, we obtain the following ordering of criteria

$$mathematics \succ physics = English \succ informatics \succ$$

$$\succ \text{ music lessons} \succ \text{physical education.}$$

In ***our second approach*** we use the following formula

$$Sim(FP_X, FP_Y) = 1 - \frac{\sum_{i,j(i\neq j)} |\mathrm{fp}_{x_{ij}} - \mathrm{fp}_{y_{ij}}|}{m^2 - m}, \tag{2}$$

where m is a number of alternatives, $fp_{x_{ij}}(fp_{y_{ij}})$ are values of fuzzy preference structures of criterion X (Y), and $i, j \in \{1, 2 \ldots m\}$. Now we obtain the following ordering of criteria

$$mathematics \succ English \succ physics \succ informatics \succ$$

$$\succ \text{ music lessons} = \text{physical education.}$$

Remark 1. Let ρ be an arbitrary metric taking values from the unit interval. We can define the metric based similarity as follows

$$Sim_\varrho(FP_X, FP_Y) = 1 - \frac{\sum_{i,j(i\neq j)} \varrho(\mathrm{fp}_{x_{ij}}, \mathrm{fp}_{y_{ij}})}{m^2 - m}. \tag{3}$$

By triangle inequality we get that Sim_ϱ is T_L- transitive, more Sim_ϱ is also $T-$transitive for each $T \leq T_L$. Similarity Sim_ϱ based on the discrete metric is T_M-transitive, subsequently this similarity is $T-$ transitive for each $t-$norm T. Moreover, we can construct, for each real number $p \in]0, 1]$, a metric ϱ_p using

$$\varrho_p(x, y) = \begin{cases} p & \text{if } x \neq y, \\ 0 & \text{otherwise.} \end{cases}$$

Then similarity Sim_{ϱ_p} is T_M- transitive, too.

Our third approach is based on scalar cardinality of fuzzy sets and fuzzy extension of set operations (see [1, 10]). We can generalize Formula (1) for $X \neq Y$ in the following way:

$$X \succ Y \iff \frac{\mathrm{card}\,(FP_X \cap_T FP_{GS})}{\mathrm{card}\,(FP_X \widetilde{\triangle} FP_{GS})} > \frac{\mathrm{card}\,(FP_Y \cap FP_{GS})}{\mathrm{card}\,(FP_Y \widetilde{\triangle} FP_{GS})}. \tag{4}$$

To avoid problems with zero in denominator, we restrict ourselves in the rest of the paper only to cardinality patterns based on function f which is positive for arbitrary real x except zero and fuzzy symmetric differences for which the following is satisfied:

$$A \neq B \Rightarrow A \widetilde{\triangle} B \neq 0.$$

For simplicity we assume the intersection based on minimum, the cardinality pattern $f = id$ and the symmetric difference pattern $h(x, y) = |x - y|$ for all $x, y \in [0, 1]$.

$$X \succ_{T_M} Y \iff \frac{\sum\limits_{i=1,j=1}^{m,m} \min(FP_X(x_{ij}), FP_{GS}(x_{ij}))}{\sum\limits_{i=1,j=1}^{m,m} |FP_X(x_{ij}) - FP_{GS}(x_{ij})|} >$$

$$> \frac{\sum_{i=1,j=1}^{m,m} \min(FP_Y(x_{ij}), FP_{GS}(x_{ij}))}{\sum_{i=1,j=1}^{m,m} |FP_Y(x_{ij}) - FP_{GS}(x_{ij})|}, \tag{5}$$

where m is number of attributes, $FP_X(x_{ij})$ (analogously $FP_Y(x_{ij})$, $FP_{GS}(x_{ij})$) is the value of the cell in the i-th row and j-th column of the table describing the fuzzy preference structure FP_X. In our case we obtain the following ordering of criteria

$$math. \succ_{T_M} English \succ_{T_M} physics \succ_{T_M} inform. \succ_{T_M}$$

$$\succ_{T_M} \text{music lessons} \succ_{T_M} \text{physical education.}$$

It is easy to see that our third approach is straightforward fuzzification of the first one. It is also obvious that the similarity measure presented above is a kind of the cardinality-based similarity measure. One family of the rational cardinality-based similarity measures which contains several appropriate candidates for fuzzification was proposed by De Baets, Meyer and Naessens in [4]. Our similarity measure is a member of that class, but it does not belongs to proposed candidates for fuzzification.

Remark 2. We can see that evaluation based on T_P provides the ordering of criteria $\succ_{T_P}$, which is the same as the ordering of criteria $\succ_{T_M}$. If orderings of criteria $\succ_{T_M}$ and $\succ_{T_P}$ are equivalent, from monotonicity it is easy to show that for each t-norm T, $T_M \geq T \geq T_P$, we get the same ordering $\succ_T$. This useful property can be generalized to any triplet of comparable t-norms.

3. Conclusion

In comparison of our approaches, the first one comes out worst due to indiference between important criteria, and the best one coming out is the last one as the criteria linearly ordered. However this is just behaviour in our specific example it can not be readily generalized without further investigation. Also the complete comparison with other authors, e.g. B. De Baets (see [4]), is essential. Especially the further study of a connections between our similarity measure and the above mentioned family of rational cardinality-based similarity measures will be the object of our future research.

Another natural question is whether it is possible to find some comparable ordering of criteria using statistical methods, for example some kind of regression. The main idea could be similar to those used in discriminant analysis or logistic regression. A dependent variable representing cases is given and we try to find some regression function(s) which allows us to distinguish our cases. In our example the dependent variable can be the global rank of students. If we succeed, then the standardized coefficients of this regression function(s) can be assumed as weights of importance of our criteria. It is obvious that it is nearly impossible to find an appropriate statistical model for our example. The first problem is the small number of cases (we have only six students). Another two important problems are ordinary variables and too many categories in our dependent variable. It is also problematic to satisfy the assumptions of the methods, such as normality of residuals

(for example due to expert evaluation). So we can conclude that our proposed methods can not be now compared directly with the statistical methods but statistical methods can be used to verify our final ordering of criteria if large samples are used. The most promising seem to be logistic regression, principal component analysis, factor analysis, ordinary regression and clustering analysis.

Acknowledgements

Dana Hliněná has been supported by Project 1ET100300517 of the Program "Information Society" and by Project MSM0021630529 of the Ministry of Education. Pavol Král' has been supported by the grant VEGA 1/2002/05.

References

[1] Benvenuti, P., Vivona, D., Divari, M. *Divergence and fuzziness measures*. Soft Computing - A Fusion of Foundations, Methodologies and Applications, Springer-Verlag Berlin, Heidelberg, 2004.

[2] J. Fodor and M. Roubens. *Fuzzy preference modeling and Multicriteria Decision Support*. In Kluwer, Dordrecht, 1994.

[3] B. De Baets and R. Mesiar. *Pseudo-metrics and T-equivalences*. In The Journal of Fuzzy Mathematics 5 (2), Los Angeles, 1997.

[4] B. De Baets, H. De Meyer, H. Naessens. *A Class of Rational Similarity Measures*. In Proc. IPMU, Madrid, July 3-7, Vol. III, pp. 1356-1361, 2000.

[5] M. Grabisch and M. Roubens. *An axiomatic approach of interaction in multicriteria decision making*. In 5th Eur. Congr. on Intelligent Techniques and Soft Computing (EUFIT 97), Aachen, Germany, 1997.

[6] M. Grabisch and M. Roubens. *Application of the Choquet integral in multicriteria decision making*. In M. Grabisch, T. Murofushi, and M. Sugeno, editors, Fuzzy Measures and Integrals - Theory and Applications, pages 348–374. Physica Verlag, 2000.

[7] D. Hliněná and P. Vojtáš *A note on an example of use of fuzzy preference structures*. Submitted.

[8] M. Šabo. *The history and new results in fuzzy preference structures*, In Proccedings of Aplimat, Bratislava, 531–537, 2005.

[9] B. Van De Walle and B. De Baets and E. Kerre. *A comparative study of completeness conditions in fuzzy preference structures*. In Proc. IFSA'97, Prague, vol III, 74–79. 1997.

[10] Wygralak, M. *Cardinalities of Fuzzy Sets*, Studies in Fuzziness and Soft Computing 118, Springer-Verlag Berlin,Heidelberg, 2003.

Využitie MS Excel pri navrhovaní regulačných diagramov priemerov a rozpätí.

Stella Hrehová

Fakulta výrobných technológií

Technická univerzita Košice

Bayerova 1

080 05 Prešov

hrehoval@post.sk

Abstrakt

V príspevku sú uvedené možnosti využitia prostriedkov tabuľkového procesora MS Excel pri tvorbe regulačných diagramov priemerov a rozpätí. Sú využité aj možnosti zjednodušenia tvorby využitím možností automatizácie činnosti pomocou jazyka Visual Basic for Application pre MS Excel.

Kľúčové slová: priemer, rozpätie, riadenie kvality, MS Excel, Visual Basic for Application

Úvod

Neustále narastajúci význam využitia štatistických metód v procese riadenia kvality má vplyv na výraznom zvýšení úrovne kvality, spoľahlivosti výrobkov a výrobného procesu. Ťažiskom štatistického riadenia kvality sú exaktné štatistické metódy a špecifické grafické výstupy. Ich praktická aplikácia bez využitia výpočtovej techniky naráža na mnohé bariéry. Jednou z možnosti ako hodnotiť spôsobilosť výrobného procesu je aj tvorba regulačných diagramov.
Regulačný diagram je nástroj štatistickej regulácie procesu, ktorý umožňuje operatívne určovať, či je proces stabilný, alebo nestabilný.

Typický regulačný diagram obsahuje :

- *Centrálnu priamku CL (center line*), ktorá reprezentuje očakávanú hodnotu regulovanej veličiny, keď je proces stabilný,
- *Hornú regulačnú hranicu UCL(upper control limit),* a *dolnú tolerančnú hranicu LCL (lower control limit),* ktoré sa tiež počítajú z údajov získaných v čase, keď bol proces stabilný,

- Body pozorovania, z ktorých sú vždy dva bezprostredne susedné spojené úsečkou

Regulačný diagram sa zostrojuje na báze získaných meraní sledovaného ukazovateľa kvality procesu. Do regulačného diagramu sa zakresľujú individuálne hodnoty alebo hodnoty nejakej výberovej charakteristiky. Regulačné hranice definujú variabilitu výberovej charakteristiky spôsobenú náhodnými príčinami.
Bod mimo regulačných hraníc indikuje možnú prítomnosť vymedziteľných príčin. Regulačné hranice možno charakterizovať ako hranice prognózovanej variability danej systémom, t.j. spôsobenej náhodnými príčinami.

Parametre regulačného diagramu priemerov určíme na základe vzťahov :

$$UCL = \overline{\overline{x}} + A_2 \overline{R}$$
$$CL = \overline{\overline{x}}$$
$$LCL = \overline{\overline{x}} - A_2 \overline{R}$$

Hodnoty A_2 sú tabelované pre rôzne hodnoty *n.*

Parametre regulačného diagramu rozpätí určíme na základe vzťahov :

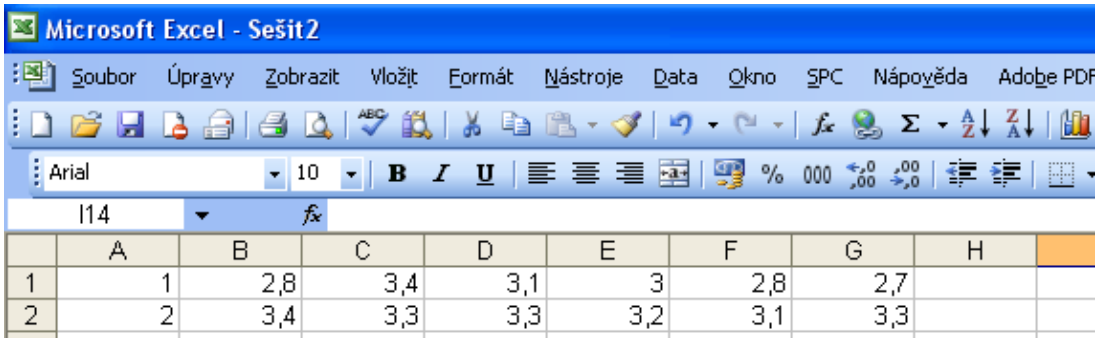
$$UCL = \overline{R} D_4$$
$$CL = \overline{R}$$
$$LCL = \vec{R} D_3$$

Hodnoty D_3 a D_4 sú pre rôzne *n* tabelované a uvedené v tabuľkách.

Využitie MS Excel pri tvorbe regulačných diagramov priemerov a rozpätí

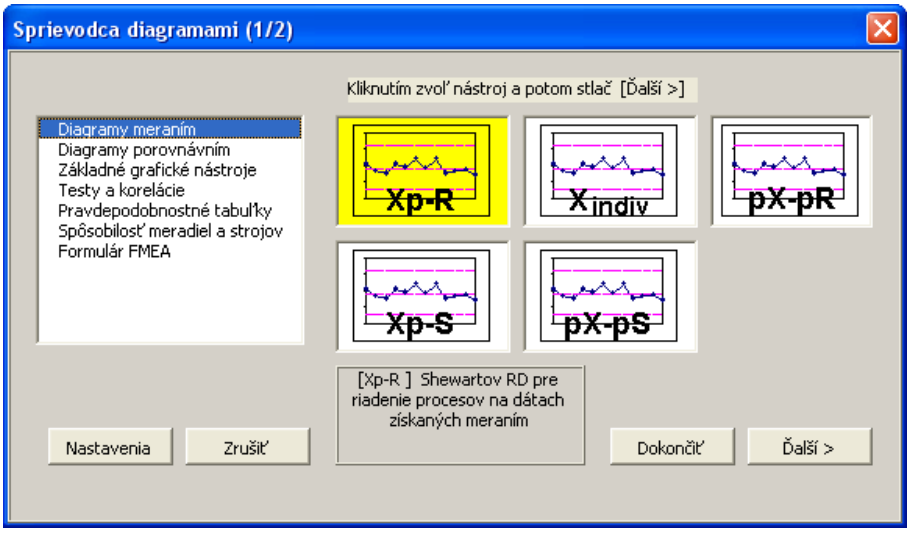
MS Excel je tabuľkový procesor, ktorý obsahuje veľa vložených funkcií. Jeho výhodou je jednoduchosť pri tvorbe grafov. Umožňuje užívateľovi vytvoriť graf v takom tvare, v akom ho požaduje.

Pre hodnotenie spôsobilosti výrobného procesu bola vytvorená nadstavba SPC firmy FBE. Po nainštalovaní sa v základnom prostredí objaví nové príkazové slovo „ SPC“. Tento softvér je určený pre štatistickú reguláciu procesov (meraním aj porovnávaním), analýzu spôsobilosti procesov, Paretovu analýzu, tvorbu korelačných a autokorelačných diagramov. Okrem štandardných postupov SPC podľa ISO, VDA a QS9000 obsahuje aj metódu pre krátke dávky a malé série (ASQC). Ponúka zjednodušenie tvorby diagramov možnosťou definovať si užívateľské tlačidlá, v ktorých sa zapamätajú všetky nastavenia a popisy. Obsahuje diagramy s formulármi pre analýzu systému merania (R-R) a spôsobilosti strojov.
Grafické výstupy sú plne kompatibilné so súbormi Microsoft Excel. Testy dát sú doplnené o testy normality pomocou Q-Q plot diagramu.



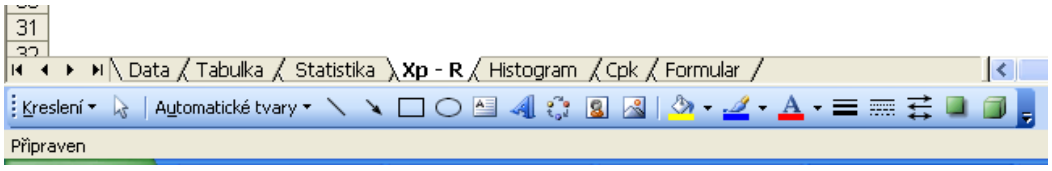
Obr. 1 Pridanie nadstavby SPC

Po vybratí možnosti „SPC“ dostaneme na výber



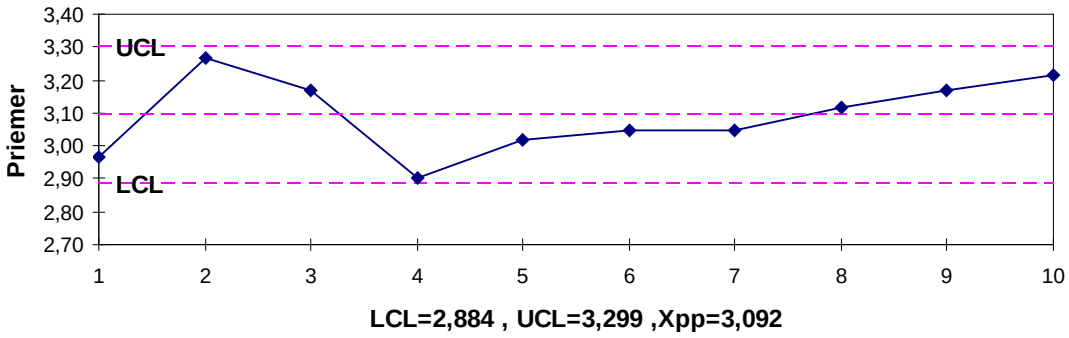
Obr. 2 Základné okno SPC

Potvrdením možností sa vytvoria nové listy v zošite s príslušným pomenovaním, takže máme možnosť sa prepínať do listov, podľa toho čo chceme práve analyzovať.

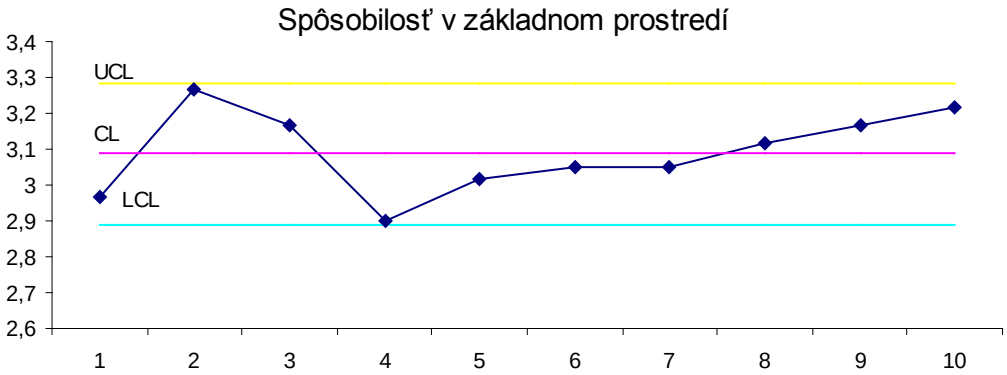


Obr.3 Označenie jednotlivých listov

V nasledujúcich dvoch obrázkoch sú grafy spôsobilosti výrobného procesu získané z nadstavby SPC (Obr. 4) a graf, ktorý bol vytvorený v základnom prostredí, tzn. ručne počítané hodnoty. (Obr. 5).



Obr. 4 Graf spôsobilosti výrobného procesu z nadstavby MS Excel - SPC



Obr. 5 Graf spôsobilosti výrobného procesu v základnom prostredí

Záver

Na trhu je ponuka veľkého množstva špecializovaného softvéru, ktoré sa zaoberajú štatistickým riadením výrobného procesu. Líšia sa svojimi možnosťami spracovávať dáta, grafickou podporou riešených úloh a pod. Predstavený systém ma výhodu v tom, že zadávanie dát sa vykonáva v známom prostredí MS Excelu.

Je potrebné však pripomenúť, že bez znalosti tvorby jednotlivých grafov spôsobilosti výrobného procesu, je použitie akéhokoľvek softvérového prostriedku len získaním výsledkov bez pochopenia podstaty.

Literatúra

[1] Chajdiak, J. : *Štatistické riadenie kvality*, STATIS, Bratislava 1998, ISBN 80-85659-12-3

[2] Terek, M. - Hrnčiarová, Ľ. : *Štatistické riadenie kvality*, EKONÓMIA, Bratislava 2004, ISBN 80-89047-97-1

[3] www.fbe.sk

POMEROVÉ A REGRESNÉ BODOVÉ ODHADY STREDNEJ HODNOTY A ÚHRNU V KONEČNOM ZÁKLADNOM SÚBORE

Ľubica Hrnčiarová, Milan Terek

1. Úvod

Bodový odhad pomernej veličiny *R* môže byť nielen cieľom, ale tiež prostriedkom k získaniu odhadu úhrnu UH_X alebo strednej hodnoty $\mu_{k,X}$ študovanej premennej *X* v konečnom základnom súbore.

Podmienkou je, aby bol presne známy, z nezávislého zdroja, úhrn UH_Y alebo stredná hodnota $\mu_{k,Y}$ pomocnej premennej *Y*. Cieľom metódy pomerového odhadu je zlepšiť presnosť odhadu úhrnu UH_X alebo strednej hodnoty $\mu_{k,X}$ študovanej premennej *X* v konečnom základnom súbore na základe využitia informácie o korelácii študovanej premennej *X* a pomocnej premennej *Y*.

Keď sa skúma závislosť medzi *X* a *Y* a je približne lineárna, ale priamka neprechádza začiatkom súradnicovej sústavy, vtedy je lepšie použiť namiesto pomerového odhadu, regresný odhad (bodový odhad založený na lineárnej regresii).

2. Pomerové a regresné bodové odhady strednej hodnoty a úhrnu v konečnom základnom súbore

2.1 Bodové odhady pomeru *R* v základnom súbore

Pri jednoduchom náhodnom výbere sa ukazuje ako „prirodzený“ bodový odhad pomeru *R* v základnom súbore, *výberový pomer* :

$$\hat{R} = \frac{U_x}{U_Y} = \frac{U_x / n}{U_y / n} = \frac{\overline{X}}{\overline{Y}}, \tag{1}$$

kde

U_X je výberový úhrn premennej *X*, U_Y je výberový úhrn premennej *Y* a

$\frac{U_x}{U_y}$ predstavuje pomernú veličinu za výberový súbor,

$\overline{X}$ je výberový priemer premennej *X*, $\overline{Y}$ je výberový priemer premennej *Y* a

n je rozsah výberu.

Tejto štatistike sa podobá *jednoduchý aritmetický priemer pomerov R_i vo výberovom súbore*:

$$\bar{r} = \frac{1}{n}\sum_{i=1}^{n} R_i = \frac{1}{n}\sum_{i=1}^{n} \frac{X_i}{Y_i}. \tag{2}$$

Súčasne budeme uvažovať ako možný bodový odhad pomeru R v základnom súbore, *štatistiku*:

$$\frac{\overline{X}}{\mu_{Y,k}}. \tag{3}$$

kde

$\mu_{k,Y}$ je stredná hodnota pomocnej premennej Y v konečnom základnom súbore .

Prednosťou tejto štatistiky je, že nie je podielom dvoch náhodných veličín.

Každá z vyššie uvedených výberových charakteristík, aby bola vhodným bodovým odhadom (estimátorom) parametra základného súbore, mala by spĺňať určité požiadavky. Bodový odhad by mal byť predovšetkým nevychýlený, alebo aspoň asymptoticky nevychýlený a ďalej by mal byť konzistentný. Keď súčasne viacej odhadov spĺňa tieto obidve požiadavky, dávame prednosť odhadu s najnižším rozptylom, t.j. najvýdatnejšiemu.

Vlastnosti uvedených bodových odhadov porovnáme v nasledujúcom príklade. Pre jednoduchosť riešenia uvažujeme základný súbor so štyrmi prvkami. Hodnoty študovanej premennej X a pomocnej premennej Y sú uvedené v tab. 1.

Tabuľka 1 Základný súbor

Prvok	x_s	y_s	R_s
A	4,26	4	1,07
B	2,02	2	1,01
C	1,60	1	1,60
D	3,30	3	1,10
	UH_x= 11,18	UH_y= 10	R = 1,118

Predpokladáme jednoduchý náhodný výber. Bodové odhady pomeru R v konečnom základnom súbore obsahuje tab. 2.

Stredné hodnoty a stredné kvadratické chyby bodových odhadov pomeru R v základnom súbore sú uvedené v tab. 3.

Keď porovnáme bodové odhady pomeru R v základnom súbore v tab. 3, bodový odhad (3) je síce najmenej vychýlený, ale má viacnásobne vyššiu strednú kvadratickú chybu odhadu ako odhad (1), t.j. výberový pomer $\hat{R}$.

Tabuľka 2 Výpočtová tabuľka bodového odhadu pomeru R v základnom súbore

Výber	$i = 2$			
$i = 1$	A	B	C	D
			Hodnoty $\bar{x}/\bar{y}$	
A	**1,07**	1,05	1,17	1,08
B	1,05	**1,01**	1,21	1,06
C	1,17	1,21	**1,60**	1,23
D	1,08	1,06	1,23	**1,10**
			Hodnoty $\frac{1}{n}\sum_i^n \frac{x_i}{y_i}$	
A	**1,07**	1,04	1,34	1,09
B	1,04	**1,01**	1,31	1,06
C	1,34	1,31	**1,60**	1,35
D	1,09	1,06	1,35	**1,10**
			Hodnoty $\bar{x}/\mu_{Y,k}$	
A	**1,70**	1,26	1,17	1,51
B	1,26	**0,81**	0,72	1,06
C	1,17	0,72	**0,64**	0,98
D	1,51	1,06	0,98	**1,32**

Tab. 3 Stredné hodnoty $E(\bullet)$ a stredné kvadratické chyby $MSE(\bullet)$ bodových odhadov pomeru R v základnom súbore

Bobový odhad pomeru R	$E(\bullet)$	$MSE(\bullet)$
$\bar{x}/\bar{y}$	1,13	0,005524
$\frac{1}{n}\sum_i^n \frac{x_i}{y_i}$	1,20	0,025034
$\bar{x}/\mu_{y,k}$	1,12	0,059557

Odhad pomeru R v konečnom základnom súbore pomocou výberového pomeru $\hat{R}$ si teória vždy najviac všímala a rovnako aj v praxi sa najviac používa.

2.2 Pomerové odhady úhrnu a strednej hodnoty v konečnom základnom súbore

Platí:

$$R = \frac{UH_X}{UH_Y} = \frac{\mu_{k,X}}{\mu_{k,Y}}$$

Jednoduchým násobením dostaneme:

$$UH_X = R \cdot UH_Y \qquad \text{a} \tag{4}$$

$$\mu_{k,X} = R \cdot \mu_{k,Y} \tag{5}$$

Predpokladajme jednoduchý náhodný výber. Keď nahradíme vo výrazoch (4) a (5) neznámy pomer R v základnom súbore jeho výberovým odhadom (1), t.j. výberovým pomerom $\hat{R} = \frac{U_x}{U_Y} = \frac{U_x / n}{U_y / n} = \frac{\overline{X}}{\overline{Y}}$, dostaneme pomerový odhad úhrnu, popr. strednej hodnoty premennej X v základnom súbore.

Pomerový bodový odhad úhrnu v konečnom základnom súbore je:

$$\hat{UH}_X = \hat{R} \cdot UH_Y = \frac{U_x}{U_y} \cdot UH_Y = \frac{\overline{X}}{\overline{Y}} \cdot UH_Y \tag{6}$$

Pomerový bodový odhad strednej hodnoty $\mu_{k,X}$ v konečnom základnom súbore je:

$$\hat{\mu}_{k,X} = \hat{R} \cdot \mu_{k,Y} = \frac{U_x}{U_y} \cdot \mu_{k,Y} = \frac{\overline{X}}{\overline{Y}} \cdot \mu_{k,Y} \tag{7}$$

kde

U_x, U_y sú výberové úhrny a

$\frac{U_x}{U_y}$ predstavuje pomernú veličinu za výberový súbor ,

$\overline{Y}$ $\overline{X}$ sú výberové priemery .

Vlastnosti odhadov (6) a (7) sú bezprostredne určené vlastnosťami pomerového odhadu $\hat{R}$. Vo všeobecnosti ide o vychýlený, ale pritom konzistentný odhad R . Vychýlenie pomerového odhadu môžeme znížiť, ak:

- rozsah výberového súboru n je veľký,
- hodnoty výberového pomeru $\frac{n}{N}$ a
- strednej hodnoty $\mu_{k,Y}$ sú vysoké, ale
- hodnota rozptylu $\sigma^2_{k,Y}$ je nízka a
- korelácia medzi premennými X, Y sa blíži k 1.

Vzorce **rozptylov pomerových odhadov úhrnu, strednej hodnoty konečného základného súboru** sa získajú jednoducho ako súčiny rozptylov odhadov pomeru v konečnom základnom súbore a hodnoty UH_Y^2 (popr. $\mu_{k,Y}^2$).

$$D\left(\hat{UH}_X\right) = UH_Y^2 \cdot D\left(\hat{R}\right) \quad \text{a} \quad D\left(\hat{\mu}_{\mathrm{k,X}}\right) = \mu_{k,Y}^2 \cdot D\left(\hat{R}\right) \tag{8}$$

Výber bez opakovania

Rozptyl odhadu úhrnu v konečnom základnom súbore je

$$D\left(\hat{UH}_X\right) \cong UH_Y^2\left(1-\frac{n}{N}\right)\frac{1}{n}\frac{1}{\mu_{k,Y}^2}\frac{\sum_{s=1}^{N}\left(X_s - RY_s\right)^2}{N-1} = N^2\left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_{s=1}^{N}\left(X_s - RY_s\right)^2}{N-1}, \tag{9}$$

kde $D\left(\hat{R}\right) = \left(1-\frac{n}{N}\right)\frac{1}{n}\frac{1}{\mu_{k,Y}^2}\frac{\sum_{s=1}^{N}\left(X_s - RY_s\right)^2}{N-1}$.

Na základe údajov z výberového súboru

$$\hat{D}\left(\hat{UH}_X\right) \cong N^2\left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_{i=1}^{n}\left(X_i - \hat{R}Y_i\right)^2}{n-1}. \tag{10}$$

Rozptyl odhadu strednej hodnoty v konečnom základnom súbore je N^2 krát menší, t.j.

$$D\left(\hat{\mu}_{k,X}\right) \cong \mu_{k,Y}^2\left(1-\frac{n}{N}\right)\frac{1}{n}\frac{1}{\mu_{k,Y}^2}\frac{\sum_{s=1}^{N}\left(X_s - RY_s\right)^2}{N-1} = \left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_{s=1}^{N}\left(X_s - RY_s\right)^2}{N-1} \tag{11}$$

Jeho bodový odhad je

$$\hat{D}\left(\hat{\mu}_{k,X}\right) \cong \left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_{i=1}^{n}\left(X_i - \hat{R}Y_i\right)^2}{n-1} \tag{12}$$

Vzorce (8),(9),(11) by zostali bez zmeny aj v prípade, keď by sme chceli zistiť stredné kvadratické chyby *MSE* (•).

Rozložením celkovej sumy štvorcov v (9), (10) dostaneme vzorce rozptylov pomerových odhadov úhrnu, strednej hodnoty konečného základného súboru:

$$D\left(U\hat{H}_X\right) = N^2\left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_{s=1}^{N}\left[\left(X_s-\mu_{k,X}\right)-R\left(Y_s-\mu_{k,Y}\right)\right]^2}{N-1} =$$

$$= N^2\left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_{s=1}^{N}\left[\left(X_s-\mu_{k,X}\right)^2-R^2\left(Y_s-\mu_{k,Y}\right)^2-2\left(X_s-\mu_{k,X}\right)\left(Y_s-\mu_{k,Y}\right)\right]^2}{N-1} = \qquad (13)$$

$$= N^2\left(1-\frac{n}{N}\right)\frac{1}{n}\left(\sigma_{k,X}^{)2}+R^2\sigma_{k,Y}^{)2}-2R\rho_{XY}^{)}\sigma_{k,X}^{)}\sigma_{k,Y}^{)}\right)$$

$$D\left(\hat{\mu}_{k,X}\right) = \left(1-\frac{n}{N}\right)\frac{1}{n}\left(\sigma_{k,X}^{)2}+R^2\sigma_{k,Y}^{)2}-2R\rho_{XY}^{)}\sigma_{k,X}^{)}\sigma_{k,Y}^{)}\right), \qquad (14)$$

kde

$$\sigma_{k,X}^{)2} = \frac{1}{N-1}\sum_{s=1}^{N}\left(X_s-\mu_{k,X}\right)^2$$

$$\sigma_{k,Y}^{)2} = \frac{1}{N-1}\sum_{s=1}^{N}\left(Y_s-\mu_{k,Y}\right)^2$$

$$C_{XY}^{)} = \frac{1}{N-1}\sum_{s=1}^{N}\left(X_s-\mu_{k,X}\right)\left(Y_s-\mu_{k,Y}\right)$$

$$\rho^{)}_{XY} = \frac{C_{XY}^{)}}{\sigma_{k,X}^{)}\sigma_{k,Y}^{)}}.$$

Na základe údajov z výberového súboru

$$\hat{D}\left(U\hat{H}_X\right) = N^2\left(1-\frac{n}{N}\right)\frac{1}{n}\left(s_X^2+\hat{R}^2s_Y^2-2\hat{R}s_{XY}\right), \qquad (15)$$

$$\hat{D}\left(\hat{\mu}_{k,X}\right) = \left(1-\frac{n}{N}\right)\frac{1}{n}\left(s_X^2+\hat{R}^2s_Y^2-2\hat{R}s_{XY}\right). \qquad (16)$$

2.3 Regresné odhady úhrnu a strednej hodnoty premennej *X* v základnom súbore

Všimneme si lineárne regresné odhady. Podobne ako pomerové odhady, sú určené sa zväčšenie presnosti využitím doplnkovej premennej *Y*, ktorá je korelovaná so študovanou premennou *X*.

Keď sa skúma závislosť medzi *X* a *Y* a je približne lineárna, ale priamka neprechádza začiatkom súradnicovej sústavy, vtedy je lepšie použiť namiesto pomerového odhadu, regresný odhad (bodový odhad založený na lineárnej regresii).

Predpokladajme, že poznáme hodnoty premenných *X, Y* všetkých štatistických jednotiek vo výbere a že je známa stredná hodnota $\mu_{k,Y}$ pomocnej premennej *Y* konečného základného súboru . Potom lineárny regresný odhad strednej hodnoty $\mu_{k,X}$ študovanej premennej *X* konečného základného súboru vypočítame podľa vzťahu:

$$\hat{\mu}_{k,X} = \overline{X} + \hat{B}\left(\mu_{k,Y} - \overline{Y}\right) \tag{17}$$

Lineárny regresný odhad úhrnu UH_X v základnom súbore premennej *X* dostaneme vynásobením odhadu (15) hodnotou *N:*

$$\hat{UH}_X = N\overline{X} + \hat{B}\left(UH_Y - N\overline{Y}\right) \tag{18}$$

Regresné odhady (17) a (18) sú odhady *konzistentné*, keď je veličina $\hat{B}$ dopredu stanovená konštanta, alebo štatistika vypočítaná z rovnakého výberu ako $\overline{X}$ a $\overline{Y}$. *Vychýlenosť a výdatnosť,* závisí od druhu veličiny $\hat{B}$. Najlepšou konštantou $\hat{B}$ je regresný koeficient *B* v základnom súbore. Regresný odhad je v tomto prípade nevychýlený a z toho vyplýva, že stredná kvadratická chyba regresného odhadu je totožná s rozptylom regresného odhadu.

Veľkosť rozptylu regresného odhadu $\hat{\mu}_{k,X} = \overline{X} + \hat{B}\left(\mu_{k,Y} - \overline{Y}\right)$ pre $\hat{B} = B$ vypočítame podľa vzťahu

$$\begin{aligned}
&D_{min}\left(\hat{\mu}_{k,X}\right) = D\left[\overline{X} + B\left(\mu_{k,Y} - \overline{Y}\right)\right] = D\left[\overline{X} - B\left(\overline{Y} - \mu_{k,Y}\right)\right] = \\
&= D\left(\overline{X}\right) + \hat{B}^2 \cdot D\left(\overline{Y}\right) - 2BC\left(\overline{X};\overline{Y}\right) = \\
&= \left(1 - \frac{n}{N}\right)\frac{1}{n}\left(\sigma_{k,X}^{)2} + B^2\sigma_{k,Y}^{)2} - 2BC_{XY}^{)}\right) = \left(1 - \frac{n}{N}\right)\frac{1}{n}\left(\sigma_{k,X}^{)2} + \sigma_{k,X}^{)2}\rho_{XY}^{)2} - 2\sigma_{k,X}^{)2}\rho_{XY}^{)2}\right) = \\
&= \left(1 - \frac{n}{N}\right)\frac{1}{n}\sigma_{k,X}^{)2}\left(1 - \rho_{XY}^{)2}\right),
\end{aligned} \tag{19}$$

kde ρ_{XY} je koeficient korelácie a $C_{XY}^{)}$ je kovariancia v základnom súbore

($B = \rho_{XY}^{)}\frac{\sigma_{k,X}^{)}}{\sigma_{k,Y}^{)}}$, $C_{XY}^{)} = \rho_{XY}^{)}\sigma_{k,X}^{)}\sigma_{k,Y}^{)}$).

Výsledok (19) názorne ukazuje, že regresný odhad $\hat{\mu}_{k,X} = \overline{X} + B\left(\mu_{k,Y} - \overline{Y}\right)$ pri známom *B* by nebol nikdy horší (menej výdatný) ako jednoduchý odhad pomocou $\overline{X}$, pretože rozptyl (19) je súčin rozptylu $D\left(\overline{X}\right)$ a dvojčlena, ktorý nie je nikdy väčší ako 1. Zhoda výdatnosti oboch odhadov môže nastať iba pri ρ_{XY} =0, t.j. pri nekorelovanosti pomocného znaku *Y* so študovaným znakom *X*. Ďalej vidíme, že naopak, čím je táto

korelácia silnejšia (či už je kladná alebo záporná), tým je výdatnejší regresný odhad v porovnaní s jednoduchým odhadom.

Veľkosť rozptylu regresného odhadu UH_X pre $\hat{B} = B$ vypočítame podľa vzťahu

$$D_{min}\left(U\hat{H}_X\right) = N^2\left(1-\frac{n}{N}\right)\frac{1}{n}\sigma_{k,X}^2\left(1-\rho_{XY}^2\right) \tag{20}$$

Keď označíme reziduálne odchýlky od základnej regresnej priamky $X_s - \mu_{k,X} + B\left(Y_s - \mu_{k,Y}\right)$ ako E_s (vo výbere ako $E_i = X_i - \overline{X} + B\left(Y_i - \overline{Y}\right)$) a ich rozptyl ako $\sigma_E^{)2}$ môžeme vzorec (19) napísať v tvare $\left(1-\frac{1}{N}\right)\frac{1}{n}\sigma_E^{)2}$. Jeho **nevychýlený bodový odhad** je

$$\hat{D}\left(\hat{\mu}_{k,X}\right) = \left(1-\frac{n}{N}\right)\frac{1}{n}S_E^2 = \left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_i^n\left[X_i - \overline{X} - B\left(Y_i - \overline{Y}\right)\right]^2}{n-1} \tag{21}$$

Bodový odhad rozptylu UH_X je

$$\hat{D}\left(U\hat{H}_X\right) = N^2\left(1-\frac{n}{N}\right)\frac{1}{n}S_E^2 = N^2\left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_i^n\left[X_i - \overline{X} - B\left(Y_i - \overline{Y}\right)\right]^2}{n-1} \tag{22}$$

Uvažujme teraz prípad, kedy $\hat{B} \neq B$. **Veľkosť rozptylu regresného odhadu** $\mu_{k,X}$ v tomto prípade je

$$D\left(\hat{\mu}_{k,X}\right) = D\left[\overline{X} - \hat{B}\left(\overline{Y} - \mu_{k,Y}\right)\right] = D\left(\overline{X}\right) + \hat{B}^2 D\left(\overline{Y}\right) - 2\hat{B}C\left(\overline{X};\overline{Y}\right) =$$

$$\left(1-\frac{n}{N}\right)\frac{1}{n}\left(\sigma_{k,Y}^{)2} + \hat{B}^2\sigma_{k,X}^{)2} - 2\hat{B}C_{XY}^{)}\right) = \left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_s^N\left[X_s - \mu_{k,X} + \hat{B}\left(Y_s - \mu_{k,Y}\right)\right]^2}{N-1}. \tag{23}$$

Jeho bodový odhad je

$$\hat{D}\left(\hat{\mu}_{k,X}\right) = \left(1-\frac{n}{N}\right)\frac{1}{n}S_E^2 = \left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_i^n\left[X_i - \overline{X} + \hat{B}\left(Y_i - \overline{Y}\right)\right]^2}{n-1}. \tag{24}$$

Bodový odhad rozptylu $U\hat{H}_X$ je

$$\hat{D}\left(U\hat{H}_X\right) = N^2\left(1-\frac{n}{N}\right)\frac{1}{n}S_E^2 = N^2\left(1-\frac{n}{N}\right)\frac{1}{n}\frac{\sum_i^n\left[X_i - \overline{X} + \hat{B}\left(Y_i - \overline{Y}\right)\right]^2}{n-1}. \tag{25}$$

Keď sa skúma závislosť medzi X a Y a je približne lineárna, priamka prechádza začiatkom súradnicovej sústavy (hodnota lokujúcej konštanty regresnej priamky sa rovná nule) a podmiemený reziduálny rozptyl je úmerný hodnotám Y, vtedy je vhodné použiť $\hat{B} = \hat{R} = \overline{X} / \overline{Y}$, regresný odhad sa zmení na pomerový odhad. *Pomerové odhady sú v teórii regresných odhadov špeciálnym prípadom širšej triedy regresných odhadov*.

3. Záver

Predpokladajme, že poznáme hodnoty premenných X, Y pre všetky štatistické jednotky vo výbere a že stredná hodnota $\mu_{k,Y}$ pomocnej premennej Y konečného základného súboru je známa ($\mu_{k,Y} = 4{,}6$). Pre jednoduchosť budeme vychádzať zo základného súboru s piatimi prvkami, z ktorých dvojica premenných X, Y nadobúda hodnoty uvedené v nasledujúcej tabuľke:

Tabuľka 4 Základný súbor

Prvok	x_s	y_s
A	1	*0*
B	4	*2*
C	7	*6*
D	9	*6*
E	11	*9*
Spolu	*32*	*23*

Jednoduchým náhodným výberom bez opakovania o rozsahu troch prvkov dostaneme výbery a vypočítame nasledujúce výberové charakteristiky :

a) regresný bodový odhad strednej hodnoty $\mu_{k,X}$ študovanej premennej X v konečnom základnom súbore:

$$\hat{\mu}_{k,X} = \overline{X} + \hat{B}\left(\mu_{k,Y} - \overline{Y}\right),$$

kde $\hat{B} = \mathrm{B}$ (najlepšie výsledky sa dosiahnu, keď sa použije ako konštanta B, regresný koeficient základného súboru , v našom prípade B = 1,08984),

b) jednoduchý bodový odhad strednej hodnoty $\mu_{k,X}$ (výberový priemer) premennej X v základnom súbore :

$$\hat{\mu}_{k,X} = \overline{X}$$ a

c) pomerový odhad strednej hodnoty $\mu_{k,X}$ študovanej premennej X v konečnom základnom súbore:

$$\hat{\mu}_{k,X} = \hat{R} \cdot \mu_{k,Y} = \frac{\overline{X}}{\overline{Y}} \cdot \mu_{k,Y} .$$

Na porovnanie bodových odhadov (b), (c), (d), (e), v tab. 5 vypočítame strednú hodnota $E(\hat{\mu}_{k,X})$ a strednú kvadratickú chybu $MSE(\hat{\mu}_{k,X}) = E(\hat{\mu}_{k,X} - \mu_{k,X})^2$ jednotlivých odhadov.

Z tab. 5 zistíme, že nevychýlený a zároveň najvýdatnejší odhad strednej hodnoty $\mu_{k,X}$ študovanej premennej *X* v základnom súbore dosiahneme regresným odhadom.

Tabuľka 5 Výpočtová tabuľka

Výbery	(a)	(b)	(c)
	$\overline{x} + B(\mu_{k,Y} - \overline{y})$	$\overline{x}$	$\frac{\overline{x}}{\overline{y}} \mu_{k,Y}$
ABC	6,10	4,00	6,89
ABD	6,77	4,67	8,05
ABE	6,34	5,33	6,68
ACD	6,32	5,67	6,52
ACE	5,89	6,33	5,82
ADE	6,56	7,00	6,44
BCD	6,59	6,67	6,57
BCE	6,16	7,33	5,94
BDE	6,83	8,00	6,49
CDE	6,38	9,00	5,91
***E*(●)**	**6,4**	**6,4**	**6,5**
***MSE*(●)**	**0,08**	**2,11**	**0,39**

Poďakovanie

Tento príspevok vznikol s prispením grantovej agentúry VEGA v rámci projektu číslo 1/3182/06 Zlepšovanie kvality produkcie strojárskych výrobkov pomocou štatistických metód.

Podrobnosti o jednoduchých, pomerových a regresných odhadoch strednej hodnoty a úhrnu v základnom súbore možno nájsť napríklad v [1]- [4].

Literatúra

[1] COCHRAN, W. G.: *Sampling Techniques*. New York: Wiley and Sons, 1977.

[2] ČERMÁK, V.- VRABEC, M.: *Teorie výběrových šetření, Část 2*. Praha: VŠE, 1998.

[3] GRAIS B.: *Méthodes statistiques. Techniques statistiques 2*. Paris: Dunod, 2000.

[4] JANIGA, I., CISKO, P. : *O rozsahoch výberu pre intervaly spoľahlivosti strednej hodnoty normálneho rozdelenia*. In APLIMAT 2006: 5th International Conference: Proceedings. Bratislava:KM SjF STU, 2006. ISBN 80-967305-5-X, s. 541- 544.

Adresy autorov: Doc. Ing. Ľubica Hrnčiarová, PhD., Katedra štatistiky FHI EU v Bratislave, Dolnozemská 1, 852 35 Bratislava, hrnciaro@dec.euba.sk
Doc. Ing. Mian Terek, PhD., Katedra štatistiky FHI EU v Bratislave, Dolnozemská 1, 852 35 Bratislava, terek@dec.euba.sk

Približné riešenie parciálnej diferenciálnej rovnice vedenia tepla

Hrubina Kamil, Vagaská Alena

Fakulta výrobných technológií

Technická univerzita v Košiciach

Bayerova 1

080 01 Prešov

paris.kh@azet.sk, vagaska.alena@fvt.sk

Abstrakt

V článku je skúmaná problematika riešenia parciálnej diferenciálnej rovnice vedenia tepla pri definovaných začiatočných a okrajových podmienkach, ktorá predstavuje matematický model procesu, alebo riadeného systému. Hľadané približné riešenie parciálnej diferenciálnej rovnice použitím diskretizácie premenných x, y vedie na sústavu diferenciálnych rovníc s konštantnými koeficientmi. Keďže v presnom riešení sústavy diferenciálnych rovníc vystupuje matica $e^{A\mu}$ a integrál z tejto matice, je skúmaný postup aproximácie pre realizáciu efektívneho algoritmu na výpočet týchto matíc.

Kľúčové slová: diferenciálne rovnice, aproximácia, maticové funkcie

1 Úvod

V článku budeme skúmať problematiku riešenia matematického modelu procesu, resp. riadeného systému, ktorý je opísaný parciálnou diferenciálnou rovnicou vedenia tepla v dvoch dimenziách so začiatočnými a okrajovými podmienkami, s využitím aproximácie.

Samotná rovnica vedenia tepla patrí medzi najdôležitejšie rovnice v technických disciplínach (v elektrotechnike, strojárenstve, v hutníctve a v iných). Metódy riešenia rovnice je možné zhruba rozdeliť na tri skupiny:

a) Metódy analytické (tieto z hľadiska riešenia všeobecných úloh, ktoré sa v praxi vyskytujú, neprichádzajú do úvahy).
b) Metódy analógové (najvhodnejšie sa javia metódy sietí, ktoré úzko súvisia s numerickými metódami).
c) Metódy numerické, ktoré v súčasnosti prekonávajú najväčší vývoj a to paralelne s vývojom počítačov. Z najpoužívanejších uveďme metódu sietí, diferenčnú metódu – metódu kolokácie a aproximácie, metódu konečných prvkov (Legras, 1978; Marčuk, 1987; Hrubina, 1993).

Cieľom článku je nájsť riešenie sústavy diferenciálnych rovníc s konštantnými koeficientami. Túto sústavu získame transformáciou parciálnej diferenciálnej rovnice vedenia tepla. Numerické riešenie sústavy diferenciálnych rovníc možeme získať aplikovaním aproximácie, ktorá umožňuje efektívny výpočet matice $e^{A\mu}$, ako aj integrálu z tejto matice.

2 Transformácia diskretizáciou premenných x, y parciálnej diferenciálnej rovnice vedenia tepla

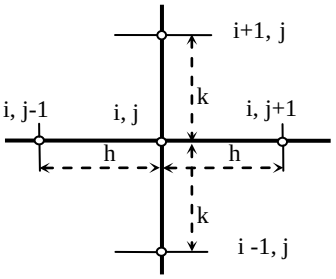
Hľadajme riešenie $\boldsymbol{u}(x, y, t)$ parciálnej diferenciálnej rovnice vedenia tepla

$$\frac{\partial^2 \boldsymbol{u}}{\partial x^2} + \frac{\partial^2 \boldsymbol{u}}{\partial y^2} = K\frac{\partial \boldsymbol{u}}{\partial t} \tag{1}$$

ktorá je definovaná v oblasti D Euklidovského priestoru E_2, a nech Γ je hranica oblasti D. Pre hľadané riešenie sú dané podmienky, ktoré poznáme:

1) Začiatočné hodnoty funkcie $\boldsymbol{u}(x, y, 0)$ v pravouhlej oblasti D
2) Hodnoty okrajových hodnôt funkcie $\boldsymbol{u}(x_i, y_i, t)$ na hranici Γ oblasti D.

Približné riešenie (1) vyjadríme, ak urobíme diskretizáciu premenných *x*, *y* v riešenom probléme tak, že pokryjeme oblasť D pravouhlou sieťou. Keďže D je oblasť v rovine R_{xy}, v tejto rovine veďme dve sústavy priamok rovnobežných so súradnicovými osami: $x = ih,\ y = jk,$ kde h a k sú ľubovoľné kladné čísla a $i, j = 0, \pm 1, \pm 2, \ldots$. Tieto priamky rozdeľujú celú rovinu na obdĺžniky so stranami *h* a *k*. Vrcholy týchto obdĺžnikov sú uzlami pravouhlej siete, obr. 1.



Obr. 1

Uzly označme $M_{ij}(x_i, y_j)$. Pre ďalší postup hľadania približného riešenia (1) definujeme funkcie

$$\varphi_{ij}(t) = \boldsymbol{u}(x_i, y_i, t)$$

a nahraďme spojitý operátor $\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ diskrétnym operátorom tým, že pre uvažovaný uzol berieme do úvahy hodnoty funkcie $\boldsymbol{u}(x, y, t)$ v susedných uzloch.

Na základe uvedených predpokladov môžeme vyjadriť aproximáciu spojitého operátora

$$\frac{\partial^2 \boldsymbol{u}}{\partial x^2} + \frac{\partial^2 \boldsymbol{u}}{\partial y^2} = \frac{\varphi_{i,\,j-1} + \varphi_{i,\,j+1} - 2\varphi_{ij}}{h^2} + \frac{\varphi_{i-1,\,j} + \varphi_{i+1,\,j} - 2\varphi_{ij}}{k^2}$$

Ďalej môžeme vyjadriť pre každý uzol vo vnútri siete oblasti D aproximáciu pravej strany rovnice (1):

$$K\frac{d\varphi_{i\,j}(t)}{dt} = \frac{\varphi_{i,\,j-1} + \varphi_{i,\,j+1} - 2\varphi_{ij}}{h^2} + \frac{\varphi_{i-1,\,j} + \varphi_{i+1,\,j} - 2\varphi_{ij}}{k^2} \tag{2}$$

Teda, použitím diskretizácie premenných x, y v rovnici získame sústavu diferenciálnych rovníc s konštantnými koeficientami, ktorú môžeme vyjadriť v maticovom tvare

$$\frac{d\boldsymbol{\varphi}(t)}{dt} = \boldsymbol{A}\boldsymbol{\varphi}(t) + \boldsymbol{f}(t); \quad \boldsymbol{\varphi}(0) = \boldsymbol{g} \quad \text{dané} \tag{3}$$

Matica $\boldsymbol{A}$ závisí od spôsobu diskretizácie, matica $\boldsymbol{f}(t)$ závisí od daných hodnôt na hranici. Odchýlka medzi riešením $\boldsymbol{u}(x, y, t)$ spojitého problému (1) a presným riešením $\boldsymbol{\varphi}(t)$ sústavy diferenciálnych rovníc (3), zrejme pri splnení začiatočných a okrajových podmienok pre obidve úlohy, je odchýlkou závislou od použitej diskretizácie (definovanej siete). Pre riešenie rovnice (1) môžeme zvoliť parameter K nezávisle od x a y, pravouhlú oblasť D a aproximáciu spojitého operátora podľa vzťahu (2). V tomto prípade je možné vyjadriť presné riešenie sústavy diferenciálnych rovníc (3) a uviesť do súvislosti odchýlku medzi presným a približným riešením, (Legras, 1978; Hrubina, 1992, 1993).

3 Numerické riešenie sústavy diferenciálnych rovníc s konštantnými koeficientami

Budeme skúmať numerické riešenie sústavy diferenciálních rovníc (3)

$$\frac{d\boldsymbol{\varphi}(t)}{dt} = \boldsymbol{A}\boldsymbol{\varphi}(t) + \boldsymbol{f}(t), \text{ kde } \frac{d\boldsymbol{\varphi}}{dt}, \boldsymbol{\varphi} \text{ a } \boldsymbol{f}(t) \text{ sú stĺpcové matice} \tag{3}$$

$$\frac{d\boldsymbol{\varphi}}{dt} = \begin{pmatrix} \frac{d\varphi_1}{dt} \\ \vdots \\ \frac{d\varphi_n}{dt} \end{pmatrix}, \qquad \boldsymbol{\varphi} = \begin{pmatrix} \varphi_1 \\ \vdots \\ \varphi_n \end{pmatrix} \quad \text{a} \quad \boldsymbol{f}(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_n(t) \end{pmatrix}$$

$\boldsymbol{A}$ je švorcová matica *n*-tého stupňa, ktorej prvky sú nezávisle od t. Funkcie $f_i(t)$ sú dané funkcie, ktorých hodnoty poznáme pre rôzne hodnoty t. Ďalej predpokladáme, že poznáme začiatočné hodnoty funkcií $\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)$. Našim cieľom je určiť numerickou metódou riešenie $\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)$ sústavy diferenciálnych rovníc (3).

Vieme, že presné riešenie sformulovaného problému so začiatočnými hodnotami je definované v maticovom vyjadrení:

$$\boldsymbol{\varphi}(t) = e^{\boldsymbol{A}t}\boldsymbol{\varphi}(0) + \int_0^t e^{\boldsymbol{A}(t-\tau)} \cdot \boldsymbol{f}(\tau)\,d\tau \tag{4}$$

Toto presné riešenie problému (3) môžeme použiť v metóde krok po kroku, ktorá nám umožní výpočet funkcií φ_i v čase $(t + \mu)$ pričom použijeme hodnoty funkcií $\varphi_i(t)$; μ je krok. Teda presné riešenie bude mať tvar:

$$\boldsymbol{\varphi}(t + \mu) = e^{\boldsymbol{A}\mu}\boldsymbol{\varphi}(t) + \int_0^{\mu} e^{\boldsymbol{A}(\mu - s)} \cdot \boldsymbol{f}(t + s)\,ds \tag{5}$$

Pre použitie relácie (5) v sieťovej oblasti je potrebné vytvoriť metódy, ktoré umožňujú výpočet matice $e^{\boldsymbol{A}\mu}$, ako aj integrál z matíc:

$$\int_0^{\mu} e^{\boldsymbol{A}(\mu-s)} \cdot \boldsymbol{f}(t+s)\,ds \tag{6}$$

3.1 Aproximácia matice $e^{\boldsymbol{A}\mu}$ a maticových funkcií

Budeme vychádzať zo základného teorému. Takže budeme uvažovať maticu $\boldsymbol{A}$, ktorá je reálna a môžeme ju vyjadriť v diagonálnom tvare. Postačujúcou podmienkou je, aby matica $\boldsymbol{A}$ bola symetrická alebo aby jej vlastné čísla boli rôzne. Teda maticu $\boldsymbol{A}$ môžeme vyjadriť v tvare $\boldsymbol{A}=\boldsymbol{V}\,\mathrm{diag}(\lambda_i)\cdot\boldsymbol{V}^{-1}$, kde $\mathrm{diag}(\lambda_i)$ je diagonálna matica, ktorej prvky sú vlastné čísla $\lambda_1,\lambda_2,...,\lambda_n$ matice $\boldsymbol{A}$; matica $\boldsymbol{V}$, ktorej stĺpce obsahujú zložky vlastných vektorov matice $\boldsymbol{A}$; $\boldsymbol{V_1}, \boldsymbol{V_2}, \ldots, \boldsymbol{V_n}$ sú vlastné vektory odpovedajúce vlastným číslam $\lambda_1,\lambda_2,...,\lambda_n$.

Do úvah zaveďme polynóm tvaru

$$p(x)=a_0+a_1x+...+a_nx^n \tag{7}$$

a priradený maticový polynóm $P(\boldsymbol{A})=a_0\boldsymbol{I}+a_1\boldsymbol{A}+a_2\boldsymbol{A}^2+....+a_n\boldsymbol{A}^n$. Jednoduchým výpočtom zistíme, že $\boldsymbol{A}^2=\boldsymbol{V}\,diag\left(\lambda_i^2\right)\boldsymbol{V}^{-1}$, **...** , $\boldsymbol{A}^q=\boldsymbol{V}\,diag\left(\lambda_i^q\right)\boldsymbol{V}^{-1}$; potom

$a_q\mu^q\boldsymbol{A}^q=\boldsymbol{V}\,diag\left(a_q\mu^q\lambda_i^q\right)\boldsymbol{V}^{-1}$ a napokon

$$P(\boldsymbol{A}\mu)=\boldsymbol{V}\,diag(p(\mu\lambda_i))\,\boldsymbol{V}^{-1} \tag{8}$$

Ak zavedieme maticovú funkciu $G(\boldsymbol{A})$ definovanú rozvojom do nekonečného radu $G(\boldsymbol{A})=g_0\boldsymbol{I}+g_1\boldsymbol{A}+...+g_q\boldsymbol{A}^q+...$ a priradíme mu funkciu

$g(x)=g_0+g_1x+...+g_qx^q+...$ rovnako dostaneme

$$G(\boldsymbol{A}\mu)=\boldsymbol{V}\,diag\left(g(\mu\lambda_i)\right).\boldsymbol{V}^{-1} \tag{9}$$

Vzťah (9) je platný za podmienky, že n hodnôt $\mu\lambda_i$ patrí do oblasti konvergencie postupnosti $g(x)$. Môžeme teda vyjadriť:

$$G(\boldsymbol{A}\mu)-P(\boldsymbol{A}\mu)=\boldsymbol{V}\,diag\left(g(\mu\lambda_i)-p(\mu\lambda_i)\right)\boldsymbol{V}^{-1} \tag{10}$$

Zaveďme ešte maticu odchýliek $\boldsymbol{\eta}$, ktorú definujeme vzťahom **Error! Bookmark not defined.Error! Bookmark not defined.** $\boldsymbol{\eta}=\left(G(\boldsymbol{A}\mu)-P(\boldsymbol{A}\mu)\right)$,

pričom položme $\varepsilon_i=g(\mu\lambda_i)-p(\mu\lambda_i)$, takže dostaneme vzťah:

$$\boldsymbol{\eta}=\boldsymbol{V}\,diag(\varepsilon_1,\ldots,\varepsilon_n)\boldsymbol{V}^{-1} \tag{11}$$

Zo vzťahu (11) vyplýva, že vlastné vektory matice $\boldsymbol{A}$ sú vlastnými vektormi aj matice $\boldsymbol{\eta}$ a odpovedajúce vlastné čísla ε_i, odkiaľ môžeme napísať $\boldsymbol{\eta V_i}=\varepsilon_i\boldsymbol{V_i}$.

Predpokladajme, že vlastné vektory $\boldsymbol{V_i}$ tvoria bázu n vektorov lineárne nezávislých, potom ľubovoľný vektor $\boldsymbol{W}$ môžeme vyjadriť $\boldsymbol{W}=c_1\boldsymbol{V_1}+c_2\boldsymbol{V_2}+...+c_n\boldsymbol{V_n}$. (12)

Označme $\boldsymbol{W_1}$ vektor, ktorého zložky sú definované vzťahom $\boldsymbol{W_1}=\boldsymbol{\eta W}$, potom relácia (12) umožňuje napísať:

$$\boldsymbol{W_1} = c_1 \varepsilon_1 \boldsymbol{V_1} + c_2 \varepsilon_2 \boldsymbol{V_2} + ... + c_n \varepsilon_n \boldsymbol{V_n} \tag{13}$$

Predpokladajme, že sme našli taký polynóm $p(x)$, že všetky ε_i v module budú menšie než dané ε. Vzťah (13) vyjadruje, že všetky súradnice c_i ľubovoľného vektora $\boldsymbol{W}$ sú násobené číslami, ktoré sú menšie než ε.

Nech matica $\boldsymbol{A}$ je symetrická. Vlastné vektory matice $\boldsymbol{A}$ tvoria bázu z n vektorov, ktorých zložky sú reálne, vektory sú lineárne nezávislé, ortogonálne a môžeme ich voliť normované. Tieto vlastnosti umožňujú zaujímavú interpretáciu vzťahu (13).

Uvažujme normu vektora $\boldsymbol{W}$, $\|\boldsymbol{W}\| = \left(\boldsymbol{W}^2\right)^{\frac{1}{2}} = \left(c_1^2 + c_2^2 + ... + c_n^2\right)^{\frac{1}{2}}$, rovnako

$\|\boldsymbol{\eta W}\| = \left(c_1^2 \varepsilon_1^2 + c_2^2 \varepsilon_2^2 + ... + c_n^2 \varepsilon_n^2\right)^{\frac{1}{2}} \le \varepsilon \left(c_1^2 + c_2^2 + ... + c_n^2\right)^{\frac{1}{2}}$ odkiaľ

$$\|\boldsymbol{\eta W}\| \le \varepsilon \|\boldsymbol{W}\| \tag{14}$$

Modul vektora odchýliek je ε krát menší než modul začiatočného vektora. Stručne vyjadrené, ak vieme nájsť polynóm $p(x)$, ktorý aproximuje funkciu $g(x)$ pre n hodnôt $\mu\lambda_i$ tak, že $|g(\mu\lambda_i) - p(\mu\lambda_i)| < \varepsilon$ pre $i = 1, 2, ..., n$, matica $P(\boldsymbol{A}\mu)$ je aproximáciou matice $G(\boldsymbol{A}\mu)$, potom chyba je charakterizovaná jednou z relácií (16) (vo všeobecnom prípade), alebo vzťahom (14), keď $\boldsymbol{A}$ je symetrická matica.

3.2 Aproximácia matice $e^{\boldsymbol{A}\mu}$ polynómom $p(x)$

Prípad symetrickej matice $\boldsymbol{A}$, ktorej vlastné čísla sú reálne a záporné. Nájsť polynóm, ktorý rovnomerne aproximuje funkciu e^x na intervale $(-\infty, 0)$ nie je možné, (Legras, 1978).

Pre aproximáciu e^x je nutná:

1) Existencia mojoranty $\boldsymbol{\Lambda}$ pre vlastné čísla; a to podľa teorému Hadamara, ktorý obsahuje tvrdenie, že pre $i = 1, 2, ..., n$ máme podmienku: $-\boldsymbol{\Lambda} \le \lambda_i < 0$
2) Voliť čísla α, ε a polynóm $p(x)$ vhodného stupňa tak, aby bola splnená podmienka: $|e^x - p(x)| < \varepsilon$, ak $-\alpha \le x \le 0$

Pre aproximáciu zvolíme krok $\mu \le \mu_0 = \frac{\alpha}{\boldsymbol{\Lambda}}$. Z tejto podmienky vyplýva:

$-\alpha \le \mu\lambda_i < 0$, takže $|e^{\mu\lambda_i} - p(\mu\lambda_i)| < \varepsilon$ pre $i = 1, 2, ..., n$, μ_0 je konštanta, ktorá je hornou limitou kroku pre ktorý vzťah (5) môže byť ešte použitý, ak približne nahradíme maticu $e^{\boldsymbol{A}\mu}$ maticovým polynómom $P(\boldsymbol{A}\mu)$, ($P(\boldsymbol{A}\mu)$ je maticový polynóm priradený k polynómu $p(x)$), (Legras, 1978).

4 Aproximácia sústav diferenciálnych rovníc

Riešenie parciálnych diferenciálnych rovníc nás často vedie k hľadaniu riešenia sústavy diferenciálnych rovníc u ktorých je počet neznámych funkcií značne veľký. V tomto prípade je obtiažné vyjadriť maticu $\boldsymbol{A}$ ako aj maticový polynóm. Musíme hľadať algoritmus, ktorým získame výpočet vektora $\boldsymbol{\psi}$, ktorý je definovaný maticovým vyjadrením $\boldsymbol{\psi} = \boldsymbol{A}.\boldsymbol{\varphi}$, čo môžeme ešte vyjadriť takto:

$$\boldsymbol{\psi} = L(\boldsymbol{\varphi}) \tag{15}$$

Zo vzťahu (15) vyplýva, že priamo nevyjadrujeme maticu $\boldsymbol{A}$.

Podobne nevyjadrujeme maticový polynóm $P(\boldsymbol{A}\mu)$, avšak môžeme vypočítať $P(\boldsymbol{A}\mu).\varphi$ následovne: $a_0, a_1, \ldots, a_p$ sú koeficienty polynómu $p(x)$ priradeného k maticovému polynómu $P(\boldsymbol{A})$; vytvoríme postupnosť: $\boldsymbol{u_0} = a_p\boldsymbol{\varphi}$, $\boldsymbol{u_1} = a_{p-1}\boldsymbol{\varphi} + \mu L(\boldsymbol{u_0})$, $\boldsymbol{u_2} = a_{p-2}\boldsymbol{\varphi} + \mu L(\boldsymbol{u_1})$, $\ldots$ $\boldsymbol{u_p} = a_0\boldsymbol{\varphi} + \mu L(\boldsymbol{u_{p-1}})$.

Jednoduchou úvahou overíme, že vektor $\boldsymbol{u_p}$ je definovaný

$$\boldsymbol{u_p} = P(\boldsymbol{A}\mu) \times \boldsymbol{\varphi} \tag{16}$$

Pre výpočet zložiek vektora $\boldsymbol{u_p}$ je možné vytvoriť pre vzťah (16) procedúru, v prípade, že poznáme koeficienty a_j. Tým vlastne môžeme vytvoriť procedúru pre výpočet matice $L(\boldsymbol{\varphi})$.

V prípade, že vlastné čísla matice $\boldsymbol{A}$ sú komplexné čísla, potom približné vyjadrenie e^z uskutočníme polynómom s premennou z definovanou v kružnici Γ s polomerom α. Ak $\boldsymbol{\Lambda}$ je majoranta v module pre vlastné čísla; a ak $p_2(z)$ približne vyjadruje e^z v kružnici Γ s odchýlkou menšou než ε, $P_2(\boldsymbol{A}\mu)$ bude ešte aproximáciou matice $e^{\boldsymbol{A}\mu}$ a to v zmysle predchádzajúcej teórie.

4.2 Numerický výpočet integrálu matice

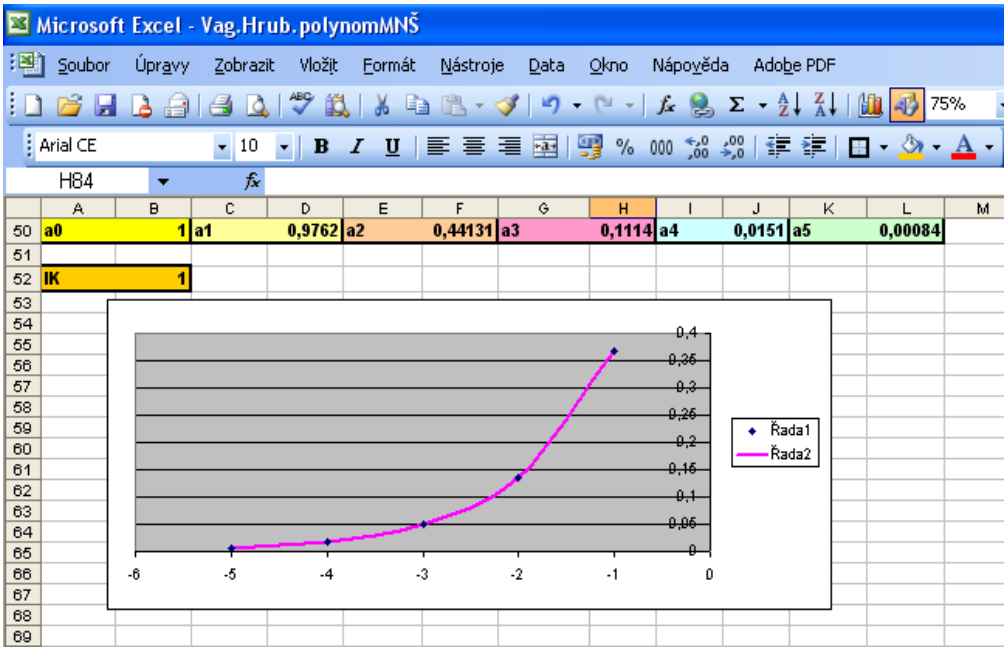
Mnohé experimenty na PC ukázali, že použitie algoritmov z klasických metód (Gauss, Cotes) aj v prípade použitia veľkého počtu uzlových bodov viedli k nedostatočným výsledkom, (Legras, 1978). Pre získanie aproximácie definovaného integrálu:

$$\int_0^{\mu} e^{\boldsymbol{A}(\mu - s)} . \boldsymbol{f}(t+s)\, ds$$

s dostatočnou presnosťou pri použití numerickej metódy, môžeme použiť algoritmus, v ktorom použijeme presné hodnoty matice $e^{\boldsymbol{A}(\mu-s)}$ a interpolujeme maticu $\boldsymbol{f}(t+s)$ polynómom s-tého stupňa. Pre túto interpoláciu môžeme vybrať z viacerých metód napr. zovšeobecnenú Cotesovu metódu, (Legras, 1978; Marčuk, 1987; Hrubina a kol. 2003).

Pre aproximáciu funkcie e^x sme vykonali niekoľko experimentov na PC s použitím programového systému MS EXCEL. Zvolili sme $\alpha = 5$, $-5 \le x \le 0$; $\varepsilon - 10^{-3}$, $n - 5$. Aplikáciou metódy najmenších štvorcov sme získali koeficienty $a_0, a_1 \ldots, a_5$ (obr. 2). Teda aproximácia funkcie e^x môže byť vyjadrená polynómom $p(x)$ piateho stupňa:

$$e^x = 0.00084\,x^5 + 0.01506\,x^4 + 0{,}11145\,x^3 + 0{,}44130\,x^2 + 0{,}97620\,x + 1$$



Obr. 2 Aproximácia funkcie e^x aplikáciou MNŠ s využitím MS Excelu

Záver

Prínosom článku je spracovanie z teoretického hľadiska problematiky aproximácie riešenia parciálnej diferenciálnej rovnice vedenia tepla v dvoch dimenziách. Rovnica vedenia tepla so začiatočnými a okrajovými podmienkami použitím diskretizácie premenných x a y v definovanej oblasti D bola transformovaná aproximáciou spojitého operátora na sústavu diferenciálnych rovníc s konštantnými koeficientami so začiatočnými hodnotami. Pre použitie presného riešenia sústavy diferenciálnych rovníc v sieťovej oblasti boli ukázané možnosti tvorby metód na výpočet matice $e^{A\mu}$ ako aj integrálu z tejto matice. Numerickou metódou krok po kroku môžeme vypočítať zložky stĺpcového vektora $\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)$, ktorý je riešením sústavy diferenciálnych rovníc.

Článok bol vypracovný v rámci riešenia vedeckého projektu VEGA č. 1/2212/05 „Research and development of mechatronic components and systems of bioservosystems on the base of pneumatic artificial muscles“; a inštitusionálnej úlohy FVT TU Košice so sídlom v Prešove č. 1/2006 „Výskum vlastnosti mechatronických dynamických systémov a možnosti ich zdokonaľovania“

Literatúra

[1] BUTKOVSKIJ, A. G. (1975). *Metody upravlenija sistemami s raspredelenymi parametrami*. Nauka, Moskva

[2] HRUBINA, K. (1992). *Optimal control problems for systems with distributed parameters and their solution using algorithms of numerical methods*. Electrical Engineering Journal r. 43., č. 6, p. 187 – 192.

[3] HRUBINA, K. (1993). *Solving optimal control problems for systems with distributed parameters by means of iterative algorithms of algebraic methods.* Kybernetika a informatika, roč. 6, č. 1, s. 48 – 64.
[4] HRUBINA, K. (1993). *Optimum control problems for systems with distributed parameters and their solution using the finite element method algorithm.* Journal of Electrical Engineering, r. 44, No 1, p. 11 – 20.
[5] HRUBINA, K. – MAJERČÁK, J. – BORŽÍKOVÁ, J. (2003). *Riešené úlohy algoritmami numerických metód s podporou počítača.* Informatech, Košice, 235 s. ISBN 80-88941-16-4
[6] HRUBINA, K. a kol. (2005) Optimal Control of Processes based on the use of Informatics Methods. Informatech. Košice, 286s. ISBN 80-88941-30-X
[7] LEGRAS, J. (1971). Méthodes et techniques de l´analyse numérique, Dunod, Paris (preklad K. Hrubina, Bratislava: Alfa 1978,., 383 s.)
[8] LIONS, J. L. (1986). *Contrôle optimal de systémes gouvernés par des équations aux derivées partielles.* Dunod Gautier-Villars, Paris
[9] MAMRILLA, D. – VAGASKÁ, A. (2002). *On the solutions of certain quasi –linear differential systems.* In: Proceedings of the 5th Scientific Conference with International Participation, Informatics and Algorithms, Prešov, September 12 – 13, 2002, FMT, Technical University, p. 120 – 123, Košice: Informatech. ISBN 80-88941-21-0
[10] MAMRILLA, D. – VAGASKÁ, A. (2004). *About cutting curve of the adjustable knife.* In: Proceedings of the 3rd International Conference, Aplimat 2004, Bratislava, February 3 – 6, The Slovak University of Technology, part II, p. 671– 674, Bratislava: FME. ISBN 80-227-1995-1
[11] MARČUK, G. I. (1987) *Metódy numerické matematiky.* Academia Praha, 526 s.
[12] MICHLIN, S. G. a kol. (1973). *Približné metódy riešenia diferenciálnych a integrálnych rovníc.* ALFA Bratislava, SNTL Praha, 274 s.

Normalizácia metód štatistického riadenia kvality

Ivan Janiga

Abstrakt.
V príspevku uvádzame najdôležitejšie informácie o medzinárodnej technickej komisii ISO/TC 69 a národnej komisii TK 71 a o ich produktoch. Uvádzame prehľad štatistických noriem publikovaných v ISO/TC 69 a noriem prebratých do sústavy STN. Prezentujeme aj rozpracované štatistické normy ISO.

Kľúčové slová: *štatistické riadenie kvality, technická norma, normalizácia štatistických metód*

1 Úvod

Riadenie kvality je podľa STN EN ISO 9000 časť manažérstva kvality zameraná na plnenie požiadaviek na kvalitu. Štatistické riadenie kvality je jedna z častí patriacich riadeniu kvality. Metódy používané v štatistickom riadení kvality pokrývajú prakticky celú oblasť aplikovanej štatistiky a neustále sa rozvíjajú. Organizáciám všetkých typov a veľkostí sa najnovšie vedecké výsledky v oblasti manažérstva kvality a aplikovanej štatistiky predkladajú vo forme technických noriem. S rozvojom medzinárodnej spolupráce v produkcii tovarov sa normalizácia v uvedených dvoch oblastiach stáva medzinárodnou.

2 Normalizácia na medzinárodnej úrovni

Najväčším svetovým neštátnym orgánom, ktorý produkuje technické normy, je medzinárodná organizácia pre normalizáciu – ISO. Členmi ISO sú národné normalizačné orgány z viac ako sto krajín. V rámci ISO pôsobia aj medzinárodné technické komisie ISO/TC 176 *Quality management and quality assurance* a ISO/TC 69 *Application of statistical methods*, ktoré majú priamy súvis s kvalitou produkcie.

Technická komisia ISO/TC 69 bola založená v roku 1948 a je zodpovedná za normalizáciu, terminológiu, prezentáciu a interpretáciu výsledkov skúšok a kontrol. Má zodpovednosť aj za normalizáciu podmienok aplikácie štatistických metód pri riadení kvality produktov. Termín produkt znamená nielen výrobok ale aj proces, službu, marketing, servis predaného produkt a pod. Technická komisia ISO/TC 69 zabezpečuje funkciu poradného orgánu v oblasti aplikácie štatistických metód pre všetky technické komisie v rámci ISO.

V minulosti bola normalizačná činnosť v oblasti štatistiky zameraná predovšetkým na terminológiu a výberové metódy na preberaciu kontrolu dodávky. Dnešný súbor noriem a technických správ pokrýva terminológiu, všeobecné metódy (odhady, testovanie hypotéz a pod.), výberové metódy, presnosť metód a výsledkov meraní (neistota, opakovatelnosť, reprodukovatelnosť a pod.) spôsobilosť meradiel a procesov, detekčná schopnosť a iné.

Na plnení pracovného programu ISO/TC 69 sa zúčastňujú experti, ktorí majú skúsenosti z aplikácie štatistických metód v rôznych špeciálnych oblastiach (automobilový priemysel, telekomunikácie, skušobne atď.). Je zrejmé, že pracovný program a technické možnosti (metódy a prístupy), ktoré ISO/TC 69 prijíma, v podstatnej miere určujú aktívne zúčastnený experti.

Technická komisia ISO/TC 69 zodpovedá za prípravu medzinárodných noriem v oblasti aplikácie štatistických metód a to predovšetkým v riadení kvality. V tejto prierezovej oblasti

neexistujú európske normy EN. ISO/TC 69 udržuje integrovaný systém všeobecných noriem, ktoré dobre zodpovedajú súčasnému štatistickému mysleniu v praxi. Normy pomáhajú organizáciám identifikovať a implementovať všetky dôležité štatistické okolnosti vtedy, keď sa vytvárajú, zbierajú, analyzujú, prezentujú, vyhodnocujú a interpretujú dáta.

Technická komisia ISO/TC 69 má pravidelnú spoluprácu s už spomenutou ISO/TC 176, ktorá vypracovala normy radu ISO 9000, a s IEC/TC 56, ktorej predmetom záujmu je spoľahlivosť. Okrem týchto dvoch komisií spolupracuje aj s ďalšími.

V rámci technickej komisie ISO/TC 69 sú vytvorené subkomisie SC, pracovné skupiny WG a ad hoc skupiny AHG. Všetky AHG a niektoré WG podliehajú priamo TC 69. Ostatné WG pracujú v rámci štyroch subkomisií SC 1, SC 4, SC 5 a SC 6.

Priamo TC 69 podliehajú:

- CAG 1 Chairman Advisory Group
- AHG 5 Support that ISO/TC69 can provide to National Statistical Office
- AHG 4 Statistical standards and software
- AHG 2 Evaluation of Conformity
- AHG 3 Committee on Six Sigma
- AHG 1 Application of ISO statistical standards to the clauses of ISO/IEC 17025
- WG 3 Statistical interpretation of data
- WG 9 Random variate generation
- WG 10 Six Sigma
- SC 1 Terminology and symbols
  - WG 2 Addenda and corrigenda on ISO 3534-1 and 2
- SC 4 Applications of statistical methods in process management
  - WG 6 Process capability and performance measures
  - WG 9 Quality capability statistics
  - WG 10 Control charts standards
  - WG 11 Capability and performance
- SC 5 Acceptance sampling
  - WG 1 Classification of sampling problems
  - WG 2 Sampling procedures for inspection by attributes
  - WG 3 Sampling procedures for inspection by variables
  - WG 4 Sequential and continuous sampling plans
  - WG 5 Parts per Million (ppm) sampling
  - WG 6 Accept-zero sampling plans
  - WG 7 Random sampling procedures
- SC 6 Measurement methods and results
  - WG 1 Accuracy of measurement methods and results
  - WG 4 Statistical aspects of the preparation and use of reference materials
  - WG 5 Capability of detection

WG 7 Statistical methods to support measurement uncertainty evaluation

3 Normalizácia na národnej úrovni

Národné ekvivalenty medzinárodných komisií ISO/TC 176 *Quality management and quality assurance* a ISO/TC69 *Application of statistical methods* sú technické komisie TK *Manažérstvo kvality a zabezpečovanie kvality* a TK 71 *Aplikácia štatistických metód v riadení kvality*. Obidve TK pôsobia na Slovenskom ústave technickej normalizácie v Bratislave. Predmetom nášho záujmu je TK 71, ktorá bola založená v roku 1996. Medzi jej hlavné činnosti patrí prevzatie ISO noriem prekladom do sústavy Slovenských technických noriem, ktoré sa označujú STN ISO, a zastupovanie P- a O- členstva Slovenskej republiky v TC69 a v jej subkomisiách. V súčasnosti zastupujeme P- členstvo v TC69 a v SC 4

Dvaja z členov TK71 pracujú ako experti v pracovnej skupine TC69/WG3 *Štatistická interpretácia dát*. V tejto WG sa vytvárajú normy radu ISO 16269

4 Tvorba medzinárodných noriem ISO

Medzinárodná norma ISO je výsledkom dohody členov ISO. Môže sa používať buď v originály, alebo zavedená prekladom do sústavy národných noriem. Tvorba noriem prebieha v šiestich etapách.

1. Etapa návrhu novej pracovnej položky NWIP, NWI, NP.
2. Prípravná etapa, ktorá končí vypracovaním prvého návrhu CD a jeho predložením na hlasovanie riadnym členom TC.
3. Etapa prípravy v komisii, ak je úspešná, končí zaregistrovaním CD ako DIS.
4. Etapa odsúhlasovania. V tejto etape musí byť anglická alebo francúzska verzia DIS rozoslaná na päťmesačné hlasovanie riadnym členom TC. Táto etapa sa končí zaregistrovaním textu ako konečný návrh medzinárodnej normy FDIS.
5. Etapa schvaľovania. Konečný návrh FDIS sa rozposiela na dvojmesačné schvaľovanie, ktorý po schválení postupuje do etapy vydania.
6. Etapa vydania končí vydaním medzinárodnej normy.

Medzinárodné normy ISO podliehajú každých päť rokov systematickým previerkam. Okrem noriem vydáva ISO aj Technické správy (TR), Technické špecifikácie (TS), Verejne dostupné špecifikácie (PAS), Priemyselné technické dohody (ITA), Pokyny resp. príručky (Guide´s), z ktorých niektoré sú spoločné s IEC.

5 Vydané a rozpracované ISO normy zo štatistických metód

Technická komisia ISO/TC 69 má na internetovej stránke aktuálne informácie o vydaných a rozpracovaných normách. Použili sme materiál, ku ktorému máme prístup, a spracovali sme ho do dvoch tabuliek. V prílohe 1 uvádzame vydané normy do . Tabuľku vydaných noriem sme doplnili o ISO normy prevzaté prekladom do sústavy Slovenských technických noriem označovaných ako STN ISO. V prílohe 2 sú uvedená rozpracované ISO normy z aplikácie štatistických metód..

6 Záver

So štatistickým riadením kvality súvisia všetky publikácie [1] až [10] citované v literatúre.

Literatúra

[1] GARAJ, I., JANIGA, I. *Dvojstranné tolerančné medze pre neznámu strednú hodnotu a rozptyl normálneho rozdelenia* . Bratislava: STU, 2002. 147 s. ISBN 80-227-1779-7.

[2] GARAJ, I., JANIGA, I. *Dvojstranné tolerančné medze normálnych rozdelení s neznámymi strednými hodnotami a s neznámym spoločným rozptylom.* Bratislava: STU, 2004. 218 s. ISBN 80-227-2019-4.

[3] GARAJ, I., JANIGA, I. *Jednostranné tolerančné medze normálneho rozdelenia s neznámou strednou hodnotou a rozptylom. One sided tolerance limits of normal distribution with unknown mean and variability.* Bratislava: STU, 2005. 214 s. ISBN 80-227-2218-9.

[4] JANIGA, I., PALENČÁR, R. STN ISO 7873 Regulačné diagramy aritmetických priemerov s výstražnými medzami. Slovenská technická norma, SÚTN, Bratislava, 1999.

[5] TEREK, M., HRNČIAROVÁ, Ľ. *Štatistické riadenie kvality*. Vydavateľstvo IURA EDITION, 2004, 234 s. ISBN 80-89047-97-1.

[6] TEREK, M., HRNČIAROVÁ, Ľ. *Analýza spôsobilosti procesu*. Vydavateľstvo EKONÓM, Ekonomická univerzita v Bratislave, 2001. 205 s. ISBN 80-225-1443-8.

[7] TEREK, M.: *Viackriteriálne rozhodovanie v Taguchiho metóde*. Slovenská štatistika a demografia 2/2006. s. 21 – 39. ISSN 1210-1095.

[8] PALENČÁR, R., RUIZ, J.M., JANIGA, I., HORNÍKOVÁ, A. *Štatistické metódy v skúšobných a kalibračných laboratóriách*. Bratislava: Grafické štúdio Ing. Peter Juriga, 2001. 380 s. ISBN 80-968449-3-8.

[9] GARAJ, I. Sequential sampling plan of Poisson distribution. In *Mechanical Enginneering: International Conference: Proceedings.* Bratislava: SjF STU, 2001. ISBN 80–227-1616-2, p. 670-675.

[10] GARAJ, I. Požiadavky na rozsah náhodného výberu jednostranných preberacích plánov meraním. In *FORUM STATISTICUM SLOVACUM*. ISSN 1336-7420. 1/2006, s. 38-43.

Tento článok vznikol s podporou grantového projektu VEGA č. 1/3182/06 Zlepšovanie kvality produkcie strojárskych výrobkov pomocou moderných štatistických metód.

Kontaktné adresy autora:
Doc. RNDr. Ivan Janiga, PhD., Katedra matematiky, SjF STU, Nám. slobody 17,
812 31 Bratislava, tel.: +421-2-57296-160, e-mail: ivan.janiga@stuba.sk

Príloha 1

Štatistické metódy –normy publikované v IS/TC 69 a prebrané do sústavy STN k 24. 4. 2007

ISO/TC 69		STN	
ISO 2602:1980	Statistical interpretation of test results -- Estimation of the mean -- Confidence interval	**STN ISO 2602: 1993**	Štatistická interpretácia výsledkov skúšok. Odhad priemeru. Interval spoľahlivosti
ISO 2854:1976	Statistical interpretation of data -- Techniques of estimation and tests relating to means and variances	**–––**	
ISO 3301:1975	Statistical interpretation of data -- Comparison of two means in the case of paired observations	**STN ISO 3301: 1993**	Štatistická interpretácia údajov. Porovnanie dvoch priemerov v prípade párových pozorovaní
ISO 3494:1976	Statistical interpretation of data -- Power of tests relating to means and variances	**–––**	
ISO 5479:1997	Statistical interpretation of data -- Tests for departure from the normal distribution	**–––**	
ISO 10725:2000	Acceptance sampling plans and procedures for the inspection of bulk materials	**STN ISO 10725: 2005**	Výberové preberacie plány a postupy pri kontrole nekusových materiálov
ISO 11453:1996	Statistical interpretation of data -- Tests and confidence intervals relating to proportions	**STN ISO 11453: 2002**	Štatistická interpretácia údajov. Testy a intervaly spoľahlivosti pre podiely
ISO 11453:1996/Cor 1:1999		**STN ISO 11453: 2002**	
ISO 11648-1:2003	Statistical aspects of sampling from bulk materials -- Part 1: General principles	**–––**	
ISO 11648-2:2001	Statistical aspects of sampling from bulk materials -- Part 2: Sampling of particulate materials	**–––**	
ISO/TR 13425:2006	Guidelines for the selection of statistical methods in standardization and specification	**–––**	**Zavedená staršia verzia:** **TNI ISO/TR 13425: 2000** Príručka pre voľbu štatistických metód v normalizácii a špecifikácii (ISO/TR 13425: 1995)

ISO 16269-6:2005	Statistical interpretation of data -- Part 6: Determination of statistical tolerance intervals	**STN ISO 16269-6**	Štatistická interpretácia dát. Časť 6: Stanovenie štatistických tolerančných intervalov. Revízia STN ISO 3207: 1993 **V pláne technickej normalizácie - bude vydaná v roku 2007**
ISO 16269-7:2001	Statistical interpretation of data -- Part 7: Median -- Estimation and confidence intervals	**STN ISO 16269-7: 2005**	Štatistická interpretácia dát. Časť 7: Medián. Odhadovanie a intervaly spoľahlivosti
ISO 16269-8:2004	Statistical interpretation of data -- Part 8: Determination of prediction intervals	**STN ISO 16269-8**	Štatistická interpretácia dát. Časť 8: Stanovenie predikčných intervalov **V pláne technickej normalizácie – bude vydaná v roku 2007**
	ISO/TC 69/SC 1		**STN**
ISO 3534-1:2006	Statistics -- Vocabulary and symbols -- Part 1: General statistical terms and terms used in probability	**–––**	**V pláne technickej normalizácie** **Zavedená staršia verzia:** **STN ISO 3534-1: 1999** Štatistika. Slovník a značky. Časť 1: Pravdepodobnosť a všeobecné štatistické termíny (ISO 3534: 1993)
ISO 3534-2:2006	Statistics -- Vocabulary and symbols -- Part 2: Applied statistics	**–––**	**Zavedená staršia verzia:** **STN ISO 3534-2: 1999** Štatistika. Slovník a značky. Časť 2: Štatistické riadenie kvality (ISO 3534-2: 1993)
ISO 3534-3:1999	Statistics -- Vocabulary and symbols -- Part 3: Design of experiments	**STN ISO 3534-3: 2003**	Štatistika. Slovník a značky. Časť 3: Navrhovanie experimentov
	ISO/TC 69/SC 4		**STN**
ISO 7870:1993	Control charts -- General guide and introduction	**STN ISO 7870: 2000**	Regulačné diagramy. Všeobecná príručka a úvod
ISO/TR 7871:1997	Cumulative sum charts -- Guidance on quality control and data analysis using CUSUM	**TNI ISO/TR 7871: 1999**	Diagramy kumulatívnych súčtov. Príručka pre riadenie kvality a analýzu údajov

	techniques		použitím metódy CUSUM (kumulatívnych súčtov)
ISO 7873:1993	Control charts for arithmetic average with warning limits	**STN ISO 7873: 2000**	Regulačné diagramy aritmetických priemerov s výstražnými medzami
ISO 7966:1993	Acceptance control charts	**STN ISO 7966: 2000**	Preberacie regulačné diagramy
ISO 8258:1991	Shewhart control charts	**STN ISO 8258: 1995**	Shewhartové regulačné diagramy
ISO 8258:1991/Cor1:1993		**–––**	
ISO 11462-1:2001	Guidelines for implementation of statistical process control (SPC) -- Part 1: Elements of SPC	**–––**	
ISO 21747:2006	Statistical methods -- Process performance and capability statistics for measured quality characteristics	**–––**	
	ISO/TC 69/SC 5		**STN**
ISO 2859-1:1999	Sampling procedures for inspection by attributes -- Part 1: Sampling schemes indexed by acceptance quality limit (AQL) for lot-by-lot inspection	**STN ISO 2859-1: 2004**	Štatistické prebierky porovnávaním. Časť 1: Preberacie plány AQL na kontrolu každej dávky v sérii
ISO 2859-1:1999/Cor 1:2001		**STN ISO 2859-1: 2004**	
ISO 2859-2:1985	Sampling procedures for inspection by attributes -- Part 2: Sampling plans indexed by limiting quality (LQ) for isolated lot inspection	**STN ISO 2859-2: 1992**	Štatistické prebierky porovnávaním. Časť 2: Preberacie plány LQ pre kontrolu izolovaných dávok
ISO 2859-3:2005	Sampling procedures for inspection by attributes -- Part 3: Skip-lot sampling procedures	**STN ISO 2859-3: 2006**	Štatistické prebierky porovnávaním. Časť 3: Občasná prebierka
ISO 2859-4:2002	Sampling procedures for inspection by attributes -- Part 4: Procedures for assessment of declared quality levels	**STN ISO 2859-4: 2006**	Štatistické prebierky porovnávaním. Časť 4: Postupy pri posudzovaní deklarovaných úrovní kvality
ISO 2859-5:2005	Sampling procedures for inspection by attributes -- Part 5: System of sequential sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection	**STN ISO 2859-5: 2007**	Štatistické prebierky porovnávaním. Časť 5: Systém sekvenčných preberacích plánov AQL na kontrolu každej dávky v sérii

ISO 2859-10:2006	Sampling procedures for inspection by attributes -- Part 10: Introduction to the ISO 2859 series of standards for sampling for inspection by attributes	**–––**	
ISO 3951-1:2005	Sampling procedures for inspection by variables -- Part 1: Specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection for a single quality characteristic and a single AQL	**STN ISO 3951-1: 2007**	Štatistické prebierky meraním. Časť 1: Špecifikácia preberacích plánov AQL jedným výberom na kontrolu každej dávky pre jeden znak kvality a jednu hodnotu AQL
ISO 3951-2:2006	Sampling procedures for inspection by variables -- Part 2: General specification for single sampling plans indexed by acceptance quality limit (AQL) for lot-by-lot inspection of independent quality characteristics	**–––**	
ISO 3951-5:2006	Sampling procedures for inspection by variables -- Part 5: Sequential sampling plans indexed by acceptance quality limit (AQL) for inspection by variables (known standard deviation)	**–––**	
ISO 8422:2006	Sequential sampling plans for inspection by attributes	**–––**	**Zavedená staršia verzia: STN ISO 8422: 1995** Preberacie plány postupným výberom pri kontrole porovnávaním (ISO 8422: 1991)
ISO 8423:1991	Sequential sampling plans for inspection by variables for percent nonconforming (known standard deviation)	**STN ISO 8423: 1995**	Preberacie plány postupným výberom pri kontrole meraním percenta nezhodných jednotiek (Známa smerodajná odchýlka)
ISO 8423:1991/Cor 1:1993		**–––**	
ISO/TR 8550:1994	Guide for the selection of an acceptance sampling system, scheme or plan for inspection of discrete items in lots	**–––**	
ISO 13448-1:2005	Acceptance sampling procedures based on the allocation of priorities principle (APP) -- Part 1: Guidelines for the APP approach	**–––**	

ISO 13448-2:2004	Acceptance sampling procedures based on the allocation of priorities principle (APP) -- Part 2: Coordinated single sampling plans for acceptance sampling by attributes	**–––**	
ISO 14560:2004	Acceptance sampling procedures by attributes -- Specified quality levels in nonconforming items per million	**–––**	
ISO 18414:2006	Acceptance sampling procedures by attributes -- Accept-zero sampling system based on credit principle for controlling outgoing quality	**–––**	
ISO 21247:2005	Combined accept-zero sampling systems and process control procedures for product acceptance	**–––**	
	ISO/TC 69/SC 6		**STN**
ISO 5725-1:1994	Accuracy (trueness and precision) of measurement methods and results -- Part 1: General principles and definitions	**STN ISO 5725-1: 2000**	Presnosť (správnosť a zhodnosť) metód a výsledkov merania. Časť 1: Všeobecné zásady a definície
ISO 5725-1:1994/Cor 1:1998		**STN ISO 5725-1: 2000**	
ISO 5725-2:1994	Accuracy (trueness and precision) of measurement methods and results -- Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method	**STN ISO 5725-2: 2000**	Presnosť (správnosť a zhodnosť) metód a výsledkov merania. Časť 2: Základná metóda stanovenia opakovateľnosti a reprodukovateľnosti normalizovanej metódy merania
ISO 5725-2:1994/Cor 1:2002		**–––**	
ISO 5725-3:1994	Accuracy (trueness and precision) of measurement methods and results -- Part 3: Intermediate measures of the precision of a standard measurement method	**STN ISO 5725-3: 2000**	Presnosť (správnosť a zhodnosť) metód a výsledkov merania. Časť 3: Medziľahlé miery zhodnosti normalizovanej metódy merania
ISO 5725-3:1994/Cor 1:2001		**STN ISO 5725-3: 2000**	
ISO 5725-4:1994	Accuracy (trueness and precision) of measurement methods and results -- Part 4:	**STN ISO 5725-4: 2000**	Presnosť (správnosť a zhodnosť) metód a výsledkov merania. Časť 4: Základné

	Basic methods for the determination of the trueness of a standard measurement method		metódy stanovenia správnosti normalizovanej metódy merania
ISO 5725-5:1998	Accuracy (trueness and precision) of measurement methods and results -- Part 5: Alternative methods for the determination of the precision of a standard measurement method	**STN ISO 5725-5: 2002**	Presnosť (správnosť a zhodnosť) metód a výsledkov meraní. Časť 5: Alternatívne metódy stanovenia zhodnosti normalizovanej metódy merania
ISO 5725-5:1998/Cor 1:2005		**–––**	
ISO 5725-6:1994	Accuracy (trueness and precision) of measurement methods and results – Part 6: Use in practice of accuracy values	**STN ISO 5725-6: 2000**	Presnosť (správnosť a zhodnosť) metód a výsledkov merania. Časť 6: Použitie hodnôt mier presnosti v praxi
ISO 5725-6:1994/Cor 1:2001		**–––**	
ISO 10576-1:2003	Statistical methods – Guidelines for the evaluation of conformity with specified requirements – Part 1: General principles	**–––**	
ISO 11095:1996	Linear calibration using reference materials	**STN ISO 11095:2002**	Lineárna kalibrácia s použitím referenčných materiálov
ISO 11843-1:1997	Capability of detection – Part 1: Terms and definitions	**STN ISO 11843-1:2002**	Detekčná schopnosť. Časť 1: Termíny a definície
ISO 11843-1:1997/Cor 1:2003		**–––**	
ISO 11843-2:2000	Capability of detection – Part 2: Methodology in the linear calibration case	**STN ISO 11843-2:2002**	Detekčná schopnosť. Časť 2: Metodika lineárnej kalibrácie
ISO 11843-3:2003	Capability of detection – Part 3: Methodology for determination of the critical value for the response variable when no calibration data are used	**STN ISO 11843-3:2005**	Detekčná schopnosť. Časť 3: Metóda na stanovenie kritickej hodnoty ozvovej premennej bez použitia kalibračných údajov
ISO 11843-4:2003	Capability of detection – Part 4: Methodology for comparing the minimum detectable value with a given value	**–––**	**V pláne technickej normalizácie**
ISO 13528:2005	Statistical methods for use in proficiency	**–––**	

	testing by interlaboratory comparisons		
ISO/TS 21748:2004	Guidance for the use of repeatability, reproducibility and trueness estimates in measurement uncertainty estimation	**STN P ISO/TS 21748: 2006**	Návod na používanie odhadov opakovateľnosti, reprodukovateľnosti a správnosti v odhadovaní neistoty merania
ISO/TS 21749:2005	Measurement uncertainty for metrological applications -- Repeated measurements and nested experiments	**–––**	
ISO/TR 22971:2005	Accuracy (trueness and precision) of measurement methods and results -- Practical guidance for the use of ISO 5725-2:1994 in designing, implementing and statistically analysing interlaboratory repeatability and reproducibility results	**–––**	

Príloha 2

Štatistické metódy – rozpracované normy v ISO/TC 69 k 3. 5. 2007

Rozpracované normy		ISO/TC 69 -vydané normy, ktoré sú v revízii	ktoré budú STN - vydané normy, v revízii
ISO/CD 16269-4	Statistical interpretation of data -- Part 4: Detection and treatment of outliers	––	
ISO/CD TR 18532	Guidance on the application of statistical methods to quality and standardization	––	
ISO/CD 28640	Random variate generation methods	––	
ISO/CD TR 29901	Applications of statistical methods -- Design and analysis of full factorial experiments ilustrated with four factors		
ISO/TC 69/SC 1 *Názvoslovie a značky* Nie sú rozpracované dokumenty			
ISO TC 69/SC 4			
ISO/DIS 7870-1	Control charts -- Part 1: General guidelines	**Revízia ISO 7870: 1993**	**Revízia STN ISO 7870: 2000**
ISO/WD 11462-2	Guidelines for implementation of statistical process control (SPC) -- Part 2: Catalogue of tools and techniques		
ISO/CD TR 12783	Process capability and performance measures		
ISO/DIS 13700	Machine performance studies -- Measured data -- Discrete parts		

Rozpracované normy		ISO/TC 69 -vydané normy, ktoré sú v revízii	ktoré budú STN - vydané normy, v revízii
ISO/CD 22514-1	Capability and performance -- Part 1: General principles and concepts		

Rozpracované normy		ISO/TC 69 -vydané normy, ktoré sú v revízii	ktoré budú STN - vydané normy, v revízii
ISO TC 69/SC 5			
ISO/CD 2859-2	Sampling procedures for inspection by attributes -- Part 2: Sampling plans indexed by limiting quality (LQ) for isolated lot inspection	**Revízia ISO 2859-2: 1985**	**Revízia STN ISO 2859-2: 2002**
ISO/WD 3951-4	Sampling procedures for inspection by variables -- Part 4: Procedures for assessment of declared quality levels		
ISO/DIS 8423	Sequential sampling plans for inspection by variables for percent nonconforming (known standard deviation)	**Revízia ISO 8423: 1991** **ISO 8423:1991/Cor 1: 1993**	**Revízia STN ISO 8423: 1995**
ISO/TR 8550-1	Guidance on the selection and usage of acceptance sampling systems for inspection of discrete items in lots -- Part 1: Acceptance sampling	**Revízia ISO/TR 8550: 1994**	**––**
ISO/CD TR 8550-2	Guidance on the selection and usage of acceptance sampling systems for inspection of discrete items in lots -- Part 2: Sampling by attributes	**Revízia ISO/TR 8550: 1994**	**––**
ISO/TR 8550-3	Guidance on the selection and usage of acceptance sampling systems for inspection of discrete items in lots -- Part 3: Sampling by variables	**Revízia ISO/TR 8550: 1994**	**––**
ISO/DIS 24153	Random sampling and randomization procedures		
ISO/CD 28801	Double sampling plans for inspection by		

Rozpracované normy		ISO/TC 69 -vydané normy, ktoré sú v revízii	ktoré budú STN - vydané normy, v revízii
	attributes with minimal sample sizes, indexed by producer's risk quality (PRQ) and consumer's risk quality (CRQ)		
	ISO TC 69/SC 6		
ISO/CD 11843-5	Capability of detection -- Part 5: Methodology in the linear and non-linear calibration cases		
ISO/AWI 27877	Concepts of precision and uncertainty for qualitative data		
ISO/AWI 28037	Use of linear calibration curves		

Dva způsoby vyhodnocení robustního návrhu

Eva Jarošová, Jiří Michálek

Abstract: The paper deals with the experiment for robust design's analysis. Two groups of factors are distinguished, i.e. control and noise factors. The noise factors vary during the normal process and contribute to the response variation. In the experiment their levels can be controlled to some extent. Robust design determines such process setting that the process is less sensitive to the noise variation. The aim of the experiment is to find the control factors' levels at which consequences of the noise variation will be reduced. That is why control factors with dispersion effects have to be identified. Various methods are used for dispersion modeling. Two modeling strategies are described in the paper. One of them represents the original Taguchi's approach and uses signal/noise ratio as a response. Only control factors are included in the model. The other approach consists in response variable modeling. Both groups of factors are involved in the model. Based on the fitted response model the transmitted variance model is computed. These two methods are applied to the data taken from literature.

Key words: analysis of variance, response surface model, design of experiment, signal/noise ratio, control and noise factors

Úvod

Cílem tohoto příspěvku je ukázat na konkrétním příkladě návrhu experimentu dva přístupy k otázce robustnosti, tj. necitlivosti vůči rušivým faktorům. S myšlenkou robustního návrhu přišel na počátku osmdesátých let G. Taguchi, který faktory ovlivňující sledovaný systém, např. výrobní proces, rozdělil na faktory řiditelné a rušivé. Obě skupiny faktorů mají vliv na výstupní veličinu. Rušivé faktory se mohou během normálního procesu nekontrolovaně měnit, během experimentu jsou však úrovně obou typů faktorů nastavovány podle zvoleného plánu. Protože rušivé faktory přispívají svou přítomností ke zvyšování variability výstupní veličiny, je cílem nalézt takové nastavení úrovní řiditelných faktorů, při němž bude výstupní veličina co nejblíže cílové hodnotě a současně bude dosahovat co možná nejmenší variability. První ukázka řešení vychází z původního přístupu navrženého Taguchim, druhá ukázka představuje využití metody odezvových ploch.

Příklad návrhu a data získaná při jeho provedení jsou půjčena z publikace [2]. Návrh se týká konstrukce a funkčnosti airbagu v automobilu. Sleduje se přitom spolehlivost a účinnost vystřelovacího zařízení. Na základě analýzy procesu byly stanoveny 4 řiditelné faktory, tj. obsah vodíku v plynové směsi se vzduchem (A), plnicí tlak plynu (B), velikost průřezu trysky (C) a velikost plynové nádržky (D) a 3 rušivé faktory, tj. doba zahřívání plnicího systému (X), skutečný tlak při nafouknutí airbagu (Y), který je ovlivněn variabilitou v materiálu a v součástkách zařízení a konečně hmotnost rozbušky (Z). Všechny faktory jsou v experimentu uvažovány na 2 úrovních, viz tab.1.

	Úroveň			Úroveň	
Faktory řiditelné	1	2	Faktory rušivé	1	2
A: Vodík (%)	12,5	14,0	*X*: Zahřívání	pod 5 min	nad 1 hod
B: Tlak (lb/in^2)	2500	3200	*Y*: Tlak (lb/in^2)	6500	7000
C: Tryska (mm^2)	120	50	*Z*: Rozbuška (μg)	170	200
D: Nádržka (cm^3)	400	600			

Tab.1 – Zkoumané faktory a jejich úrovně

V příkladu byly uvažovány 3 výstupní veličiny: hodnota maximálního tlaku ve zkušebním tanku (*EC1*), rychlost šíření tlaku ve zkušebním tanku (*EC2*) a tlak při nafouknutí (*EC3*). V tab. 2 jsou u každé z těchto veličin uvedeny dvě hodnoty, nejhorší a nejlepší. Tyto hodnoty jsou potřebné pro výpočet nové výstupní veličiny *OEC*, jejíž konstrukce podle [2] má umožnit posuzování vlivu faktorů na všechny 3 výstupní veličiny najednou. Váhy jednotlivých výstupních veličin použité při výpočtu *OEC* jsou uvedeny v tab. 2.

Výstupní veličina	Nejhorší	Nejlepší	Kritérium	Váha
EC1: Maximální tlak (kPa)	400	600	Větší je lepší	0,5
EC2: Rychlost šíření (kPa/ms)	25	8	Menší je lepší	0,3
EC3: Tlak (MPa)	70,5	45	Menší je lepší	0,2

Tab. 2 – Popis výstupních veličin

Pro řiditelné faktory byla vybrána ortogonální oblast L8 (termín používaný Taguchim představuje dílčí faktoriální návrh 2^{4-1}), pro rušivé faktory ortogonální oblast L4 (návrh 2^{3-1}). Obě oblasti jsou kříženy, takže experiment obsahuje celkem 32 zkoušek. Při každé zkoušce se měřily hodnoty výstupních veličin *EC1*, *EC2*, *EC3*. Nastavení úrovní u 8 kombinací řiditelných faktorů a 4 kombinací rušivých faktorů je patrné z tab.3, kde jsou pro představu uvedeny alespoň hodnoty výstupní veličiny *EC1*.

					Oblast L4			
				X	1	1	2	2
				Y	1	2	1	2
Oblast L8				*Z*	1	2	2	1
A	*B*	*C*	*D*					
1	1	1	1		450	445	470	411
1	1	2	2		462	596	511	460
1	2	1	2		463	489	509	435
1	2	2	1		519	527	422	432
2	1	1	2		440	493	482	489
2	1	2	1		404	445	445	491
2	2	1	1		537	597	545	565
2	2	2	2		486	556	508	532

Tab.3 – Uspořádání experimentu

Z každé trojice hodnot *EC1*, *EC2* a *EC3* byly podle vzorce

$$OEC = \left|\frac{EC1-400}{600-400}\right| \cdot 0,5 + \left(1 - \left|\frac{EC2-8}{25-8}\right|\right) \cdot 0,3 + \left(1 - \left|\frac{EC3-45}{70,5-45}\right|\right) \cdot 0,2$$

vypočteny hodnoty veličiny *OEC*. Sloupce v následující tabulce odpovídají jednotlivým kombinacím rušivých faktorů.

OEC1	*OEC2*	*OEC3*	*OEC4*
49,95	41,25	59,07	33,93
36,87	67,63	65,59	62,45
49,28	43,82	36,07	53,06
56,08	70,90	60,50	45,78
35,51	40,47	40,27	37,06
61,50	73,76	60,96	81,05
57,58	63,90	53,86	59,47

Tab. 4 – Vypočtené hodnoty veličiny OEC

Vzhledem ke způsobu konstrukce veličiny *OEC* je při vyhodnocení podle Taguchiho použito kritérium „větší je lepší“.

Taguchiho přístup

Pro vyhodnocení experimentu byla použita kritéria založená na poměru signál/šum, jak je u Taguchiho přístupu obvyklé. Kritérium „větší je lepší“ má tvar

$$S/N = -10Log\frac{1}{n}\sum_{1}^{n} y_i^{-2}, \qquad (1)$$

kritérium „menší je lepší“ je dáno vztahem

$$S/N = -10Log\frac{1}{n}\sum_{1}^{n} y_i^{2}, \qquad (2)$$

kde y_i je hodnota výstupní veličiny při i-té kombinaci rušivých faktorů a n je počet kombinací rušivých faktorů. Pro každý řádek tab. 4 byly pro uvažované výstupní veličiny vypočteny průměry a hodnoty *S/N* . Na takto získané hodnoty byla aplikována analýza rozptylu.

Analýza rozptylu pro S/N

Snahou je odhalit ty řiditelné faktory, které nejvíce přispívají k vyšším hodnotám veličiny *OEC* při co nejmenší variabilitě. Tabulka analýzy rozptylu v Minitabu má tvar:

Analysis of Variance for SN ratios „larger is better“

Source	DF	Seq SS	Adj SS	Adj MS	F	P
A	1	4,1664	4,16641	4,16641	663,65	0,025
B	1	1,8596	1,85963	1,85963	296,21	0,037
C	1	1,1726	1,17258	1,17258	186,78	0,046
D	1	1,8939	1,89387	1,89387	301,67	0,037
A*B	1	5,7846	5,78457	5,78457	921,40	0,021
A*C	1	4,9127	4,91275	4,91275	782,53	0,023
Residual Error	1	0,0063	0,00628	0,00628		
Total	7	19,7961				

Všechny faktory a uvedené interakce jsou statisticky významné na hladině 0,05. Rovnice lineárního modelu (procedura *Factorial*) má při kódování úrovní -1 a 1 tvar

$$odh\ S/N = 33{,}9217 + 0{,}7217A + 0{,}4821B - 0{,}3828C + 0{,}4866D + 0{,}8503AB - 0{,}7836AC$$

Nejvyšší hodnotu, tedy hodnotu 37,6288, má odhad *S/N* při nastavení $A = 1$, $B = 1$, $C = -1$, $D = 1$. Celkový průměr, s nímž můžeme tuto hodnotu srovnávat, je 33,9217.

Analýza rozptylu pro průměry veličiny OEC

Tentokrát zkoumáme pouze vliv řiditelných faktorů na úroveň hodnot veličiny *OEC*.

Analysis of Variance for Means

Source	DF	Seq SS	Adj SS	Adj MS	F	P
A	1	109,539	109,539	109,539	100,53	0,063
B	1	40,804	40,804	40,804	37,45	0,103
C	1	43,980	43,980	43,980	40,36	0,099
D	1	58,848	58,848	58,848	54,01	0,086
A*B	1	249,622	249,622	249,622	229,08	0,042
A*C	1	225,224	225,224	225,224	206,69	0,044
Residual Error	1	1,090	1,090	1,090		
Total	7	729,106				

Necháme-li v modelu stejné členy jako při analýze *S/N*, bude mít tvar

$$odh\ OEC = 52{,}465 + 3{,}7A + 2{,}258B - 2{,}345C + 2{,}712D + 5{,}586AB - 5{,}306AC\ .$$

Pokud uvažujeme nastavení úrovní faktorů zjištěné pomocí modelu *S/N*, zvýší se průměrná hodnota veličiny *OEC* na hodnotu 74,426. Otázkou ovšem je, do jaké míry je přírůstek statisticky významný.

Je samozřejmě možné aplikovat uvedený postup zvlášť na každou ze 3 výstupních veličin. V žádném z uvažovaných případů se však neprokázal významný vliv ani u jednoho řiditelného faktoru či jejich interakce.

Model odezvy

Druhý způsob vyhodnocení spočívá v modelování hodnot původní výstupní veličiny (odezvy), přičemž do modelu jsou zařazeny jak řiditelné, tak rušivé faktory. Významné hlavní efekty či interakce řiditelných faktorů svědčí o vlivu těchto faktorů na střední hodnotu odezvy, významné interakce řiditelných a rušivých faktorů naznačují, kterými

řiditelnými faktory můžeme ovlivnit velikost variability sledované veličiny. Používá se tzv. duální přístup, to znamená, že z původního modelu s oběma typy faktorů se odvodí dva modely, jeden pro střední hodnotu, druhý pro rozptyl odezvy (viz např. [1]). Odvozené modely obsahují již jen řiditelné faktory. Model střední hodnoty umožňuje vybrat úrovně řiditelných faktorů vhodné pro posun střední hodnoty odezvy žádoucím směrem, pomocí modelu rozptylu lze odhadnout, při kterých úrovních řiditelných faktorů by měla být variabilita odezvy menší.

Rovnici původního modelu můžeme zapsat ve tvaru

$$y(\mathbf{x},\mathbf{z}) = \beta_0 + \mathbf{x}^T\boldsymbol{\beta} + \mathbf{x}^T\boldsymbol{\Gamma}\mathbf{x} + \mathbf{z}^T\boldsymbol{\gamma} + \mathbf{x}^T\boldsymbol{\Delta}\mathbf{z} + \varepsilon, \tag{3}$$

kde člen $\mathbf{x}^T\boldsymbol{\beta}$ představuje hlavní efekty a interakce řiditelných faktorů, člen $\mathbf{z}^T\boldsymbol{\gamma}$ hlavní efekty rušivých faktorů (interakce rušivých faktorů se obvykle neuvažují) a člen $\mathbf{x}^T\boldsymbol{\Delta}\mathbf{z}$ odpovídá interakcím řiditelných a rušivých faktorů (matice $\boldsymbol{\Delta}$ obsahuje příslušné parametry). Předpokládáme $\varepsilon \approx N(0,\sigma^2)$ a dále nulové střední hodnoty a konstantní rozptyly rušivých faktorů a jejich nekorelovanost. Rušivé faktory *X*, *Y*, *Z* v našem případě tedy považujeme za náhodné veličiny s rozptyly po řadě σ_x^2, σ_y^2, σ_z^2.

V našem modelu uvažujeme celkem 7 faktorů, 4 řiditelné a 3 rušivé. Počet kombinací v návrhu experimentu odpovídá součinu $2^{4-1}\cdot 2^{3-1}$, tedy 32. Hodnoty odezvy jsou určeny výsledky zkoušek při jednotlivých kombinacích úrovní. Odezvu tvoří postupně veličiny *EC1*, *EC2*, *EC3* a *OEC*. Výsledky uvádíme jen pro *EC1* a *OEC*.

Pro dobrou interpretaci výsledků je při konstrukci modelu třeba zařadit jen nejdůležitější efekty. Při rozhodování o zařazení příslušného efektu lze využít t-testů, je však třeba zajistit, aby byl odhad směrodatné chyby efektů založen na dostatečném počtu stupňů volnosti. Při větším počtu faktorů je identifikace modelu poměrně pracná. Počet členů v modelu je zpočátku vysoký, protože bychom předem neměli vylučovat existenci interakcí vyššího řádu a. k dispozici není obdoba stepwise metody používané v regresi. Při postupných úpravách modelu je výhodné použít např. Paretův graf standardizovaných efektů. Uvažujeme-li hladinu významnosti $\alpha = 0{,}05$, dostaneme pro odezvu *EC1* rovnici

$$odh\ EC1 = 488 + 12{,}94A + 19{,}62B + 14{,}5X + 20{,}19AB - 16{,}81AC + 9{,}06DX\ .$$

Člen s interakcí *DX* byl doplněn vzhledem k přítomnosti efektu rušivého faktoru *X*, abychom vůbec mohli aplikovat výše uvedený postup (efekt této interakce byl ze všech interakcí s rušivým faktorem *X* největší). Při nevýznamnosti interakce není ovšem úspěšnost robustního návrhu zaručena.

Odhad střední hodnoty veličiny *EC1* při pevných úrovních řiditelných faktorů je při výše uvedených předpokladech

$$odh\ E(EC1) = 488 + 12{,}94A + 19{,}62B + 20{,}19AB - 16{,}81AC\ ,$$

odhad rozptylu

$$odh\ D(EC1) = (14{,}5 + 9{,}06D)^2\sigma_x^2 + \sigma^2\ .$$

Koeficienty v rovnicích odpovídají kódování úrovní faktorů -1 a 1 a lze je interpretovat takto: Požadujeme-li co nejvyšší hodnoty odezvy *EC1*, měli bychom volit *A* = 1, *B* = 1 a *C* = -1, volba *D* = -1 by měla znamenat menší rozptyl. Ze znázornění interakce *DX* na

obr.1a lze soudit, že při úrovni D = -1 je efekt rušivého faktoru X menší. K vyhledání optimální kombinace řiditelných faktorů je možné využít metodu odezvových ploch a hledat průnik řešení pro oba modely, tak jak to umožňuje např. program Design-Expert. V naší práci tento program nevyužíváme, protože interpretace výsledků je zřejmá.

Model pro odezvu *OEC* má tvar

$$odh\ OEC = 52{,}466 + 3{,}7A + 2{,}259B - 2{,}344C + 1{,}603X + 5{,}586AB - 5{,}306AC - 3{,}9BX\ ,$$

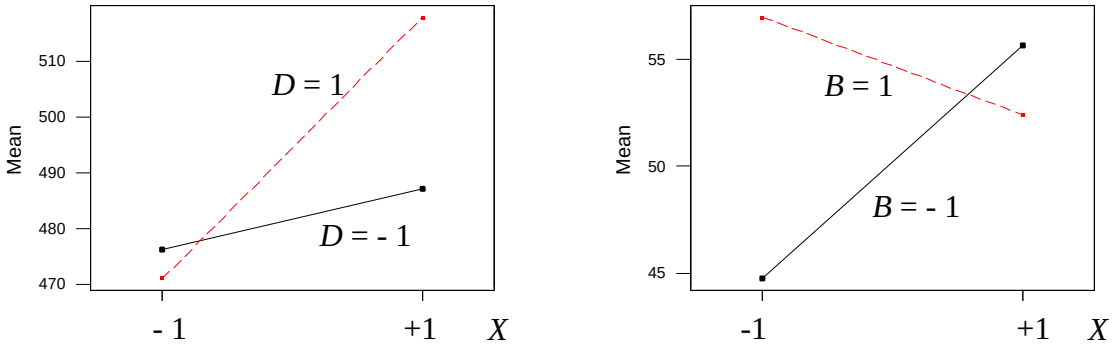
odvozený model střední hodnoty je

$$odh\ E(OEC) = 52{,}466 + 3{,}7A + 2{,}259B - 2{,}344C + 5{,}586AB - 5{,}306AC$$

a odvozený model rozptylu

$$odh\ D(OEC) = (1{,}603 - 3{,}9B)^2 \sigma_x^2 + \sigma^2 .$$

Pro dosažení co nejvyšších hodnot odezvy *OEC* je třeba volit A = 1, B = 1 a C = -1. Faktor D střední hodnotu *OEC* významně neovlivňuje. Volba úrovně B = 1 zároveň znamená určitou redukci variability, jak je patrné z rovnice pro rozptyl, případně z obr.1b, kde úsečka s menším sklonem odpovídá úrovni B = 1.



Obr. 1a,b – Graf interakcí z modelu odezvy EC1 (a) a OEC (b)

Porovnání výsledků obou přístupů

Pokud jde o veličinu *OEC*, vedou oba přístupy k podobnému řešení. Optimální nastavení úrovní faktorů *A*, *B*, *C* pro získání co největších hodnot *OEC* je stejné. U faktoru *D*, kde by se podle prvního přístupu měla volit úroveň 1, druhý přístup významný vliv nepotvrdil. Odlišnost závěrů není překvapivá. Příčinou může být jednak použití charakteristiky *S/N*, která v sobě kombinuje charakteristiky úrovně a variability, jednak skutečnost, že identifikace modelu na základě experimentu 2^{4-1} je problematická. Pokud bychom chtěli zkoumat vliv řiditelných faktorů na samotnou variabilitu, nabízí Minitab v souvislosti s Taguchiho návrhy také analýzu rozptylu, kdy odezvou je místo poměru *S/N* výběrová směrodatná odchylka. V tomto případě vycházejí efekty *A*, *B*, *C*, *AB* a *AC* významné, ale při nastavení A = 1, B = 1 a C = -1 je odhad směrodatné odchylky větší. Nebereme-li při této aplikaci v úvahu zřejmé porušení předpokladu normality, mohli bychom soudit, že při zvoleném nastavení faktorů převážilo zvýšení střední hodnoty *OEC* nad zvýšením variability. To je ovšem zase v rozporu s výsledky

druhého přístupu, podle nějž vhodným nastavením řiditelných faktorů docílíme zvýšení hodnot *OEC*, ale vliv těchto faktorů na variabilitu odezvy způsobenou rušivými faktory je minimální.

Rozdílnost závěrů v případě *EC1* je zřejmá, u dalších dvou výstupních veličin byla situace podobná. Přístup založený na analýze *S/N* neodhalil, na rozdíl od přímého modelování těchto veličin, žádné významné efekty.

Taguchiho přístup k vyhodnocování robustních návrhů je statistiky kritizován pro nedostatečnou teoretickou odůvodněnost. Sporné je používání kritérií v podobě poměru *S/N*, aplikace analýzy rozptylu na směrodatné odchylky znamená porušení předpokladu normality, při aplikaci na průměry se zanedbává existence heteroskedasticity. Taguchiho ortogonální návrhy nepředpokládají existenci interakcí vyššího řádu a často se a priori i některé dvoufaktorové interakce považují za nevýznamné. Náš příspěvek ukazuje další ze slabin Taguchiho přístupu k robustnímu návrhu. Počet stupňů volnosti pro odhad směrodatné chyby efektů je často nedostatečný a může tak být přeceněn význam efektů, které ve skutečnosti významné nejsou, případně naopak.

Literatura:

[1] Myers, R.H., Montgomery, D.C.: *Response Surface Methodology*. J. Wiley&Sons, New York, 2002

[2] Roy, R.K.: *Design of Experiments using the Taguchi Approach*. J. Wiley & Sons, New York, 2001

Adresy autorů:

Eva Jarošová
Fakulta informatiky a statistiky
VŠE v Praze
nám. W. Churchilla 4
130 67 Praha 3
jarosova@vse.cz

Jiří Michálek
ÚTIA AVČR
Pod vodárenskou věží 4
180 00 Praha 8
michalek@utia.cas.cz

Metódy neparametrickej regresnej analýzy

Jana Kalická

kalicka@math.sk

Abstrakt - Cieľom článku je prezentovať niektoré metódy neparametrickej regresnej analýzy. Neparametrické regresné metódy je vhodné použiť v prípadoch, keď regresná krivka nedostatočne flexibilne vystihuje reálne dáta alebo dáta sú výberom z iného ako normálneho rozdelenia.

Kľúčové slová: neparametrická štatistika, Kendalov a Spearmanov korelačný koeficient, jadrová funkcia, vážený priemer, poradie

1. Korelačná analýza

Mnohé problémy technickej praxe vyžadujú analyzovať n dvojíc meraní $\{X_i, Y_i\}_{i=1}^n$, ktoré sú realizáciou náhodného vektora (X, Y). Jednou z úloh ich analýzy je odhaliť a popísať mieru závislosti medzi X a Y. Najčastejšie používanou mierou závislosti je korelačný koeficient:

$$\rho_{xy} = \frac{E(x - E(X))(Y - E(Y))}{\sqrt{Var(X)Var(Y)}} \quad , \tag{1}$$

ktorého odhad z nameraných dát je v tvare:

$$r_{xy} = \frac{\sum_{i=1}^{n}(X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \overline{X})^2 \sum_{i=1}^{n}(Y_i - \overline{Y})^2}}, \tag{2}$$

pričom $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ a $\overline{Y} = \frac{1}{n}\sum_{i=1}^{n} Y_i$. Tento korelačný koeficient odhaľuje iba mieru lineárnej závislosti a nie je invariantný vzhľadom na rastúcu transformáciu náhodných premenných. Preto je vhodné vypočítať aj neparametrické miery korelácie- Kendalov alebo Spearmanov korelačný koeficient. V ďalšom označme $R(X_i) = R_i$ ($Q(Y_i) = Q_i$) poradie prvku X_i (Y_i) v usporiadanej vzostupnej postupnosti meraní $\{X_i\}_{i=1}^n$ (postupnosti $\{Y_i\}_{i=1}^n$). Ak sú vo výbere pozorovania rovnakých hodnôt napr. $X_i = X_j$, potom ich poradia sú tiež rovnaké čísla, a síce priemer ich poradí. Spearmanov korelačný koeficient vyrátame zo vzorca:

$$r_s = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^{n}(R_i - Q_i)^2 \tag{3}$$

Poznámka 1.

V prípade väčšieho počtu rovnakých hodnôt sa odporúča použiť modifikovaný vzorec pre Spearmanov korelačný koeficient , pozri [2].

Dvojicu meraní $\{X_i, Y_i\}$ a $\{X_j, Y_j\}$ nazveme súhlasnou (concordant), ak $(X_i - X_j)(Y_i - Y_j) > 0$ a nesúhlasnou (discordant), ak platí opačná ostrá nerovnosť.

Označme n_c počet súhlasných, a n_d počet nesúhlasných dvojíc, n_x nech je počet dvojíc, kde $X_i = X_j$ a n_y počet dvojíc, kde $Y_i = Y_j$. Potom Kendalov korelačný koeficient vypočítame zo vzorca:

$$r_k = \frac{n_c - n_d}{\sqrt{(n_c + n_d + n_x)(n_c + n_d + n_y)}} \quad . \tag{4}$$

Poznámka 2.
V systéme Mathematica sú pre výpočet Spearmanovho a Kendalovho korelačného koeficienta implementované príkazy: *SpearmanRankCorrelation[výberX, výberY]* a *KendallRankCorrelation[výberX, výberY]*.

2. Metóda vážených priemerov

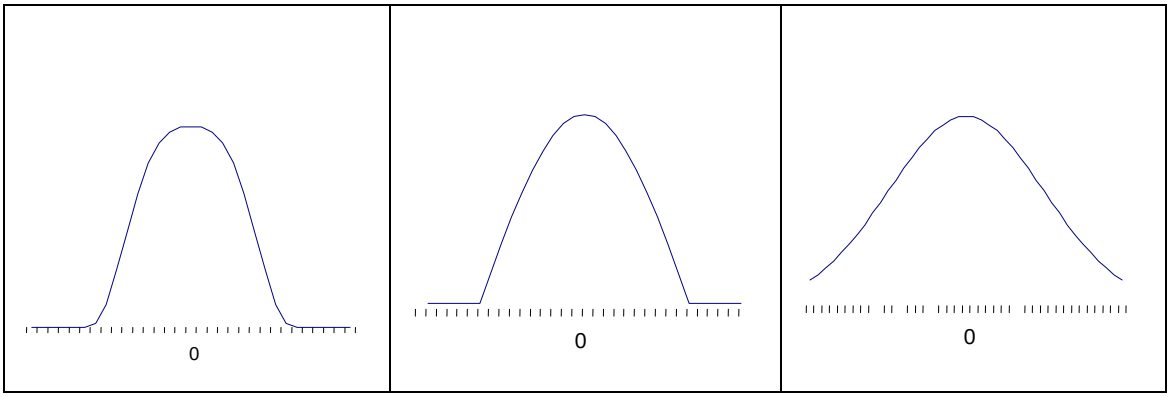
Metóda vážených priemerov je jednou z pomerne jednoduchých metód neparametrickej regresie. Základnou myšlienkou tejto metódy je myšlienka, že najväčšiu informáciu o pozorovaní Y_0, ktoré prislúcha hodnote X_0 možno získať predovšetkým z tých hodnôt Y_i, ktorých príslušná hodnota X_i je veľmi blízka hodnote X_0. Preto by takým hodnotám vo váženom priemere mala prislúchať najväčšia váha. Označme $z_i = \frac{X_i - X_0}{h}$, kde *h* je dĺžka intervalu (šírka okna, bandwith), z ktorého budeme vyberať hodnoty X_i. Dĺžka okna sa volí zvyčajne dvoma spôsobmi, buď je fixná alebo sa mení. Ak sa mení, tak je to interval, ktorý obsahuje k (vopred volené číslo) najbližších hodnôt X_i okolo X_0. Aby sme hodnotám z_i priradili príslušnú váhu, zadefinujeme *jadrovú funkciu* $K(z)$, ktorá je symetrická okolo nuly, s rastúcim $|z|$ klesá a pre $|z| > 1$ nadobúda hodnotu 0. Po výpočte váh $w_i = K(z_i)$ odhadneme $\hat{Y}_0$ v tvare:

$$\hat{Y}_0 = \hat{f}(X_0) = \frac{\sum_{i=1}^{n} w_i Y_i}{\sum_{i=1}^{n} w_i} \quad . \tag{6}$$

Najčastejšie používané jadrové funkcie sú uvedené v tabuľke 1. a tvar niektorých z nich je na obrázku 1. (v poradí: trikubická, Gaussova, kosínusová).

Tab. 1

jadrová funkcia	K(z), $\lvert z\rvert \le 1$
trojuholníková	$1 - \lvert z\rvert$
trikubická	$\left(1 - \lvert z\rvert^3\right)^3$
Gaussova	$\frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2} z^2)$
kosínusová	$\frac{\pi}{4} \cos(\frac{\pi}{2} z)$



Obr. 1.

3. Metóda založená na poradiach meraní

Táto metóda umožňuje odhadnúť (za určitých podmienok) $\hat{Y}_i$ pre meranie X_i a tiež $\hat{X}_i$ pre príslušné meranie Y_i. Metóda je založená na poradiach meraní a k výpočtom odhadov využíva klasickú regresnú priamku preloženú bodmi $\{R_i, Q_i\}$. Postupujeme v uvedenom poradí: vypočítame regresnú priamku poradí $y = a + bx$, kde

$$a = \frac{(1-b)(n+1)}{2}, \qquad b = \frac{\sum_{i=1}^{n} R_i Q_i - \frac{n(n+1)^2}{4}}{\sum_{i=1}^{n} R_i^2 - \frac{n(n+1)^2}{4}} \quad . \tag{6}$$

Vypočítame poradia odhadnuté pomocou priamky poradí pre každé X_i a Y_i. T. j. pre každé poradie R_i vypočítame $\hat{Q}_i = \hat{Q}(Y_i) = a + bR_i$ a pre Q_i zase $\hat{R}_i = \hat{R}(X_i) = \frac{Q_i - a}{b}$. Je zrejmé, že takto získané poradia nemusia byť nutne celými číslami. Navyše v hodnotách, pre ktoré $\hat{R}_i$ je menšie ako 1 a väčšie ako n, sa hodnoty $\hat{X}_i$ nedajú vypočítať. Pomocou každého odhadnutého poradia zrátame $\hat{X}_i$ a $\hat{Y}_i$. Ak $\hat{R}_i$ zodpovedá niektorému poradiu prvku X_j, potom $\hat{X}_i = X_j$. Ak $\hat{R}_i$ je hodnota medzi poradiami prvkov $X_j < X_k$ potom $\hat{X}_i$ vypočítame v tvare:

$$\hat{X}_i = X_j + \frac{\hat{R}_i - R_j}{R_k - R_j}(X_k - X_j). \tag{7}$$

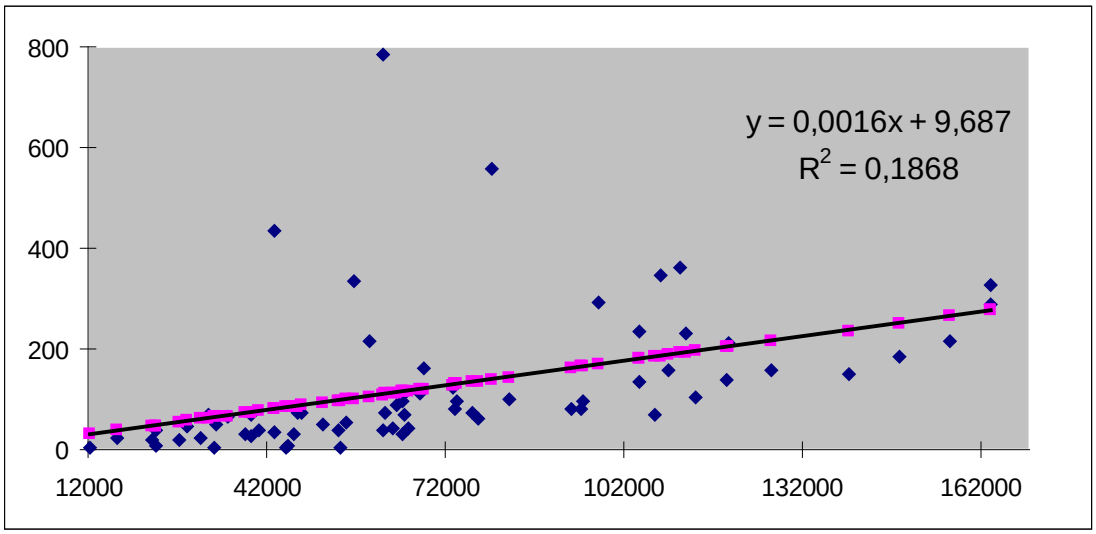
Podobne ak $\hat{Q}_i$ je poradím niektorého prvku Y_j, potom $\hat{Y}_i = Y_j$, inak $\hat{Q}_i$ bude hodnotou medzi poradiam prvku $Y_j < Y_k$ a $\hat{Y}_i$ vypočítame v tvare:

$$\hat{Y}_i = \hat{Y}_j + \frac{\hat{Q}_i - Q_j}{Q_k - Q_j}(Y_k - Y_j) \quad . \tag{8}$$

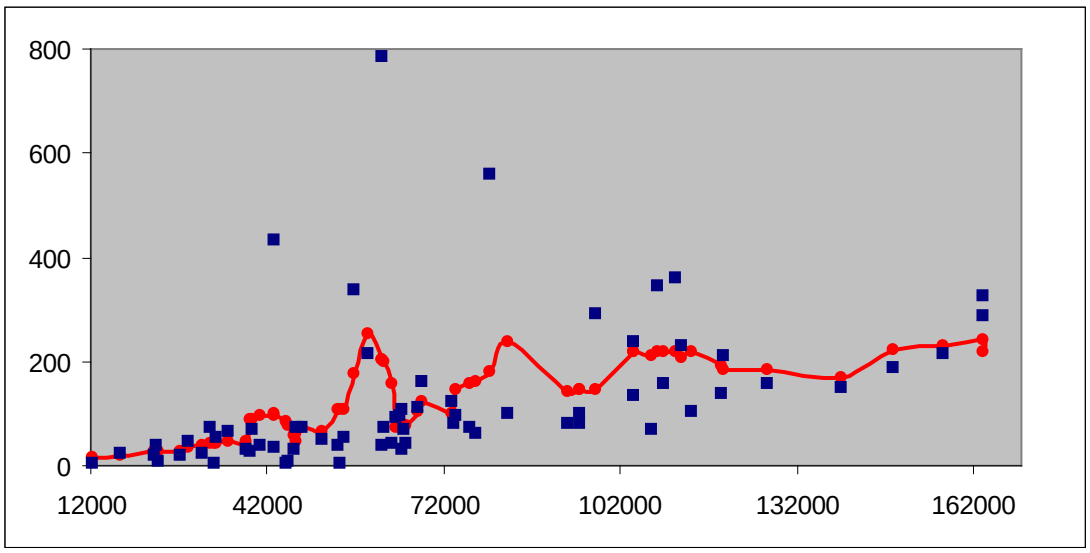
Posledným krokom tejto procedúry je vykreslenie lomenej čiary, prechádzajúcej bodmi $(X_i, \hat{Y}_i)$ a $(\hat{X}_i, Y_i)$.

4. Aplikácia metód v praxi

Obidve popísané metódy aplikujeme na súbore dvojíc dát (pozri [4]), ktoré popisujú počty zdravotníckych pracovníkov Y_i vzhľadom na počet obyvateľstva X_i v 68 okresoch Slovenska. Korelačný koeficient je r_{xy} =0,432. Spearmanov a Kendalov korelačný koeficient sú vyššie čísla, r_s =0,735 a r_k =0,546. Dvojice dát spolu s regresnou priamkou, tvarom regresnej priamky a koeficientom determinácie (R^2) sú na obrázku 2. Dvojice dát a krivka vyhladená metódou vážených priemerov je na obrázku 3. Kvôli porovnaniu s ostatnými metódami ešte spočítame strednú kvadratickú chybu odhadu MSE(1), MSE(1)=15 511,13.



Obr.2.



Obr. 3.

Spočítame teraz vážené priemery hodnôt (trikubická jadrová funkcia) pri pohyblivej dĺžke okna, obyčajné aritmetické priemery vždy 7 po sebe idúcich hodnôt

a dvojice dát $(X_i, \hat{Y}_i)$ a $(\hat{X}_i, Y_i)$. Zrátame tiež stredné kvadratické chyby odhadov MSE(2), MSE(3) a MSE(4), pričom v poslednom prípade budeme uvažovať pri výpočte MSE iba $(X_i, \hat{Y}_i)$. Výsledky sú opäť uvedené v tabuľke 2.

Tab. 2.

MSE(1)	MSE(2)	MSE(3)	MSE(4)
15511,13	13194,21	14682,18	16772,03

V danom prípade sa ukázala najvhodnejšou metóda vážených priemerov.

5. Záver

V článku boli popísané dve metódy neparametrickej regresnej analýzy, metóda vážených priemerov a metóda založená na poradiach meraní. Na konkrétnom aplikačnom príklade bolo ukázané, že tieto metódy sa za istých okolností dávajú lepšie výsledky ako použitie klasickej regresnej analýzy.

Poďakovanie: Tento článok vznikol za podpory grantu 1/3014/06.

Zoznam použitej literatúry

[1] Anděl, J.: Matematická statistika, SNTL/ALFA Praha (1985)
[2] Conover, W., J.: Practical nonparametric statistics, John Wiley & sons, (999)
[3] Fox, J.: Introduction to nonparametric regression, http://socserv.mcmaster.ca
[4] Health statistics yearbook of the Slovak republic (2004), Ústav zdravotníckych informácií a štatistiky, http://www.uzis.sk

Autor: Jana Kalická, Stavebná fakulta STU, Radlinského 11, 813 068 Bratislava

Autocorrelated Errors of Least Weighted Squares

Jan Kalina
KPMS MFF UK
Sokolovská 83
186 75 Praha 8
Czech Republic

This paper modifies the Durbin-Watson test for weighted regression and applies the result to least weighted squares, which is one of robust regression methods with a high breakdown point. Such test conditioning on the weights is exact, in contrast to the asymptotic test of Kalina (2004). This work is supported by the grant 402/06/408 (Robustification of the generalized method of moments) of the Grant Agency of the Czech Republic.

1 Durbin-Watson Test for Least Squares

Throughout this paper the regression model

$$Y_t = \beta_1 X_{t1} + \cdots + \beta_p X_{tp} + e_t, \quad t = 1, \ldots, n \tag{1}$$

or in the matrix notation $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$ is assumed to be observed as a time series in equidistant time intervals. An intercept is not required in the model although it may be present. Later we also need the following notation. Let $\boldsymbol{\mathcal{I}}_n$ denote the unit matrix of size $n \times n$, let matrix $\mathbf{M}$ be defined by $\mathbf{M} = \boldsymbol{\mathcal{I}}_n - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ and matrix $\mathbf{A}$ of size $n \times n$ by

$$\mathbf{A} = \begin{pmatrix} 1 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 1 \end{pmatrix}.$$

Durbin and Watson (1950, 1951) proposed the test of autocorrelation of errors for model (1). The one-sided test considers the null hypothesis of independent errors $\mathbf{e}$ against the alternative of positive autocorrelation of the first order. Denoting the vector of residuals of the least squares regression by $\mathbf{u} = (u_1, \ldots, u_n)^T$, the one-sided test rejects H_0 for small values of the test statistic, which can be expressed as

$$d = \frac{\sum_{t=2}^{n}(u_t - u_{t-1})^2}{\sum_{t=1}^{n} u_t^2} = \frac{\mathbf{u}^T\mathbf{A}\mathbf{u}}{\mathbf{u}^T\mathbf{u}} = \frac{\mathbf{e}^T\mathbf{M}\mathbf{A}\mathbf{M}\mathbf{e}}{\mathbf{e}^T\mathbf{M}\mathbf{e}}. \tag{2}$$

The distribution of (2) depends on $\mathbf{X}$, so Durbin and Watson consider the scelet decomposition of $\mathbf{M}$. This allows to use the theorem of Poincaré to find such lower and upper bounds for the critical value which do not depend on $\mathbf{X}$; their result is however valid only when an intercept is present in the model (1).

2 Durbin-Watson Test for Weighted Regression

After introducing the notation we modify the classical Durbin-Watson test for the context of weighted regression. Kalina (2004) also mentions the weighted regression case to show the contrary to least weighted squares. Here we get further and express the distribution of the test statistics for different situations always as a ratio of two quadratic forms.

We assume the model (1), where $\mathbf{X}$ is a matrix of (non-random) constants of size $n \times p$ with $\mathsf{rank}(\mathbf{X}) = p$ with $p < n$. The vector of random errors follows the multivariate normal distribution

$$\mathbf{e} \sim \mathsf{N}\left(\mathbf{0}, \sigma^2\mathbf{W}^{-1}\right) \tag{3}$$

where $0 < \sigma^2 < \infty$ and the diagonal matrix $\mathbf{W} = \mathsf{diag}\{w_1, \ldots, w_n\}$ contains fixed positive weights $w_1, \ldots, w_n$. Less reliable observations are assumed to have errors with a large variance, which means a large value of the corresponding diagonal element of $\mathbf{W}^{-1}$ and a small value of the weight. The least squares estimator is a special case with equal weights.

Let $\mathbb{R}$ denote the set of all real numbers. The weighted least squares estimator is defined as

$$\mathbf{b}_W = \arg \min_{(b_1,\ldots,b_p)\in\mathbb{R}^p} \sum_{t=1}^{n} w_t(Y_t - b_1X_{t1} - \cdots - b_pX_{tp})^2$$

and is equal to $\mathbf{b}_W = (\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W}\mathbf{Y}$. Further we introduce the notation $\mathbf{X}^* = \mathbf{W}^{1/2}\mathbf{X}$ and define matrices $\mathbf{M}_W$ and $\mathbf{M}^*$ by

$$\mathbf{M}_W = \boldsymbol{\mathcal{I}}_n - \mathbf{X}(\mathbf{X}^T\mathbf{W}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{W} \quad \text{and} \quad \mathbf{M}^* = \boldsymbol{\mathcal{I}}_n - \mathbf{X}^*(\mathbf{X}^{*T}\mathbf{X}^*)^{-1}\mathbf{X}^{*T}.$$

Residuals of the weighted regression equal

$$\mathbf{u} = \mathbf{Y} - \mathbf{X}\mathbf{b}_W = \mathbf{M}_W\mathbf{Y} = \mathbf{M}_W(\mathbf{X}\boldsymbol{\beta} + \mathbf{e}) = \mathbf{M}_W\mathbf{e}.$$

The matrix $\mathbf{M}_W$ is symmetric only in special cases. Next we formulate some properties of various matrices which will be used later. These can be verified easily; we refer to Rao (1973) for basics of the matrix theory.

Lemma 1. *Using the notation of above, it holds that*

- $\mathbf{M}_W\mathbf{M}_W = \mathbf{M}_W$*;*
- $\mathbf{M}_W^T\mathbf{W}\mathbf{M}_W = \mathbf{W}^{1/2}\mathbf{M}^*\mathbf{W}^{1/2}$*;*
- *the matrices* $\mathbf{M}_W^T\mathbf{M}_W$*,* $\mathbf{M}_W^T\mathbf{W}\mathbf{M}_W$*,* $\mathbf{M}_W^T\mathbf{A}\mathbf{M}_W$*,* $\mathbf{M}_W^T\mathbf{W}^{1/2}\mathbf{A}\mathbf{W}^{1/2}\mathbf{M}_W$ *and* $\mathbf{M}^*$ *are symmetric and positive semidefinite of rank* $n - p$*.*

For weighted regression we now modify the Durbin-Watson test under different assumptions on the errors $\mathbf{e}$. Firstly, under (3) it is reasonable to compute the Durbin-Watson statistic with weighted residuals of the weighted regression $\sqrt{w_1}u_1, \ldots, \sqrt{w_n}u_n$. Then the test statistic has the form

$$\frac{\sum_{t=2}^n (\sqrt{w_t}u_t - \sqrt{w_{t-1}}u_{t-1})^2}{\sum_{t=1}^n w_t u_t^2} = \frac{\mathbf{u}^T\mathbf{W}^{1/2}\mathbf{A}\mathbf{W}^{1/2}\mathbf{u}}{\mathbf{u}^T\mathbf{W}\mathbf{u}} \tag{4}$$

and its null distribution is described below in Theorem 1.

Theorem 1. *Let us denote positive eigenvalues of*

$$\mathbf{W}^{-1/2}\mathbf{M}_W^T\mathbf{W}^{1/2}\mathbf{A}\mathbf{W}^{1/2}\mathbf{M}_W\mathbf{W}^{-1/2}$$

by $\gamma_1, \ldots, \gamma_{n-p}$ *and positive eigenvalues of* $\mathbf{M}^*$ *by* $\lambda_1, \ldots, \lambda_{n-p}$*. Under the assumption (3), the equality in distribution*

$$\frac{\mathbf{u}^T\mathbf{W}^{1/2}\mathbf{A}\mathbf{W}^{1/2}\mathbf{u}}{\mathbf{u}^T\mathbf{W}\mathbf{u}} \overset{\mathcal{D}}{=} \frac{\sum_{i=1}^{n-p}\gamma_i E_i^2}{\sum_{i=1}^{n-p}\lambda_i E_i^2}$$

is true for the test statistic (4), where $E_1, \ldots, E_{n-p}$ *are independent random variables with* $\mathsf{N}(0,1)$ *distribution.*

Proof: applying Lemma 1, we give a sketch of the proof only for the denominator. Starting with

$$\mathbf{u}^T\mathbf{W}\mathbf{u} = \mathbf{e}^T\mathbf{M}_W^T\mathbf{W}\mathbf{M}_W\mathbf{e},$$

the scelet decomposition gives $\mathbf{M}_W^T\mathbf{W}\mathbf{M}_W = \mathbf{S}^T\mathbf{\Lambda}\mathbf{S}$, where $\mathbf{S}$ is an orthonormal matrix of size $n \times n$ and $\mathbf{\Lambda}$ is a diagonal matrix $\mathbf{\Lambda} = \mathsf{diag}\{\lambda_1, \ldots, \lambda_n\}$ containing eigenvalues of $\mathbf{M}_W^T\mathbf{W}\mathbf{M}_W$ on the diagonal. Let $\mathbf{t}$ be the random vector $\mathbf{t} = \mathbf{S}\mathbf{e}$, which follows the normal distribution $\mathbf{t} \sim \mathsf{N}(\mathbf{0}, \mathbf{S}\mathbf{S}^T) = \mathsf{N}(\mathbf{0}, \sigma^2\boldsymbol{\mathcal{I}}_n)$. Therefore under H_0

$$\mathbf{u}^T\mathbf{W}\mathbf{u} \stackrel{\mathcal{D}}{=} \mathbf{e}^T\mathbf{S}^T\mathbf{\Lambda}\mathbf{S}\mathbf{e} = \mathbf{t}^T\mathbf{\Lambda}\mathbf{t} = \sum_{i=1}^{n}\lambda_i t_i^2.$$

Only $n - p$ of the summands are non-zero and the scale-invariance of (4) allows us to put $E_i = t_i/\sigma$ for $i = 1, \ldots, n-p$. Similar reasoning gives the expression for the numerator.

Simulating $E_1, \ldots, E_n$ approximates the null distribution of the test statistic and the critical value depends on the weights and also the design matrix $\mathbf{X}$. While this computation requires to obtain the inverse of a possibly large matrix, software computes this using the same QR-decomposition which is used for computing estimators of linear regression parameters.

The following Theorem 2 assumes homoscedastic errors. It is an analogy of Theorem 1 for the test statistic not downweighting the residuals.

Theorem 2. *Let us denote positive eigenvalues of* $\mathbf{M}_W^T\mathbf{A}\mathbf{M}_W$ *by* $\gamma_1, \ldots, \gamma_{n-p}$ *and positive eigenvalues of* $\mathbf{M}_W^T\mathbf{M}_W$ *by* $\lambda_1, \ldots, \lambda_{n-p}$. *Assuming*

$$\mathbf{e} \sim \mathsf{N}\left(\mathbf{0}, \sigma^2\boldsymbol{\mathcal{I}}_n\right)$$

the equality in distribution

$$\frac{\mathbf{u}^T\mathbf{A}\mathbf{u}}{\mathbf{u}^T\mathbf{u}} \stackrel{\mathcal{D}}{=} \frac{\sum_{i=1}^{n-p}\gamma_i E_i^2}{\sum_{i=1}^{n-p}\lambda_i E_i^2}$$

is true for the test statistic (2), where $E_1, \ldots, E_{n-p}$ *are independent random variables with* $\mathsf{N}(0,1)$ *distribution.*

3 Durbin-Watson Test for Least Weighted Squares

Víšek (2001*a*) proposed the least weighted squares (LWS) estimator as a robust regression method with a high breakdown point. In the model (1) let

$$u_i(\mathbf{b}) = Y_i - b_1 X_{i1} - b_2 X_{i2} - \cdots - b_p X_{ip}, \quad i = 1, \ldots, n$$

denote the residual corresponding to the i-th observation for a given estimate $\mathbf{b} = (b_1, \ldots, b_p)^T \in \mathbb{R}^p$ of the parameter $\boldsymbol{\beta}$. Let us order squared residuals

$$u^2_{(1)}(\mathbf{b}) \leq u^2_{(2)}(\mathbf{b}) \leq \cdots \leq u^2_{(n)}(\mathbf{b}).$$

The *least weighted squares* (LWS) estimator of $\boldsymbol{\beta}$ is defined by

$$\mathbf{b}_{LWS} = \arg\min_{\mathbf{b}\in\mathbb{R}^p} \sum_{k=1}^{n} w_k u^2_{(k)}(\mathbf{b}).$$

Non-negative weights $w_1, w_2, \ldots, w_n$ must be specified at first. However only their magnitudes are known a priori and the weights are assigned to the data after a permutation, which is determined automatically only during the computation based on the residuals. The least trimmed squares (LTS) represent a special case of least weighted squares with weights equal to zero or one only.

Kalina (2007) describes an algorithm giving a tight approximation to the true value of the estimate and studies computational aspects of the estimator. The paper also lists different references on theoretical properties of the estimator.

Víšek (2001*b*) studies the Durbin-Watson test statistic computed from residuals of the least trimmed squares regression. The paper proves its asymptotic equivalence with the test statistic computed from residuals of the least squares fit. While the residuals corresponding to outliers are also included in the test statistic, the least weighted squares would have the advantage that they downweight outlying observations. Kalina (2004) examines the Durbin-Watson test for least weighted squares and proves the test statistic to be asymptotically equivalent to the classical statistic in the least squares context.

The exact Durbin-Watson test for least weighted squares conditioning on fixed weights corresponds to the test for weighted regression in Chapter 2. It is a parametric test assuming the distribution of random errors. Now we argue in favour of the approach of Theorem 1.

For normal data Čížek (2006) presents a method for the optimal choice of weights for the LWS estimator. The weights are chosen which estimate the variability of each particular observation. Such estimator is asymptotically efficient and asymptotically equivalent to the least squares estimator without weighting also when the weights are very different from equal ones. The assumption (3) is therefore natural for the Durbin-Watson test.

For data with outliers the downweighting of residuals is actually the motivation for introducing the LWS estimator and the assumption (3) is again justified. The next theorem describes the test.

Theorem 3. *The p-value of the Durbin-Watson test for least weighted squares against the one-sided alternative of positive autocorrelation based on (4) assuming (3) is equal to the probability*

$$P\left[\frac{\sum_{i=1}^{n-p}\gamma_i E_i^2}{\sum_{i=1}^{n-p}\lambda_i E_i^2} \leq \frac{\mathbf{u}^T\mathbf{W}^{1/2}\mathbf{A}\mathbf{W}^{1/2}\mathbf{u}}{\mathbf{u}^T\mathbf{W}\mathbf{u}}\right]$$

with $E_1, \ldots, E_{n-p}, \gamma_1, \ldots, \gamma_{n-p}$ *and* $\lambda_1, \ldots, \lambda_{n-p}$ *defined in Theorem 1.*

References

[1] Čížek P. (2006): Efficient robust estimation of regression models. *Tilburg University Discussion Paper* No. 08/2006, Tilburg.

[2] Durbin J., Watson G.S. (1950, 1951): Testing for serial correlation in least squares regression I, II. *Biometrika* **37, 38**, 409 – 428, 159 – 178.

[3] Kalina J. (2004): Durbin-Watson test for least weighted squares. In Antoch J. (ed.): *COMPSTAT 2004, Proceedings in computational statistics,* Physica-Verlag, Heidelberg, 1287 – 1294.

[4] Kalina J. (2007): Robust methods for template matching. *Folia Fac. Sci. Nat. Univ. Masaryk. Brunensis, Mathematica*, 9 pp. In print.

[5] Rao C.R. (1973): Linear methods of statistical induction and their applications. Wiley, New York. Second edition.

[6] Víšek J.Á. (2001*a*): Regression with high breakdown point. *Proceedings of ROBUST 2000*, JČMF and Czech statistical society, Prague, 324 – 356.

[7] Víšek J.Á. (2001*b*): Durbin-Watson test statistic for the least trimmed squares. *Bulletin of the Czech Econometric Society* No. 14, 1 – 40.

On a construction of joint distributions

Martin Kalina, Ahmed Mohammed, Oľga Nánásiová, Štefánia Václavíková
Dept. Mathematics, Slovak Univ. of Technology
Radlinského 11
sk-813 68 Bratislava
{*kalina, mohammed, olga, stefi*}*@math.sk*

Abstract

In this paper we construct the system of finite two-dimensional probability distributions (a convex set) by constructing of its vertices. A vertex we get in choosing a way of filling a table $n \times m$, where at each position (each cell) we put always the maximal possible value for a given pair of marginal distributions. Finally, we make an estimation of the number of vertices.

The work on this paper was inspired mainly by [2]. In this paper we want to show a possibility, how we can construct joint distributions when given marginal ones. We will work with finite random variables. The system of joint distributions is a convex set. We will construct the 'vertices' and their convex combinations will generate further joint distributions. To each vertex we may compute the correlation of the corresponding random vector. When choosing a particular joint distribution, which is a convex combination of vertices, the correlation of the corresponding random vector can be computed as the convex combination of correlations of vertices. The basics on joint probability distributions the reader can find, e.g., in [9].

Another possibility how to construct joint distributions is first to get marginal (cumulative) distribution functions F and G, then to choose some copula $C : [0,1]^2 \to [0,1]$ and to construct the joint distribution function, $H : R^2 \to [0,1]$, using Sklar Theorem:

$H(x,y) = C(F(x), G(y))$

The copula C contains the whole information on the dependence of random variables. This approach was described, e.g., by P. Volauf in [10].

Construction

We are given random variables X and Y, achieving n and m different values with non-zero probability, respectively. By a *vertex* we will understand a joint probability distribution which we get in the following way:

- We have to fill in a table $n \times m$. We choose an element of this table, say (i,j), and we put $P_{ij} = \min\{P(X_i),\, P(Y_j)\}$. For $1 \le k \le n$ and $1 \le t \le m$ let us denote

$$P_1(X_k) = \begin{cases} P(X_k), & \text{if } k \neq i \\ P(X_i) - P_{ij}, & \text{if } k = i \end{cases} \qquad P_1(Y_t) = \begin{cases} P(Y_t), & \text{if } t \neq j \\ P(Y_j) - P_{ij}, & \text{if } t = j \end{cases}$$

At least one of the values, $P_1(X_i)$ or $P_1(Y_j)$, is zero. The rest of the i-th row and/or j-th column, corresponding to which of the two values is zero, we fill in with zeros.

- In the second step we choose an element of the rest of this table, (i_2, j_2), and we put $P_{i_2 j_2} = \min\{P_1(X_{i_2}), P_1(Y_{j_2})\}$. For $1 \leq k \leq n$ and $1 \leq t \leq m$ let us denote

$$P_2(X_k) = \begin{cases} P_1(X_k), & \text{if } k \neq i_2 \\ P_1(X_{i_2}) - P_{i_2 j_2}, & \text{if } k = i_2 \end{cases}$$

$$P_2(Y_t) = \begin{cases} P_1(Y_t), & \text{if } t \neq j_2 \\ P_1(Y_{j_2}) - P_{i_2 j_2}, & \text{if } t = j_2 \end{cases}$$

At least one of the values, $P_2(X_{i_2})$ or $P_2(Y_{j_2})$, is zero. The rest of the i_2-th row and/or j_2-th column, corresponding to which of the two values is zero, we fill in with zeros.

- Repeating these steps we construct the joint probability distribution, which is a vertex in the convex set of joint probability distributions.

We illustrate the construction by the following example:

Example 1 The construction of a vertex:

- First step – we choose the cell $(3, 2)$

X^Y	1	2	3	4	$P(X_i)$
1		0			0.2
2		0			0.3
3		0.2			0.5
$P(Y_j)$	0.1	0.2	0.4	0.3	

Table 1: First step of the construction

- Second step – cell $(3, 4)$

X^Y	1	2	3	4	$P(X_i)$
1		0		0	0.2
2		0		0	0.3
3	0	0.2	0	0.3	0.5
$P(Y_j)$	0.1	0.2	0.4	0.3	

Table 2: Second step of the construction

- Third step – cell $(1,1)$

X^Y	1	2	3	4	$P(X_i)$
1	0.1	0		0	0.2
2	0	0		0	0.3
3	0	0.2	0	0.3	0.5
$P(Y_j)$	0.1	0.2	0.4	0.3	

Table 3: Third step of the construction

- The last step – we fill in the rest of the table, since this is already given uniquely

X^Y	1	2	3	4	$P(X_i)$
1	0.1	0	0.1	0	0.2
2	0	0	0.3	0	0.3
3	0	0.2	0	0.3	0.5
$P(Y_j)$	0.1	0.2	0.4	0.3	

Table 4: Last step of the construction

Yet we compute the correlation of random variables X and Y:

$$EX = 2.3,\ EY = 2.9,\ DX = 0.61,\ DY = 0.89,\ E(XY) = 7 \quad \Rightarrow \quad \rho \doteq 0.447 \qquad (1)$$

□

- If we want to construct the joint probability distribution with maximal joint distribution function, we have to choose the cell $(1,1)$ as the starting one and then to proceed in the 'right-down' direction, or to start in the (n,m) cell and to proceed in the 'left-up' direction
- In case we want to construct the joint probability distribution with the minimal joint distribution function, the starting cell should be $(1,m)$ and the direction 'left-down', or the starting cell $(n,1)$ and the direction 'right-up'.

In the next example we use random vectors (X,Y) with marginal distributions equal to those from Example 1. We construct the joint probability distributions with maximal and minimal joint distribution functions, respectively. The correlation coefficients will be in these cases also the maximal and minimal possible ones by fixed marginal distributions.

X^Y	1	2	3	4	$P(X_i)$
1	0.1	0.1	0	0	0.2
2	0	0.1	0.2	0	0.3
3	0	0	0.2	0.3	0.5
$P(Y_j)$	0.1	0.2	0.4	0.3	

Table 5: The joint probability distribution with maximal distribution function

Example 2 Table 5 contains the joint probability distribution with maximal distribution function for our fixed marginal distributions. We may construct it following the path $(1,1)$, $(1,2)$, $(2,2)$, $(2,3)$, $(3,3)$, $(3,4)$.

The correlation coefficient we can compute using the means and variances of X and Y from formula (1). $E(XY) = 7.5$, hence

$$\rho(X,Y) \doteq 0.999.$$

Table 6 contains the joint probability distribution with minimal distribution function for our fixed marginal distributions. We may construct it following the path $(3,1)$, $(3,2)$, $(3,3)$, $(2,3)$, $(2,4)$, $(1,4)$.

X^Y	1	2	3	4	$P(X_i)$
1	0	0	0	0.2	0.2
2	0	0	0.2	0.1	0.3
3	0.1	0.2	0.2	0	0.5
$P(Y_j)$	0.1	0.2	0.4	0.3	

Table 6: The joint probability distribution with maximal distribution function

$$E(XY) = 6.1 \quad \Rightarrow \quad \rho(X,Y) \doteq -0.909.$$

Concerning the number of vertices we have the following result:

Theorem 1 *We are given random variables X and Y, achieving n and m different values with non-zero probability, respectively. Assume $n \geq m$. The number of vertices, $V_{n,m}$, is bounded by:*

$$\frac{n!}{(n-m+1)!} \leq V_{n,m} \leq m^{n-m} \cdot n! \cdot (m!)^2 \cdot \frac{(m-1)1}{4}$$

In case $n = m$ and the marginal probability distributions are uniform, we get

$$V_{n,n} = n!$$

Some remarks to the space of random variables

Let us denote by $\tilde{\mathbf{\Omega}}$ the space of all random variables defined on Ω.

The covariance of random variables $X, Y \in \tilde{\mathbf{\Omega}}$ is, in fact, their dot-product when considering the space $\tilde{\mathbf{\Omega}}$ as a vector space. From this point of view the correlation of random variables X and Y, $\rho(X,Y)$ is the cosine of the angle between X and Y and if $\rho(X,Y) = 0$, then the random variables X and Y are orthogonal to each other.

However, also another orthogonality is possible in the space $\tilde{\mathbf{\Omega}}$. Namely – something like a set-complementarity. This may be expressed in several ways. The following is obvious:

$$P(Y = 0|X \neq 0) = 1 \quad \Leftrightarrow \quad P(\{\omega \in \Omega; X(\omega) \cdot Y(\omega) = 0\}) = 1$$

This orthogonality plays an important role in defining of conditional probability on special algebraic structures, e.g. on so-called *MV-algebras* (many-valued algebras), see [3] (but also on other algebraic structures, see [5]). These are important when handling with vague data.

Another generalization of Boolean algebras then MV-algebras, is so-called *ortho-modular lattice*, OML for brevity (for details on OMLs see [8]). A possible model of an OML is the union of certain system of Boolean algebras. If we deal with general algebraic structures, a strange situation (from the point of view of Kolmogorovian model) may occur (see [4, 5]), namely

$$\rho(X,Y) = 0 \not\Leftrightarrow \rho(Y,X) = 0$$

How this happens? Coming back to the construction of a joint distribution on a Boolean algebra, we must assign a probability to each pair $(X = x, Y = y)$ and hence

$$P(X = x, Y = y) = P(X = x \wedge Y = y) = P(Y = y, X = x)$$

and hence we get $\rho(X,Y) = \rho(Y,X)$.

When constructing a joint distribution on an OML (union of Boolean algebras), saying strictly, the joint distribution is given uniquely by marginal distributions. If we wish to compute the correlation of X, Y on OML, we must 'fill the whole table' similarly to the case when working on a Boolean algebra. However, the event $X = x \wedge Y = y$ is not an element of our OML, hence assigning some value $P(X = x, Y = y) = p_{xy}$ does not mean assigning the same value to $P(Y = y, X = x)$. The value $P(X = x, Y = y) = p_{xy}$ can be interpreted in such a way that first the event $X = x$ has occurred and then the event $Y = y$. This means that different values $p_{xy} \neq p_{yx}$ imply different interaction according to the time axes (see [1, 6, 7]).

Acknowledgment. This work was supported by Science and Technology Assistance

Agency under the contract No. APVV-0375-06, and by the VEGA grant agency, grant number 1/3014/06.

References

[1] M. Bohdalová, M. Minárová, O. Nánásiová, A note on algebraic approach to uncertainty, **Forum Statisticum Slovacum**, 3 (2006), 31-39.

[2] F. Durante, R. Mesiar, C. Sempi, On a family of copulas constructed from the diagonal section, **Soft Comp.** 10(6) (2006), 490-494.

[3] M. Kalina, O. Nánásiová, Conditional states and joint distributions on MV-algebras, **Kybernetika** 42 (2006), 129-142.

[4] A. Khrennikov, O. Nánásiová, Representation theorem of observables on a quantum logic, **Soft Computing**, (2006).

[5] O. Nánásiová, Map for simultanous measuremants for a quantum logic, **Int. J. of Theoretical Physics**, 42 (2003), 1889-1903.

[6] O. Nánásiová, Principle conditioning, **Int. J. of Theoretical Physics**, 43 (2004), 1383-1395.

[7] O. Nánásiová, M. Minárová, A. Mohammed, Measure of "symmetric difference", In: MAGIA 2006, STU Bratislava, 2006, 55-60.

[8] P. Pták, S. Pulmannová, Quantum Logics, Kluwer Acad. Press, Bratislava 1991.

[9] A. Rényi, Wahrscheinlichtkeitsrechnung mit einem Anhang über Informationstheorie, VEB Deutcher Verlag der Woissenschaften, Berlin 1962 (czech translation Teorie pravděpodobnosti, Academia, Praha 1972).

[10] P. Volauf, O asociácii náhodných veličín a kopuliach, **Forum Statisticum Slovacum** 3 (2005), 91-98. (in slovak)

Elektronická učebnica predmetu Štatistika*na EF

Pavol Kráľ, Alena Kaščáková, Gabriela Nedelová

Abstract
The main aim of the contribution is to present the electronic textbook Statistics which is intended for educational purposes at the Faculty of Economics Matej Bel University Banská Bystrica.
Keywords: e-learning, MOODLE, Statistics, LaTeX, pdfTeX, PDF

1. Úvod

Hlavným zámerom nášho článku je predstaviť prvotnú verziu multimediálnej učebnice štatistiky, ktorá vznikla v rámci riešenia projektu dištančné vzdelávanie ekonomických predmetov na EF. Pôvodným zámerom bola elektronická učebnica pokrývajúca potreby základného kurzu štatistiky pre všetky odbory EF UMB bez využitia LMS MOODLE. Tento zámer bol v priebehu riešenia projektu niekoľkokrát modifikovaný vzhľadom na rôzne skutočnosti. Najdôležitejšou bolo zlúčenie EF UMB s Fakultou financií UMB, ktoré viedlo k modifikácii učebných osnov jednotlivých predmetov základného kurzu štatistiky. Druhou skutočnosťou bolo vydanie vlastnej základnej literatúry pre predmety Štatistika 1 a Štatistika 2 (knihy [1, 2] pre odbory CR, FBI, EMP) a príprava klasickej printovej učebnice pre predmet Štatistika (odbor VES). Vzhľadom na tieto okolnosti bol pôvodný zámer modifikovaný nasledujúcim spôsobom. Elektronická učebnica bude nielen základnou študijnou literatúrou pre predmet Štatistika, ale mala by slúžiť aj ako pomôcka pre štatistické spracovanie kvalifikačných prác a výskumných projektov na EF UMB. Pripravené sú dve rôzne implementácie učebnice. Prvou je výučbové CD. Druhou je integrácia materiálov do kurzu predmetu Štatistika v LMS MOODLE.

2. Obsah učebnice

Výučbové CD obsahuje pdf súbor s textom učebnice, dátové súbory vo formátoch xls, sav (natívny formát SPSS) a txt, videosekvencie s postupom riešenia v SPSS, R a Exceli, inštalačné súbory programu R. Zaradenie programu R bolo motivované snahou umožniť aj externým študentom, ktorí nemajú prístup k programovému vybaveniu EF UMB, pracovať s plnohodnotným štatistickým softvérom. Online kurz v LMS MOODLE okrem

*Hoci presným názvom predmetu je Štatistika 1, používame pre jednoduchosť v celom článku názov Štatistika, aby sme zdôraznili odlišnosť tohto predmetu určeného pre odbor VES od predmetu rovnakého názvu, ktorý je určený pre odbory CR, EMP a FBI na EF UMB.

toho plne využíva možnosti systému Moodle, t.j. otvára napríklad možnosť konzultácie s vyučujúcim alebo dalšími študentami, obsahuje ďalšie testové úlohy a mnoho ďalších vecí. Elektronická učebnica obsahuje teoretické základy nasledujúcich celkov:

1. popisná štatistika
2. základy pravdepodobnosti
3. úvod do induktívnej štatistiky
4. regresia a korelácia
5. analýza kategoriálnych znakov
6. indexy.

To je v úplnom súlade s osnovou tohto predmetu:

1. Predmet a úlohy štatistiky. Základné pojmy a definície.
2. Charakteristiky úrovne. Priemery. Stredné hodnoty polohy. Kvantily.
3. Variabilita, miery variability. Rozklad rozptylu.
4. Meranie dynamiky ekonomických javov.
5. Pravdepodobnosť-základné pojmy. Náhodná premenná. Pravdepodobnostná funkcia, funkcia rozdelenia hustoty pravdepodobnosti, distribučná funkcia.
6. Druhy a technika výberov. Bodový a intervalový odhad.
7. Testovanie hypotéz.
8. Regresná analýza.
9. Korelačná analýza.
10. Analýza závislosti kvalitatívnych znakov

Jednotlivé časti obsahujú z nášho hľadiska nevyhnutné množstvo teoretických informácií, ktoré sú potrebné pre interpretáciu počítačových výstupov Excelu a štatistických programov SPSS a R. Tento obsah jednotlivých častí je v súlade s učebnými osnovami predmetu Štatistika pre odbor VES na EF UMB, ale obsahuje aj niektoré doplnkové časti, ktoré priamo nadväzujú na základnú osnovu predmetu a sú veľmi zaujímavé pre praktickú aplikáciu štatistiky napr. pri spracovaní kvalifikačných prác. Tieto časti sú: základy štatistického spracovania dotazníkov, sila testu, neparametrické testy, ANOVA, regresná diagnostika, analýza časových radov. Každý teoretický celok je ilustrovaný na niekoľkých vybraných problémoch, ktoré majú podľa možností ekonomický charakter. Každý problém obsahuje vzorové riešenie s použitím štatistického softvéru (ak je to možné, použijú sa všetky tri programi - SPSS, R a Excel). Kde sme to považovali za

vhodné z hľadiska spracovania vlastnej analýzy čitateľom, je vzorové riešenie doplnené o videosekvenciu ilustrujúcu ako pracovať s príslušnou procedúrou v nami použitom štatistickom softvéri. V závere jednotlivých celkov sa nachádzajú príklady a testové otázky na overenie dostatočného osvojenia prebraného učiva.

3. Technické detaily

Našou snahou pri technickom spracovaní učebnice bolo dosiahnuť jej platformovú nezávislosť pri zachovaní všetkých výhod elektronického formátu akou je napr. možnosť implementácie videosekvencií, preto sme zvolili formát pdf v kombinácii s elearningovým systémom Moodle. Prehliadače PDF (napríklad známy Adobe Reader) sú voľne prístupné pre všetky bežne používané platformy (Windows, Linux, Mac OS), čím je zaručená bezproblémová prenositeľnosť medzi systémami. Použitie systému Moodle samozrejme nie je ovplyvnené použitým operačným systémom a v podstate ani použitým webovým prehliadačom, pretože bez problémov funguje pri použití ľubovoľného z troch najrozšírenejších webových prehliadačov, či už bezplatných Opera a Mozilla Firefox alebo Internet Exploreru, ktorý je súčasťou operačných systémov Windows. Učebnica vo formáte pdf bola vytvorená s využitím typografického systému TEX (LATEX 2_ε a pdfTEX z distribúcie MiKTEX 2.5). Vzorové riešenia boli vytvorené prostredníctvom SPSS 13, R 2.4.1 a MS Excel 2003, videosekvencie boli vytvorené programom Camtasia Studio 2, pričom tieto sekvencie sú vo formáte avi aj ako animované gif. Dátové súbory sú vo formáte xls, sav a txt. Pdf súbory samozrejme obsahujú hypertextové odkazy, popup okienka, obrázky, animácie a mnohé ďalšie prvky slúžiace na uľahčenie osvojenia predkladanej problematiky. Pri tvorbe pdf súborov v systémoch LATEX 2_ε a pdfTEX boli použité napríklad balíčky hyperref, fancytooltips, tikzpicture a voľne prístupný balík súborov AcroTEX education bundle. Na tvorbu obrázkov boli použité okrem balíka tikzpicture aj programy CorelDraw 11, Gnuplot, Inkscape.

4. Záver

V našom príspevku stručne predstavujeme elektronickú učebnicu Štatistika, ktorá je určená najmä pre študentov odboru VES na EF UMB v Banskej Bystrici. Z hľadiska technického spracovania sa značne odlišuje od ďalších elektronických materiálov vytvorených na EF UMB s použitím programov Zoner a Lectora (pozri [3]). Hlavnou odlišnosťou je použitie formátu PDF, ktorý je štandardom pre elektronické publikovanie. Aktuálnu podobu tejto učebnice nájdete na `http://194.160.44.57/` v sekcii projekty KKMI. Uvedená elektronická učebnica by sa mala začať prakticky používať v zimnom semestri školského roku 2007/2008. Na základe praktických skúseností z jej používania predpokladáme jej ďalšie zdokonaľovanie a doplňovanie, pričom by v budúcnosti mohla byť voľne prístupná v MOODLE EF UMB. Okrem toho na túto elektronickú učebnicu by mala nadviazať učebnica viacrozmerných štatistických metód so zameraním na ekonomické aplikácie, ktorá je hlavným cieľom riešenia projektu KEGA Platformovo nezávislá voľne prístupná elektronická učebnica: Viacrozmerné štatistické metódy so zameraním na riešenie problémov ekonomickej praxe.

Poďakovanie:
Autori článku boli podporení fakultným grantom Dištančné vzdelávanie štatistických predmetov na EF UMB v Banskej Bystrici.

Literatúra:

[1] Kanderová, M., Úradníček, V.: Štatistika a pravdepodonosť pre ekonómov, 1. časť, OZ Financ, Banská Bystrica 2005. ISBN 80-968702-9-7.

[2] Kanderová, M., Úradníček, V.: Štatistika a pravdepodonosť pre ekonómov, 2. časť, OZ Financ, Banská Bystrica 2007. ISBN 978-80-969535-1-6.

[3] Kráľ, P., Balciar A.: Dištančné vzdelávanie štatistických predmetov na EF UMB, Forum Statisticum Slovacum 3, 2006, str. 185–187. ISSN 1336-7420.

Adresa autorov:

Mgr. Pavol Kráľ, PhD., Ing. Alena Kaščáková, PhD., RNDr. Gabriela Nedelová, PhD.
Katedra kvantitatívnych metód a informatiky
Ekonomická fakulta, Univerzita Mateja Bela
Tajovského 10, 975 90 Banská Bystrica
e-mail: pavol.kral@umb.sk, alena.kascakova@umb.sk, gabriela.nedelova@umb.sk

Automatic self assesment and evaluation in teaching statistics using the system Moodle

Zuzana Krivá

Abstract: Several years ago we have successfully implemented the course managements system Moodle (completely free to use) into educational process of several study programs at the Faculty of Civil Engineering of the Slovak Technical University in Bratislava. We will shortly mention advantages which use of this system has brought to us. The paper concerns on describing one of them: automatic self assessment and evaluation of students. As a suitable subject to demonstrate this feature we have chosen the subject *Informatics -Elementary statistics in Excel* of the study program Environmental Engineering.

Key words: the course management system, on line courses, randomization of quizzes and tests, evaluation, quiz quesstion, elementary statistics, self assessment.

DESCRIPTION OF THE SUBJECT

The course **Informatics - Elementary statistics in Excel** demands mathematical skills on the level of the 1st semester and ability to use PC. As soon as the students handle the work with formulas and charts, we add Excel's forms, i.e. interactive elements as spin button controls and scrollbars, because the experiment is the main method to explore and simulate various statistical events depending on varying parameters.

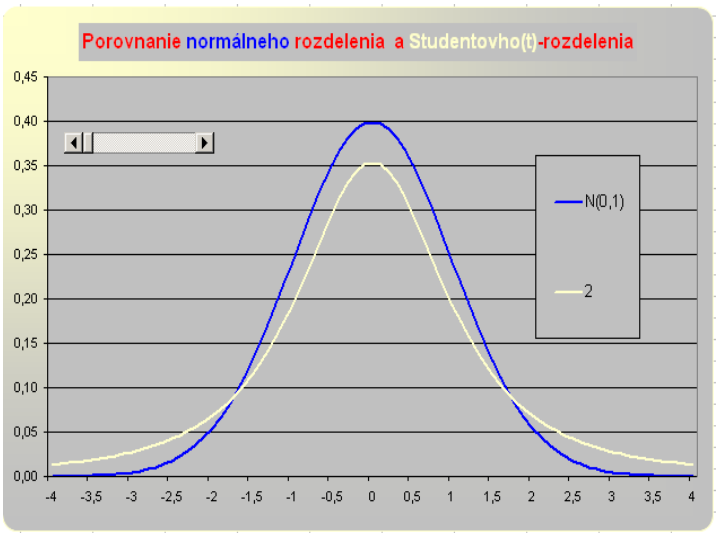


FIGURE 1
COMPARISON OF THE STANDARD NORMAL DISTRIBUTION (IN DARK BLUE) AND STUDENT DISTRIBUTION (YELLOW) WITH DEGREES OF FREEDOM SET WITH HELP OF A SCROLLBAR.

As a realistic and time related goal we stated to learn the students introduction to the theory of probability, Bernoulli scheme, basic discrete distributions and their characteristics, Gauss function, normal distribution and its characteristics, point and interval estimates, introduction to hypothesis testing, correlation, regression and trends. Except of standard and built-in tools we use add-ins Descriptive statistics, Random number generation and Histogram. We place emphasize on the ability to verify the results, whenever it is possible.

The subject Informatics – Elementary statistics in Excel is intended for full time students, but it structure can be easily transformed and adapted, and it can be used also for part time and distant students and various kinds of shorted courses. The subject is supported by a text book[], which is divided into two parts. The first part is more theoretical and the second one more practical. That's why there are two on line courses: one with lectures, which is more theoretical and supports the lectures and one more practical, which supports seminars.

The URL of the main page is http://www.math.sk/moodle. Then select Bakalárske štúdium (Bachelor's study, Figure 2). Select Informatika - Štatistika v Exceli /prednášky (lectures) or Informatika - Štatistika v Exceli /Cvičenia (seminars). Now you must log in. You have two possibilities: you can log in as a guest clicking on *Login as a guest* (Prihlásenie ako hosť*)* (if the course allows this type of access), but in this mode you cannot do tests. If you want to do them and to communicate with your teacher then you must create a new account by clicking on *New Account* (Nové konto). Then you enter the moodle clicking on *Login.* Most of courses require password (Prihlasovací kľúč). In the case of this course you must type in *nirvana*.

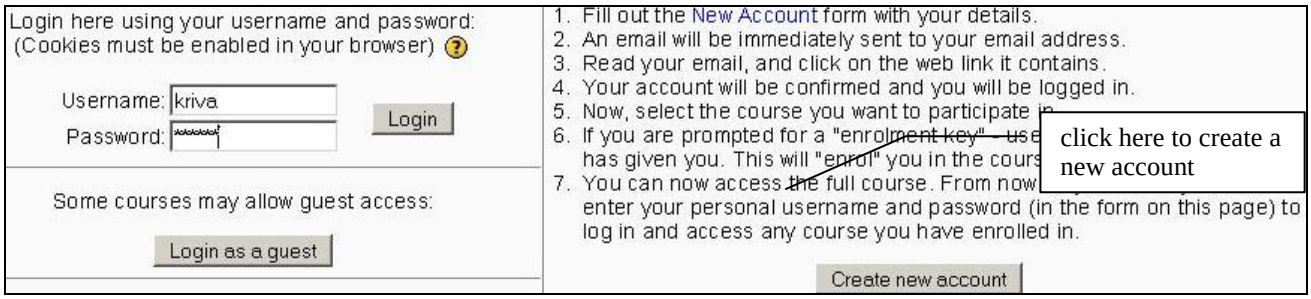


FIGURE 2

THE MAIN MENU OF MOODLE INQUIRING THE TYPE OF STUDY (ON THE LEFT). LOGGING IN THE MOODLE(ON THE RIGHT).

The system Moodle was appreciated as a great help in many aspects. We shortly emphasize the problems, which it helps to solve:

- good availability of lectures/lecture notes (important especially for part-time and distance students),
- less expensive delivering of study materials, not depending on number of students,
- updating of lectures: it is easy, comfortable, very interactive,
- unequal level of skills: supporting both less and more able students,
- it enables the higher interactivity in making experiments, it supports access to large data sets and examples prepared by teachers,
- it offers a good tool for self assessment , important for part-time and distance students,
- with its full randomization of quizzes and tests, including the random questions from a database of questions and automatic evaluation, it offers a tool for fast and correct evaluation of students, important e.g. in courses suffering from lack of time,
- the possibility to assign and correct problems via Internet
- comfortable communication between a teacher and a student.

The moodle on line subject we described was often used also for teaching some electives and part time students with typical problems: lack of time and great amount of students. Use of automatic testing became a must. Soon after launching the subject we could also clearly see the necessity to press the students to prepare by testing during the lessons. Thus, there were two kinds of tests we need to perform: short tests for self assessment performed during the lessons and tests for evaluation of the students by a teacher.

The **short tests** for self assessment typically contain 3 multiply choice questions. The questions are same for all students. After the test their solutions are discussed. Similar questions are used in final tests. A teacher has a possibility to view the result of every student and moodle offers the statistics of the test. Thus the test is a feedback not only for students but also for a teacher.
The main requirements on the short tests are:

- to be short and time restricted
- to be automatically evaluated
- their results and statistics are available.

The **final tests** are written during the last week of the semester, they require more questions, this time different for every student. The time restriction is natural, but the tests must contain amount of questions, that time for their solution is about 80 minutes and they cover all the semester material. The questions must be different, because the students have the possibility to communicate via e-mail and to watch great amount of students is difficult.

The main requirements on the final tests are:

- to contain different questions
- to be automatically evaluated
- to show the correct answers and to be automatically evaluated.

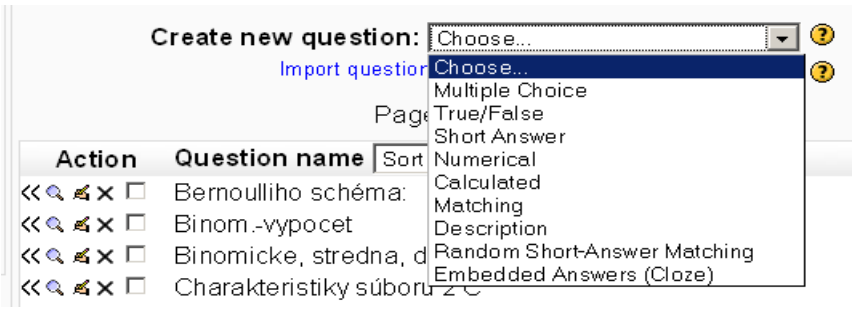
In the following text, we will show several ways how to prepare test. First we show how to prepare short test with the same set of questions for all students which are suitable for self-assessment. Then we show use of random question to generate test with different questions, and at the end, we will show, how to organize database of questions to prepare the test as well as possible.

SIMPLE TESTS FOR SELF ASSESSMENT

Tests included in lessons are intended for self assessment: they should help the student to check understanding of basic notions of the lesson by themselves. A teacher asks a question and specifies a multiple choice answer. Each attempt is automatically marked and the results can be viewed anytime the student, (but also the teacher) needs. In this way students prepare for the final test, which is composed of similar test questions. The test is a feedback not only for students but also for a teacher.

FIGURE 3
MULTIPLE CHOICE QUESTION. THE MARKED OPTION BUTTONS SHOW THE ANSWERS OF A STUDENT.

Creating a new test question. Every question must belong to some category. Thus the first step we must perform is to select (or to create a new) category. After we select a type of question: here, the questions are of multiple choice type. There are two types of multiple choice questions - single answer and multiple answer. We usually use multiple answer: in response to a question (that may include an image or an Excel table, or a link to some Internet file containing data), the respondent chooses from multiple answers. After choosing a type of question, we must fill in a corresponding form – we must type a question, set of possible answers and set evaluation.(see Figure4). Assigning 100% to the question means, that the correct question is evaluated by 1 and the wrong by 0. We would like to note, that there are many other possibilities how to evaluate.



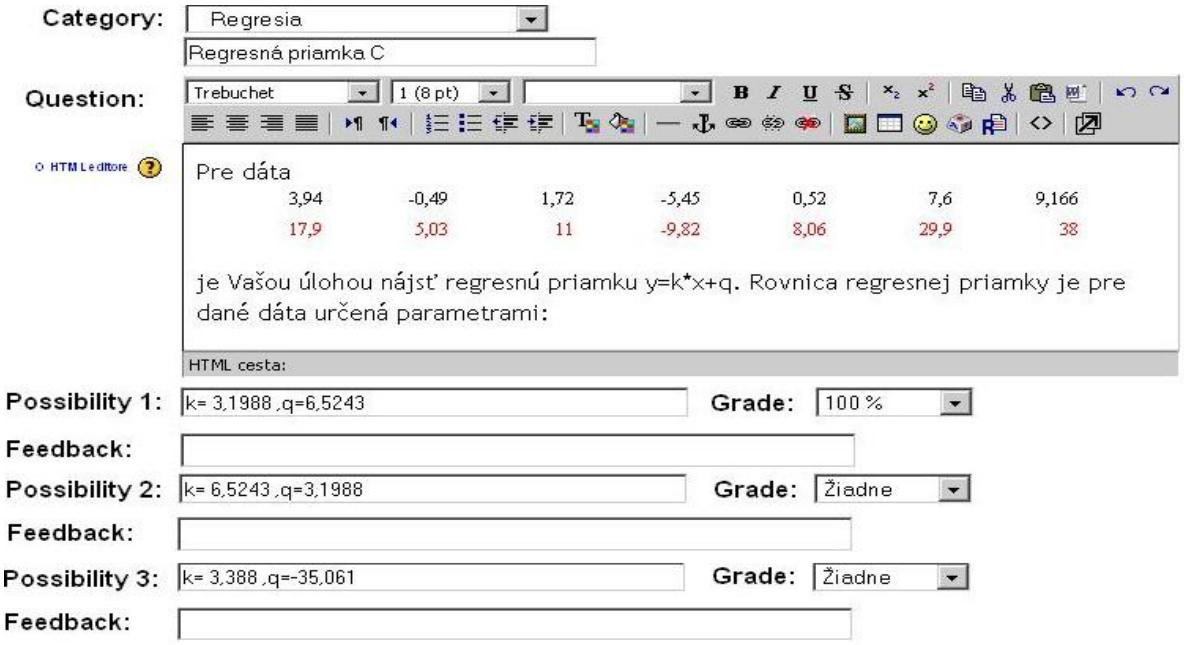


FIGURE 4
FORM FOR CREATING A TEST QUESTION.

CATEGORY OF QUESTIONS

Rather than keeping all questions in one big list, we can create categories to keep them in. Each category consists of a name and a short description. Each category can also be "published", which means that the category (and all questions in it) will be available to all courses on this server, so that other courses can use the questions in their quizzes. This feature can be reset, whenever we want. Categories can also be created or deleted at will (see Figure 6). You can also arrange the categories in a hierarchy so that they are easier to manage. The 'Move into category' field lets you move a category to another category. By clicking on the arrows in the 'Up/Down' field, you can change the order in which the categories are listed. The upper part of Figure 6 displays creating a category 'New', which is a subcategory if a category 'IS'. The description is 'Demo category'. The category will not be public. To create it, click *Add* button.

USING RANDOM QUESTIONS

Let us consider the situation, that we have three categories of questions T1,T2 and T3 for short tests. Every category contains 3 questions. We want to create a new test, where the questions are selected from existing categories in random way: e.g. two questions from T1, two questions from T2 , etc. We can use a tool called *random question*. As we have already mentioned, questions are of several types. A random question in a quiz is replaced by a randomly-chosen question from the category that was set. In this way we obtain a set of different tests covering the same topic. The easiest way how to add a random question to a test is to click on 'Add *1* random question' in the lower part of the Create new question section. The questions and answers can be permuted.

1		↓	Random question	T1
2	↑	↓	Random question	T1
3	↑	↓	Random question	T2
4	↑	↓	Random question	T2
5	↑	↓	Random question	T3
6	↑	↓	Random question	T3

FIGURE 5
Using Random Questions

HIERARCHY OF THE CATEGORIES

The quality and the success of the final test depends also on a good structure of question categories. E.g. let us have a category of 6 questions on interval estimates of 3 types, every type 2 questions. If we select two random questions from this category, it can happen, that all questions test the same type of interval. On the other hand, sometimes we want only one question on this topic. Our

experience is, that if we create a good structure of question categories, we are able to create good tests with varying number of questions and various level of details.
The new versions of moodle enable the tree structure of categories: i.e. every category can have subcategories. An example of structure of categories is displayed in Figure 6. The top category Cvic.IZP2 is intended for the second year of environmental engineering and it contains four subcategories. One of them, IS (interval estimates) contains another three categories: IS1,IS2 and IS3.

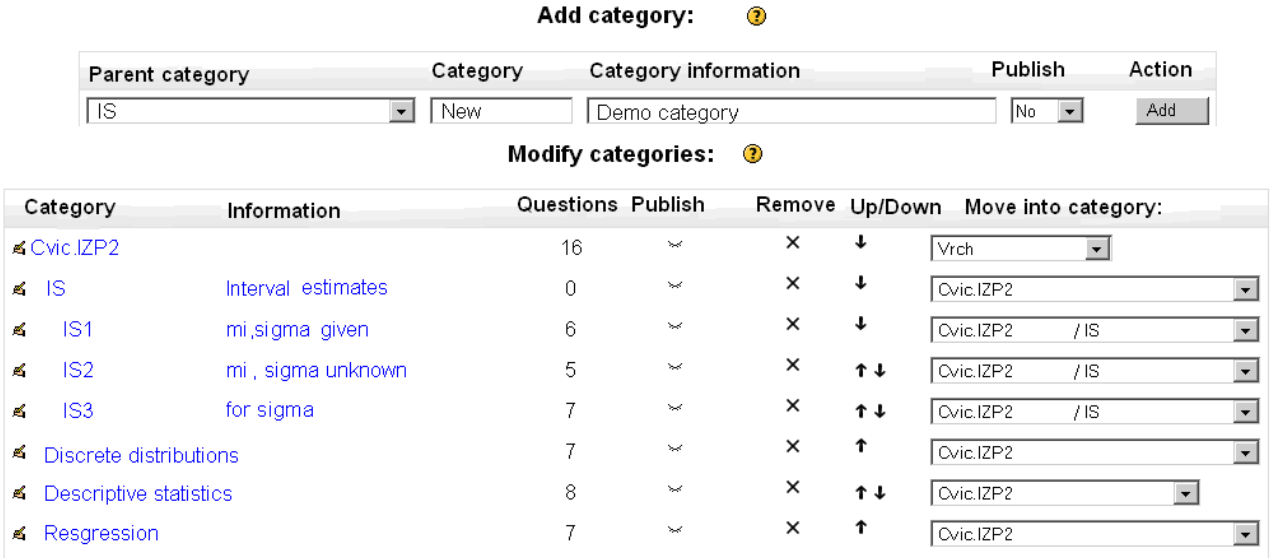


FIGURE 6
Hierarchy of categories

When we create a random question and select a category IS, that a question about any interval type is generated. If we select e.g. IS1, question of the first type interval is generated. We can set one question for every interval type.
In this example, the 'publish' field is not set. The author (owner) of this category see all his categories, but in other courses, these categories are not available. If I want to make my categories available to other courses, the top category must be set to public. In fact, it is not possible to see only S1. Its parent category, and also the parent category of this parent category must be set to public, but it means, that all Cvic.IZP2 category must be public. The 'publish' field can be switch on and off arbitrarily.

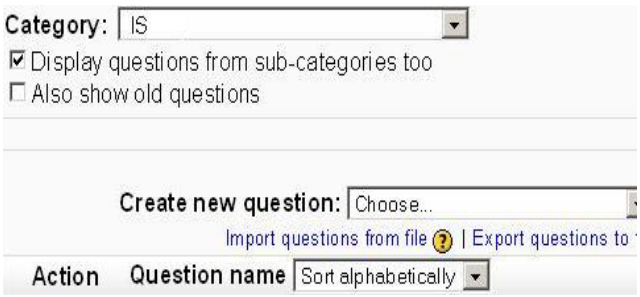


FIGURE 7
Displaying of subcategories is optional

We can move question only between categories of ours. If we work with categories and questions of some other course, the option for moving questions between categories is not avialable.

QUESTIONS AS A TEXT FILE

Preparing the database of questions, we often face these two problems:

1. we want to have several questions differing just by numbers
2. we want to print the database questions.

Both can be solved by ***importing/exporting*** the questions (see Figure 7).

This function allows us to export a complete category of questions to a text file. In some file formats some information can be lost when the questions are exported. This is because many formats do not possess all the features that exist in Moodle questions. Also some question types may not export at all. Good advice is to check exported data before using it in a production environment. There are several supported formats for import and export of database questions: in this paper we will concern only on one of them - format *.gift.*

GIFT is the most comprehensive import/export format available for exporting Moodle quiz questions to a text file. It was designed to be an easy method to for teachers write questions as a text file. It means also, that we can prepare the quiz questions locally at home, without use of moodle, and importing them into the system, we can create new categories.

The basic structure of a multiple choice question is simple: the list of answers is closed in {}, wrong answers are prefixed with a tilde (~) and the correct answer is prefixed with an equal sign (=) . They can be placed into the text arbitrarily. Questions must be separated by an empty line.

Example:

```
Tossing a cube labeled  1-6 ten times,  the probability of getting  a 3 at most 4
times   is {
        =0,984538033
        ~13/119
        ~0,154160914
        ~0,024180374
}.

The reference $F$4 is an example of  {
         ~relative
         ~mixed
         =absolute } reference.
```

In addition to these basic question types, this format offers the following options: line comments, question name, feedback and percentage answer weight.

- *line Comments.* It is a piece of text, that will not be imported into Moodle. This can be used to provide headers or more information about questions. All lines that start with a double backslash will be ignored displaying in moodle.

```
// Subheading: difficult question below
What's 2 plus 2? {=4,~5,~3}
```

- *question Name.* It can be specified by placing it first and enclosing it within double colons:

```
::Bernoulli scheme:: Is it difficult? {=no,~yes}
```

DISPLAYING THE LIST OF QUESTIONS

We want to display all questions of a category pokus. We set the category in the section for creating new questions and click on the button *Export questions to a file*. New dialoque box is displayed (see Figure 8). Leave the gift format set and again, click *Export questions to file*. The list of question is displayed and can be printed.

FIGURE 8
Displaying list of questions with help of export.

Note: The first question asks about probability of getting a 3 at most 10 times, tossing a cube 51 times. Other questions differ just by numbers.

MODIFYING QUESTIONS

If we click on *Download the exported category file,* we get the source in gift format, containing the questions (similar to Figure 10). The file, to which we exported the questions is often useless, because we do not have an access to it, however, we can display it and save it locally.

The list of questions displayed in Figure 8 can be created in several ways. First, we could type the first question in moodle, export it, download it, copy the questions and modify them. We can modify them directly in a text file, or in moodle, editing the questions. Which method is better depends on a type of question and its format (simple text, table picture, link etc.). Also we can directly type the first question in a gift file.

In the next text we display the possibility, which can be used at least in the case of simple question like thoese in Figure 8. We can use directly Excel and his functions for generating random numbers.In Excel we prepare sheet like this, following the *gift* format:

	A	B	C	D	E	F	G
1							
2	::Bernoulliho schéma::Aká je pravdepodobnost', ze ak					95	-
3	krát hodíte kockou, trojka Vám padne najviac				5	-krát?{	
4	=	0,000749355					
5	~	1/19					
6	~	0,100749355					
7	~	0,000557075					
8	}						

FIGURE 9

The cells containing the formulas (now displaying like numbers are set in following way:

F2: =21+CELÁ.ČÁST(NÁHČÍSLO()*100) [=21+INT(RAND()*100)]
E3: =1+CELÁ.ČÁST(NÁHČÍSLO()*20) [=1+INT(RAND()*20)]
B4: =BINOMDIST(E3;F2;1/6;1)
B5: =E3/F2

B6: =B4+0,1
B7: =BINOMDIST(E3;G2;1/6;0)

Pressing function key F9 causes generating new random numbers and a new question, in fact. We copy them into a gift file. Be aware, that you must work in a simple text editor (or save as a text, not in Word format). The part of such text is displayed below in Figure 10. The spaces need not be deleted. The questions are correctly displayed in Moodle.

```
::Bernoulliho schéma 5::Aká je pravdepodobnosť, ze ak                    119
-
krát hodíte kockou, trojka Vám padne najviac          13      -krát?{
 =  0,054160914
      ~    13/119
      ~  0,154160914
      ~  0,024180374
}

::Bernoulliho schéma 6::Aká je pravdepodobnosť, ze ak                    110
-
krát hodíte kockou, trojka Vám padne najviac          16      -krát?{
 =  0,327740225
      ~    8/55
      ~  0,427740225
      ~  0,089220593
}
...
```

FIGURE 10
Part of a gift file obtained by copying from Excel.

CONCLUSION

Already the first tests, which did not use random questions, proved that the moodle tests could be a very useful tool for evaluation of students. Then “bugs” (e.g. a weak student finished the test very fast and very well) had to be removed. Applying of random questions partially removed these problems, but still, we must be careful when set the properties of tests. At the end, these tests became an excellent tool for comfortable and fast evaluation, moreover, appreciated also by students as complex and objective. There are several moodle on line courses on statistics in our server. We suggest organizing a good database of questions, created and used by teacher of these courses.

REFERENCES

[1] http: //moodle.coms
[2] http: //enzv.opf.stu.cz/moodle
[3] Huba, M. a kol., "WWW a vzdelávanie", Vydavateľstvo STU, Bratislava 2003, ISBN 80-227-1999-4
[4] Kalická, J. a Krivá Z..,Praktická štatistika v Exceli, skriptum,Vydavateľstvo STU, Bratislava 2006, ISBN 80-227-2295-2

Author: Zuzana Krivá, Faculty of Civil Engineering,, Sloval Technical University, Radlinského 11, 813068 Bratislava,kriva@math.sk

Quality Planning With PLS Predictive Models in Brewery

Kupka, Karel
TriloByte Statistical Software
Stare Hradiste 574/17
Pardubice 53352, Czech Republic
E-mail: kupka@trilobyte.cz

Introduction

Partial least squares regression is a well-recognised method for modelling relationship between multivariate variables. Since its first publication this methodology has been used in econometrics, chemometrics, biometrics, sensometrics and recently, some applications in technology and qualimetrics also appeared [2], [5], [6]. We show an application in brewery, which can be however easily transferred into other branches. The goal is to predict input process parameters to generate desired output quality parameters using PLS regression.

Partial Least Squares

To find relationship between two groups of variables and predict one from another a statistical procedure based on the Partial Least Squares (PLS) approach was used. PLS was first published in form of an iterative algorithm by Herman Wold [1] in the late 1960s. Later the method was identified mathematically and widely used, see for example [4]. Though PLS is a widely recognised method, it is included only in a few commercial software packages, such as Unscrambler (CAMO) [9], or QC-Expert (TriloByte) [8]. The latter was used throughout this paper.

Assume that we have n rows of the measured process parameters, that form a matrix $\mathbf{X}(n \times p)$, and the same number of measured corresponding quality parameters in matrix $\mathbf{Y}(n \times q)$, where not necessarily $n > p$, q. In order to extract maximum information (variability) into lower-dimensional space PLS uses an analogy to well known orthogonal principal component approach (PCA) and transform **X** and **Y** into scores and loadings:

$$\begin{aligned} \mathbf{X} &= \mathbf{TP}' + \mathbf{E} \\ \mathbf{Y} &= \mathbf{UQ}' + \mathbf{F} \end{aligned} \qquad (1)$$

To ensure maximum relevance of chosen **X**-components to **Y** values, the two transformations are tied together by common scores matrix **T**. Dimensionality (number of columns) of **T** is typically smaller than that of **X** and **Y** and columns in **T**, **U** are orthogonal. This improves stability of the model and maximizes information gain. Noise and irrelevant useless information concentrates in „garbage“ matrices $\mathbf{E}(n \times p)$ and $\mathbf{F}(n \times q)$. Writing $\mathbf{U} = \mathbf{TB}$ (where **B** is a square diagonal matrix) gives us the tool for prediction of **Y** from **X**:

$$\hat{\mathbf{Y}} = \mathbf{TBQ}' \qquad (2)$$

with $\mathbf{T} = \mathbf{XP}^{-}$ computed from the new **X**-data, $\mathbf{A}^{-}$ denoting Moore-Penrose pseudoinversion of a matrix. Denoting $\mathbf{R} = \mathbf{BQ}'$ gives

$$\begin{aligned} \mathbf{X} &= \mathbf{TP}' + \mathbf{E} \\ \mathbf{Y} &= \mathbf{TR} + \mathbf{F} \end{aligned} \qquad (3)$$

which demonstrates the connection between **X** and **Y** through **T**.

Loadings and scores can be used to plot a PLS-Biplot, a graphical diagnostic tool analogical to Gabriel biplots [3] used with principal component analysis. With this plot we can see a projection of values in

matrices **X** or **Y** into 2d-plane by plotting first two columns of **T** or **U** respectively. First two rows of **P**' or **Q**' will then give projection of the original **X** or **Y**-variables into the first two PLS-principal components. If we then have coordinates of a certain position in PLS-Biplot $\mathbf{t} = (t_1, t_2)$, we can predict corresponding row of values of **X** and **Y** using

$$\mathbf{x} = \mathbf{t}\mathbf{P}_2', \text{ from which also } \mathbf{y} = \mathbf{t}\mathbf{B}\mathbf{Q}' \qquad (4)$$

where $\mathbf{P}_2$ are the first two columns or **P**.

Unlike classical least squares or Principal Component Regression (PCR), in PLS regression, **X** and **Y** are equivalent, or interchangable. This property predetermines PLS for use in biotechnology to predict properties or activity of a product from chemical composition or structure and vice versa. Multivariate statistical modeling is thus a very fertile platform for newly emerging interdisciplinary sciences.

Software for predictive modelling

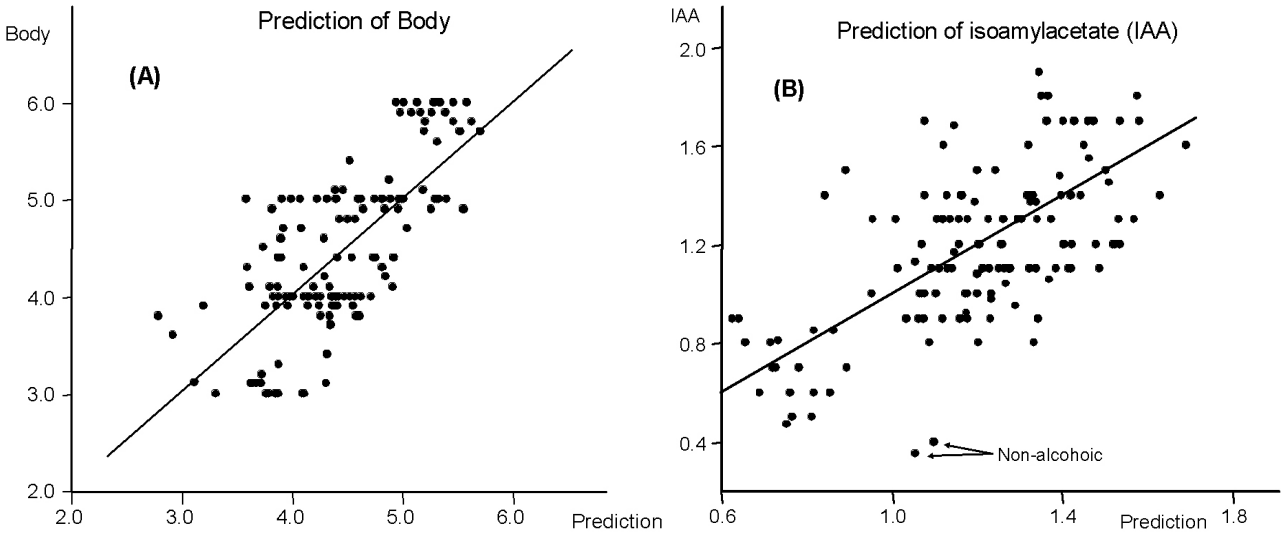
The PLS methods are not yet too common in commercially available packages. It is implemented in Norwegian software Unscrambler by CAMO and Czech software QC-Expert by TriloByte. General recommended criteria for selecting a PLS-R software include validation features to check relevance and stability of the model, plotting and visualization capabilities and user-friendliness. With some knowledge and effort it can be programmed into opened systems like Matlab or S-Plus. Other statistical and data mining packages (e.g. S-Plus, I-Miner, SPSS, SAS) involve other modelling methods like generalized additive models, neural networks, regression and classification trees, logistic regression, N-neibourhood, Bayes models, and others. GUHA methodology is implemented in a public domain software available from http://www.cs.cas.cz.

Modelling Relationship between chemical composition and sesoric value relationship

From a major Czech brewery we had a set **X** of data from chemical lab (chromatography, spectroscopy, etc.) and a corresponding set **Y** with the descriptive sensoric assessment of the same beer samples. Table **X** had over 25 columns of concentrations of the chemical compounds for each sample, table **Y** had 12 columns corresponding to different sensory attributes (like sweet, bitter, tart, caramel, astringent, etc.) marked from 0 to 10. There was total of 140 beer samples for the data analysis. These two data tables were related by computing the PLS-R model, which allows to estimate with certain precision sensoric marking from known chemical composition. Even more attractive possibility for the quality engineers and production managers was to try to predict the chemical composition for a desired set of sensory properties which would help to design new different products. The two plots in Figure 1 illustrate (A) prediction of isoamyl-acetate (banana flavour) concentration form a given sensory profile and (B) prediction of sensory attribute „body" (which can be described as the fullness of the beer's taste) of the beer from known chemical composition. The two marked points on the plot (B) are outlying non-alcoholic beer samples, which were then removed from the data. As the plots suggest, there are significant relationships between chemical and sensoric domains.

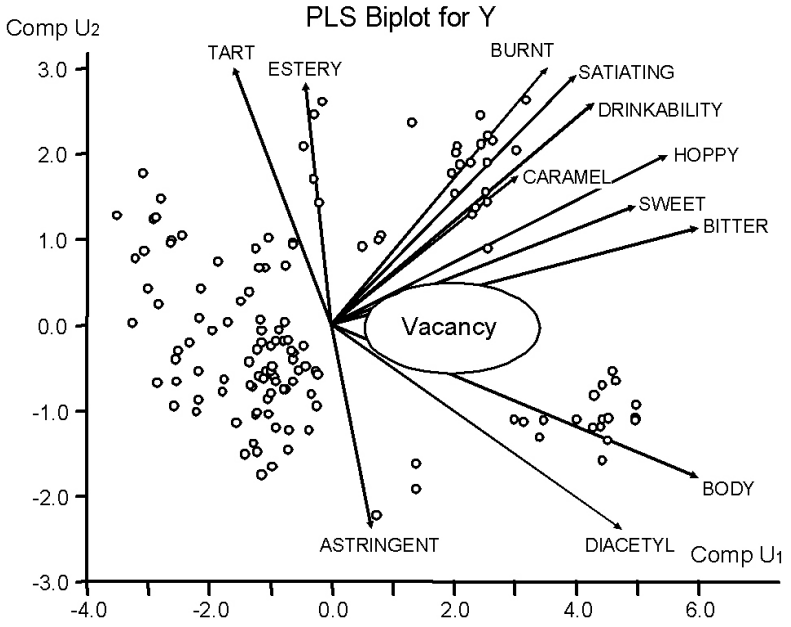
As the data from technology are often polluted with various false values or outliers, some robust modifications of PLS algorithm [7] may be used. As tests of statistical significance for parameters in PLS are complicated or unavailable, cross-validation techniques are necessary to prove prediction capability of the calculated model. In this work, an 80% repeated random validation on testing samples were used with good results. The data used in was clean of outliers, as tested by robust multivariate hotelling statistic, a classical non-robust PLS2-R algorithm was employed.

Figure 1 Prediction of Y variable: Body of the taste from chemical composition (A), Level of isoamyl acetate in beer predicted from different tastes (B).



The data collected within 3 months of production from both the sensoric evaluation and chemical analysis was used for constructing PLS-Biplot, a tool which extracts the most possible information into a 2D or 3D picture so that one can see distribution of the measured samples from all brands over the sensoric or chemical domain space. The matrices **T**, **P**, **U** and **Q** are used for construction of this plot. Individual samples are represented with dots, different beer brands form clusters. Figure 2 shows a PLS-Biplot in sensoric domain. Products are distributed rather unevenly with vacances. Those gaps may indicate uncovered market oportunities, which may thus leave room for competition. It may be of concern of the company's leadership to fill those gaps with appropriate products designed by PLS-R statistical modelling. The $\mathbf{T}_2$-coordinates of the center of vacancy (2, -0.2) were transformed using (4) into the space of sensoric values and then into the chemical composition space. Though the transformation is carried out through a reduced dimension it may be a useful lead for model-aided product design.

Figure 2 PLS-Biplot of the beer samples in the sensoric evaluation domain



Conclusion

The PLS-2 regression methods have been used in analysis of multivariate data to give a model of relationship between controllable process and chemical parameters and the subjective sensorically evaluated quality parameters of food products. The model fitted about 40% of the overall variance in data and can be used to improve stability and planning product quality parameters. A surprising side effect of PLS-Biplot was the discovery of possible gaps in product portfolio. Taking advantage of the linear character of the evaluated PLS model, an aid was given how to set up the technology to fill the gap. The use of the PLS to plan and achieve desired quality parameters was suggested.

REFERENCES

[1] Wold H.: Estimation of principal components and related models by iterative least squares. In Multivariate analysis (ed. P.R. Krishnaiah), Academic Press, New York, 1966.

[2] Martens H., Martens M.: Multivariate Analysis of Quality, J. Wiley 2000

[3] Gabriel K. R.: The biplot graphic display of matrices with application to principal component analysis, Biometrica 58 (1971), pp. 453-467

[4] Geladi P. Kowalski B.: Partial Least Squares Regression: A Tutorial, Analytica Chimica Acta 185 (1986), pp. 1-17

[5] Brereton R.G.: Chemometrics, Data Analysis for the Laboratory and Chemical Plant, J.Wiley 2002

[6] Massart D.L.: Handbook of Chemometrics and Qualimetrics, Part B, Elsevier 2003

[7] Hubert, M., Vanden Branden K.: Robust Methods for Partial Least Squares Regression, Original Research Article, mia.hubert@wis.kuleuven.ac.be, 2003

[8] Kupka, K. QCExpert – Advanced Statistical Methods, TriloByte, http://www.trilobyte.cz, 2005

[9] The Unscrambler User Manual, Camo Software, http://www.camo.no

ABSTRACT

During food design and production there is a lot of measurable quantitative, semi-quantitative and qualitative parameters. They characterize (a) properties of raw materials, (b) conditions of the production process like temperatures, flow rates, yeast parameters, and (c) properties of the final product. The latter is usually measured by three ways – objective instrumental measurements (chemical and biochemical analysis, physical and mechanical testing), objective descriptive sensoric measurements (evaluations by trained panelists) and subjective hedonic measurements (via consumer market surveys).

In brewery and beer industry it is obvious that chemical composition of a beer brand determines in some way its subjective sensoric properties. It is however very difficult if at all possible to find any simple relationship usable to estimate sensoric perception of the product given only its chemical composition. This is mainly because the relationships are non-linear with too many interactions between assays. Interaction between assays A and B means for example that A will have a different effect on the taste depending on presence or absence of B. Another problem is non-linearity of human perception itself: it will distinguish between moderate tastes much more sensitively than between two say, extremely bitter tastes. Many methods have been used as approximate tools for relating objective and subjective data, including neural networks, regression trees, logistic models, generalized linear models. In recent years it appears promising to use so-called bi-linear models, among which the Partial Least Squares regression (PLS-R) has a dominant position.

Aproximácia funkcií v experimentálnych metódach

Anna Macurová

Fakulta výrobných technológií

Technická univerzita v Košiciach

Bayerova 1

080 01

e-mail: macurova@fvt.sk

Abstrakt

V príspevku sú uvedené ukážky využitia niektorých funkcií pri vyhodnocovaní experimentov. Navrhovanie (plánovanie) experimentov sa uskutočňuje za účelom zlepšovania technologického procesu a k tomu prispievajú aproximácie funkcií v experimentálnych metódach.

Kľúčové slová: Vyhodnotenie experimentov, aproximácie funkcií, štatistické funkcie.

Významným procesom pri zhotovovaní súčiastok je *obrábanie*, ktorého cieľom je dať materiálom alebo polovýrobkom tzv. funkčnú presnosť charakterizovanú rozmermi, a požadovaným stavom obrobených povrchov.

Presnosť obrábania, nerovnosti a odchýlky pri obrábaní

Rozvoj teórie obrábania kovov spôsobuje zvyšovanie požiadaviek na zmenšovanie nepresností rozmerov a odchýlok tvaru súčiastok a na kvalitu obrobených súčiastok.

Presnosťou obrábania rozumieme [2] stupeň zhodnosti obrobenej súčiastky s výkresom súčiastky a technickými požiadavkami. Presnosť súčiastok definujú *tolerancie rozmerov a odchýlok tvaru a vzájomnej polohy.*

Nepresnosti tvaru, rozmerov a polohy súčiastok je možné charakterizovať ako odchýlky skutočných rozmerov od nominálnych, ktoré sú definované toleranciou,odchýlky od správneho geometrického tvaru (ovalita, kužeľovitosť, atď.), odchýlky vzájomnej polohy súčiastok a montážnych jednotiek.

Odchýlky, ktoré vznikajú pri obrábaní sú rozdelené do niekoľkých skupín:

1.*Teoretické odchýlky* sú odchýlky geometrického tvaru súčiastok od teoretického tvaru.

2. *Odchýlky zapríčinené nepresnosťou výrobného stroja* závisia od presnosti práce stroja.

3. *Odchýlky spôsobené zmenami teploty* . Ich príčinou sú zmeny teploty vzduchu a ohrev materiálu teplom, ktoré vzniká pri obrábaní.

4. *Odchýlky, ktoré spôsobujú upínacie sily* vznikajú pri upínaní súčiastok, keď sa deformuje nielen materiál, ale aj povrch súčiastok v mieste kontaktu s upínacím elementom.

5. *Odchýlky zapríčinené rozmerovým opotrebovaním nástroja* .

Ďalšie odchýlky môžu vzniknúť nesprávnou voľbou nastavovanej základne materiálu a dodatočným uvoľňovaním zvyškových napätí, ktoré vznikajú pri výrobe polovýrobkov a pri obrábaní súčiastok.

Kvalita obrobeného povrchu

Kvalita obrobeného povrchu je hodnotená [1] ako integrovaná charakteristika strojových súčiastok a je vyjadrená

a) geometriou obrobeného povrchu – najmä odchýlkami od ideálneho stavu,
b) hodnotením dosiahnutého stupňa drsnosti povrchu [1],
c) fyzikálno – mechanickými vlastnosťami povrchovej vrstvy – tvrdosť, spevnenie a zvyškové napätia,
d) fyzikálno –chemickým stavom povrchu.

Geometrický stav obrobeného povrchu sa podrobuje rozboru určovaním veľkosti a príčin vzniku *odchýlok* .

Obrobený povrch vždy obsahuje určitú odchýlku od požadovaného geometrického tvaru.

Súbor mikronerovností povrchu meraných na určitej dĺžke *l* je vyjadrovaný **drsnosťou** obrobeného povrchu. Aby bolo možné kvantitatívne hodnotiť drsnosť povrchov, bola prijatá európska norma normy ISO 4287: Geometrická špecifikácia povrchu – Charakter povrchu. **Drsnosť povrchu** je definovaná ako časť geometrických odchýlok s relatívne malou vzdialenosťou nerovností. Citovaná norma definuje **skutočný povrch** ako povrch, ktorý ohraničuje súčiastku a oddeľuje ju od okolitého prostredia. **Menovitý povrch** je ideálne hladký, menovitý tvar je určený výkresom alebo inou technickou dokumentáciou. Veličiny drsnosti sa vyhodnocujú od základného povrchu, ktorým je v priestore patrične posunutý menovitý povrch, a to v reze kolmom na základný, vzniknutý rez predstavuje priečny profil. Najčastejším hodnotiacim kritériom drsnosti povrchu je **stredná aritmetická odchýlka profilu** ***Ra*** . je to stredná aritmetická hodnota absolútnych odchýlok profilu v rozsahu základnej dĺžky. Výšková charakteristika drsnosti povrchu určená vzdialenosťou medzi čiarou výstupkov profilu a čiarou priehlbní profilu v rozsahu základnej dĺžky je **najväčšia výška nerovností profilu** ***Rz*** [1].

Drsnosť obrobeného povrchu pri sústružení

Drsnosť povrchu definovanú na výkrese možno považovať za limitnú hodnotu nerovností, ktoré sa môžu v technologickom procese obrábania dosiahnuť. Jednoduchý prístup k identifikácii makrogeometrie obrobeného povrchu vychádza z priameho kopírovania tvaru rezného klina na obrobený povrch.

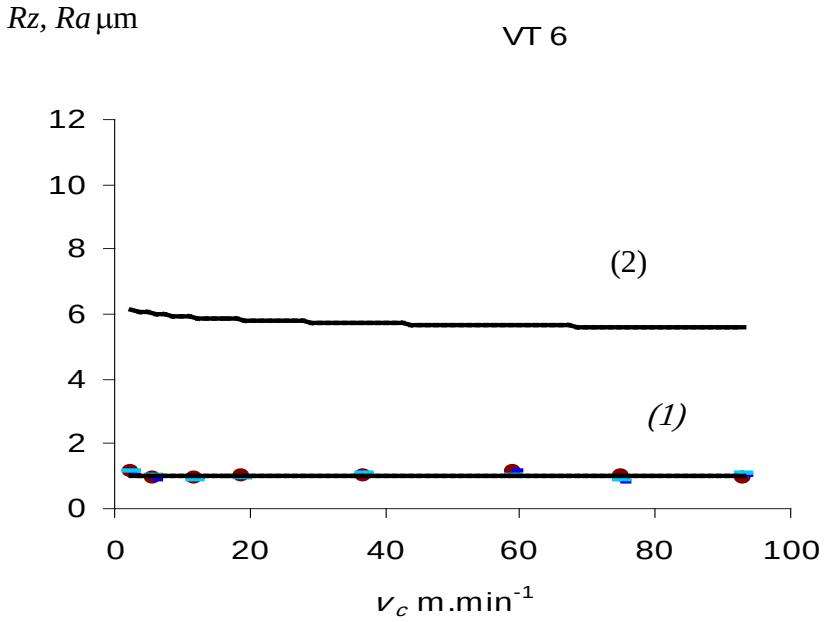
Mierou kvality obrobeného povrchu, t.j. jeho drsnosti, je veľkosť jeho vyvýšenín a prehĺbenín, ktoré na obrobenej ploche zanechal hrot nástroja. Ich veľkosť a tvar závisí na druhu obrábania a na rezných podmienkach.

Závislosť $Rz = Rz(v_c)$ v experimentoch potvrdzuje doterajší experimentálny poznatok, že s rastúcou hodnotou reznej rýchlosti v_c klesá hodnota najväčšej výšky nerovnosti obrobeného povrchu Rz. Výnimkou sú hodnoty reznej rýchlosti od 0 – 5 m.min^{-1}. Podľa dostupných prameňov takéto rezné rýchlosti v_c sa v praxi nevyužívajú, pretože dôsledkom ich zavedenia by bola nízka výrobnosť.

Aproximácie experimentálnych závislostí $Ra = Ra(v_c)$ a $Rz = Rz(v_c)$ grafmi funkcií pre materiál VT 6

Konštantné parametre rezného procesu:

posuv f = 0,2 mm, hĺbka rezu a_p = 0,5 mm, $r_\varepsilon = 1$ mm, polomer obrobku d = 53,5 mm.



Obr. 1 Aproximáciu $Ra = Ra(v_c)$ vyjadruje graf (1), aproximáciu $Rz = Rz(v_c)$ graf (2).

Vyhodnotenie experimentálnych závislostí $Rz = Rz(v_c)$ a $Ra = Ra(v_c)$

Experimentálne získaná závislosť drsnosti obrobeného povrchu od reznej rýchlosti je vyjadrená nasledujúcim vzťahom pre najväčšiu výchylku nerovnosti obrobeného povrchu sústružením

$$Rz = \frac{k_z}{v_c^{m_z}} \tag{1}$$

kde k_z, m_z nezávisia len od obrábaného materiálu.

Na základe experimentálnych charakteristík predpokladáme, že, veličina k_z vo vzťahu (1) je funkciou posuvu f a strojového času τ_s, t.j.

$$k_z = k_z(f, \tau_z) \tag{2}$$

Z predchádzajúcich experimentálnych výsledkov sú hodnoty k_z, m_z zo vzťahu (1) $m_z = 3.10^{-1}$ pre oceľ 13 270, oceľ 12 050.1 a pre titánovú zliatinu VT 6 nadobúda hodnotu $m_z = 3.10^{-2}$.

Strednú aritmetickú odchýlku profilu Ra vyjadríme analogicky, t.j. funkciou

$$Ra \approx \frac{k_a}{v_c^{m_a}} \tag{3}$$

Z experimentálnych závislostí navrhované vyjadrenie vzťahu (1) pre závislosť najväčšej výšky nerovnosti obrobeného povrchu je v tvare

$$Rz \approx \frac{k_z(f, \tau_s)}{v_c^{m_z}} \tag{4}$$

Mocninová funkcia (4) vyjadruje najväčšiu výška nerovnosti obrobeného povrchu Rz od reznej rýchlosti v_c.

Na základe experimentov je vyjadrená aproximovaná závislosť mocninovou funkciou $Rz \approx \frac{k_z(f, \tau_s)}{v_c^{m_z}}$, kde veličina k_z je funkciou posuvu a strojového času τ_s a m_z je konštanta, ktorá je určená pre obrábaný materiál experimentálne. Hodnoty meraní sú vyhodnotené pomocou funkcií popisnej štatistiky. Pomocou štatistických testovacích metód nasleduje určovanie veličín k_z a m_z, ktoré závisia od ostatných parametrov rezného procesu a podmienok experimentu. Budeme uprednostňovať skúmanie medzi dvoma kvalitatívnymi znakmi.

Literatúra

[1] BÁTORA, B., VASILKO, K.: *Obrobené povrchy*. Trenčín: Trenčianska univerzita. 2000, 183 s.. ISBN 80-88914-19-1

[2] KALPAKJAN, S.: *Manufacturing Engineering and Technology*. New York: Addiso-Wesley Publishing Company. 1990, 1199 s. ISBN 0-201-12849-7

[3] OBMAŠČÍK, M.: *Metrológia chýb a neistôt*. Žilina: MASM. 1998, 61 s. ISBN 80-85348-3

Ohraničenosť riešení diferenciálnej rovnice druhého rádu

Anna Macurová, Dušan Macura

Fakulta výrobných technológií Technickej univerzity v Košiciach
Bayerova 1, 080 01 Prešov

e-mail:macurova@fvt.sk

Fakulta prírodných a humanitných vied Prešovskej univerzity v Prešove

17. novembra 1, 081 16 Prešov

e-mail:macura@unipo.sk

Abstrakt

V príspevku je vyšetrovaná ohraničenosť riešení danej nelineárnej diferenciálnej rovnice druhého rádu pomocou polárnych súradníc. Transformáciou je získaný systém integrálnych rovníc, ktoré umožňujú vyjadriť ohraničenosť riešení toho systému.

Kľúčové slová: Nelineárna diferenciálna rovnica druhého rádu, ohraničenosť riešení nelineárneho diferenciálneho systému prvého rádu, polárne súradnice.

Analýzou diferenciálnej rovnice druhého rádu v tvare

$$(M_{AM}(t)+M_L)x''+B(x')x'+K(x)x=u \qquad (1)$$

je možné vyjadriť niektoré vlastnosti funkcií, ktorými popisujeme pohyb technického zariadenia (napríklad umelého svalu)[1]. Rovnica (1) je často v tvare

$$(M_{AM}(t)+M_L)x''+B(x')x'+K(x)x=F_L+F_{AM}$$

alebo

$$(M_{AM}(t)+M_L)x''+B(x')x'+K(x)x=M_Lg+F_{AM} \qquad (2)$$

Význam premenných v rovnici (1) je napríklad nasledovný: x, ktoré je funkciou času t je $x=x(t)$ a označuje dráhu závažia s hmotnosťou $M_{AM}(t)$ a x' je prvá derivácia dráhy podľa t , vyjadruje rýchlosť zmeny, x'' je druhá derivácia dráhy podľa t, vyjadruje zrýchlenie, $K(x)$ je funkcia, ktorá vyjadruje vlastnosti materiálu, z ktorého je pohybujúce sa teleso, funkcia $B(x')$ vyjadruje vlastnosti telesa, ktoré zabezpečuje pohyb, g je tiažové zrýchlenie, M_L je hmotnosť závažia, ktoré zabezpečuje pohyb telesa, $M_{AM}(t)$ je

hmotnosť závažia pre pohyb zariadenia, F_L je sila, ktorá zabezpečuje pohyb telesa, F_{AM} je sila, ktorá vzniká pri pohybe piestu, ktorý umožňuje činnosť telesa [1].

Budeme vyšetrovať ohraničenosť riešení diferenciálnej rovnice (1), kde označíme $M(t) = M_{AM}(t) + M_L$, $u = F_L + F_{AM}$, $F_L = M_L g$, zároveň $M(t) \neq 0$, $u = u(t)$ sú spojité funkcie na intervale J.

Rovnicu (1) vyjadríme pomocou systému diferenciálnych rovníc [5]

$$x_1'(t) = x_2(t)$$

$$x_2'(t) = -\frac{K(x_1(t))}{M(t)} x_1(t) - \frac{B(x_2(t))}{M(t)} x_2(t) + \frac{u}{M(t)}, \qquad (3)$$

kde $x_1 = x_1(t) = x$, $x_2 = x_2(t) = x'$.
Nech $\bar{x}(t) = (x_1(t), x_2(t))^T$ je všeobecné riešenie systému (3). O každom riešení $\bar{x}(t)$, pre ktoré $x_1(t_0) = x_1^0$, $x_2(t_0) = x_2^0$, $t_0 \in J$ predpokladáme, že existuje na intervale J. Označme $h > t_0 > 0$ pravý koncový bod intervalu J a $J_0 = \langle t_0, h)$ [6].
Riešenia rovnice (1) budeme vyšetrovať v tomto príspevku v zjednodušenej technickej situácii.

Položme v rovnici (1) $B(x_2) = 0$, $u(t) = 0$, potom diferenciálna rovnica (1) je vyjadrená systémom

$$x_1' = x_2$$

$$x_2' = -\frac{K(x_1)}{M(t)} x_1, \qquad (4)$$

kde $-\frac{K(x_1)}{M(t)} \in C_0(D, R) = C_0(D \equiv J \times R, R)$.

a$_1$) Nech v systéme (4) je $K(x_1) = 0$, potom všetky riešenia $x(t)$, $t \in J_0$ systému (4) sú také, že $x_2(t) = c$, $c \in R$ je konštanta a $x_1(t) = c(t - t_0)$, $c \in R$ [6].

a$_2$) Nech v systéme (4) je $K(x_1) \neq 0$.

Nech ku každému netriviálnemu riešeniu $x(t)$, $t \in J_0$ systému (4) existuje funkcia $r(t) > 0$ a funkcia $v(t)$ taká, že $v(t) \in C_1(J, R)$, $v(t) \neq 0$, kde $C_1(J, R)$ je priestor derivovateľných reálnych funkcií jednej reálnej premennej t definovaných na intervale J.
Nech pre všetky riešenia $x_i(t)$, $t \in J_0$, $i = 1,2$ platí

$$x_1(t) = r(t)\cos v(t)$$

$$x_2(t) = r(t)\sin v(t) \qquad (5)$$

Funkciu $r(t)$ nazývame polárnou funkciou a funkciu $\nu(t)$ uhlovou funkciou [6].

Systém (4) vyjadrený pomocou rovníc (5) je v tvare

$$r'(t)\cos\nu(t) - r(t)\sin\nu(t)\nu'(t) = r(t)\sin\nu(t)$$

$$r'(t)\sin\nu(t) + r(t)\cos\nu(t)\nu'(t) = -\frac{K(t, r(t)\cos\nu(t))}{M(t)} r(t)\cos\nu(t) \qquad (6)$$

Uskutočníme na systéme (6) úpravy [5], [6] a získame rovnice

$$\frac{r'(t)}{r(t)} = \sin\nu(t)\cos\nu(t) - \frac{K(t, r(t)\cos\nu(t))}{M(t)}\cos\nu(t)\sin\nu(t) \qquad (7)$$

$$\nu'(t) = -(\sin\nu(t))^2 - \frac{K(t, r(t)\cos\nu(t))}{M(t)}(\cos\nu(t))^2,\ \nu(t) \neq (2k+1)\frac{\pi}{2},\ k \in Z$$ je celé číslo.

Označíme pomocou symbolov $I_i,\ i = 1,2$ nasledujúce integrály takto

$$I_1 = \int_{t_0}^{h}\left(\sin y(t)\cos y(t) - \frac{K(t, r(t)\cos y(t))}{M(t)}\sin y(t)\cos y(t)\right)dt$$

$$I_2 = \int_{t_0}^{h}\left(-(\sin y(t))^2 - \frac{K(t, r(t)\cos y(t))}{M(t)}(\cos y(t))^2\right)dt,\ \nu(t) \neq (2k+1)\frac{\pi}{2},\ k \in Z, \qquad (8)$$

kde zápis $y(t)$, $t \in J_0$ vyjadruje spojitú funkciu.

Definícia. Riešenie $x(t)$ systému (4) nazývame $x_i - ohraničeným,\ i \in \{1,2\}$, ak $x_i(t)$ je ohraničená funkcia na intervale J_0 .V ostatných prípadoch $x(t)$ je $x_i - neohraničené$ riešenie. Riešenie $x(t)$ systému (4) nazývame $x_i - zhora$ ($x_i - zdola$) *neohraničeným,* $i \in \{1,2\}$, ak $x_i(t)$ je zhora (zdola) neohraničená funkcia na intervale J_0 [6].

Veta. Nech pre všetky spojité funkcie $y(t),\ t \in J_0$ existujú integrály I_1, I_2 (8) ako v tvrdeniach a) až f). Potom o každom netriviálnom riešení $x(t),\ t \in J_0$ systému (4) platí, že je

a) x_1, x_2 – neohraničené, ak $I_1 = \infty,\ I_2 = \pm\infty$,

b) x_1, x_2 – neohraničené, ak $I_1 = \infty,\ I_2 = L,\ L \in R$ je konštanta, $L + u(t_0) \neq \frac{k\pi}{2},\quad k \in Z$ je celé číslo,

c) x_1, x_2 – ohraničené také, že $x_1(t) \to 0$, ak $I_1 = -\infty,\ I_2 = \pm\infty$,

d) x_1, x_2 – ohraničené, také, že $x_1(t) \to 0$, ak $I_1 = -\infty$, $I_2 = L_1$, kde $L_1 \in R$ je konštanta, $L_1 + u(t_0) \neq \frac{k\pi}{2}$, $k \in Z$ je celé číslo,

e) x_1, x_2 – ohraničené, ak $I_1 = N$, $N > 0$, $I_2 = \pm\infty$,

f) x_1, x_2 – ohraničené, ak $I_1 = N$, $N > 0$, $I_2 = L$, $L \in R$ je konštanta, $L + u(t_0) \neq \frac{k\pi}{2}$, $k \in Z$ je celé číslo.

Dôkaz. Ak integrujeme (8) v intervale $\langle t_0, h \rangle$, dostaneme

$$r(h) = r(t_0) \exp \int_{t_0}^{h} \left(\sin u(t) \cos u(t) - \frac{K(t, r(t)\cos u(t))}{M(t)} \sin u(t) \cos u(t) \right) dt$$

$$v(h) = v(t_0) + \int_{t_0}^{h} \left(-(\sin u(t))^2 - \frac{K(t, r(t)\cos u(t))}{M(t)} (\cos u(t))^2 \right) dt \qquad (9)$$

Vzhľadom na predpoklad, ak existuje konečná (nevlastná) hodnota $r(h) > 0$ $(r(h) = \infty)$, t.j. ak polárna funkcia $r(t)$ je ohraničená (neohraničená), tak každé netriviálne riešenie $x(t), t \in J_0$ systému (4) je x_1, x_2 – ohraničené (x_1, x_2 – neohraničené). Situácia $x_1(t) \to 0$, $x_2(t) \to 0$ nastáva práve vtedy, ak $I_1 = -\infty$ resp. ak $r(h) = 0$ [4], [5],[6].

Literatúra

[1] BALARA, M., BORŽÍKOVÁ, J.: The Mathematical Model of the Pneumatic Artificial Muscles.In: Balara, M.a kol.: *Optimalizácia a robustifikácia mechatronických dynamických systémov*. 2006. TU Košice, s.110.

[2] GREGUŠ, M., ŠVEC, M., ŠEDA, V.: Obyčajné diferenciálne rovnice. 1985. Bratislava.

[3] KURZWEIL, J.: Obyčejné diferenciální rovnice.1978. Praha, s.424.

[4] MACURA, D., MACUROVÁ, A.: On a Transformation of Certain Generalized c-Hyperbolic Coordinates. DAAAM. 2004. Vienna. Austria. s. 261-262.

[5] MACUROVÁ, A., MAMRILLA, D.: A Remark on Transformation of Certain Nonlinear Differential Systems by Means of Hyperbolic Coordinates. *Internaional Scientific Conference on Mathematics*. Herľany, TU Košice, 1999, s. 113-117.

[6] MAMRILLA, D.: O systémoch kvázilineárnych diferenciálnych rovníc prvého rádu. Prešov, 2004. 65 s.

[7] PONTRJAGIN, L.S.: Obyčajné diferenciálne rovnice. 1982. Moskva

Commutative and non commutative s-maps

Olga Nánásiová, Katarína Trokanová, Ivan Žembery

Abstract

In this paper we will study the properties of the set of all s-maps on an quantum logic. We will show, that it is possible to construct such sequence of non commutative s-maps, which converges to commutative s-map.

Introduction

1. Introduction and the basic notions

In [3] was introduced s-map on a quantum logic and there has been shown that it is not commutative in generally. Let L be a quantum logic and let p be s-map. We say, that s-map p is commutative if for each $a, b \in L$, $p(a,b) = p(b,a)$. If L is a Boolean algebra, then $p(a,b) = p(a \wedge b, a \wedge b)$ for each $a, b \in L$. It means, that it is the measure of intersection. It cannot be proved that two observables are compatible by using s-map, but it is possible to recognize their non compatibility [4]. In this paper we will study the set of all s-maps on a quantum logic.

In this part we will briefly recall the notions as an orthomodular lattice, a state, s-map and their basic properties. These notions have been introduced and more details can be found in [5]-[7].

Definition 1. 1 *Let L be a lattice (a nonempty set with a partial ordering $\leq$, the lattice operations supremum $\vee$ and infimum $\wedge$) with the greatest element I and the smallest element O. Let $\perp$: $L \to L$ be an unary operation on L. Then $(L, O, I, \vee, \wedge, \perp)$ with the following properties:*

1. *$\forall a \in L\ \exists! a^{\perp} \in L$ such that $(a^{\perp})^{\perp} = a$ and $a \vee a^{\perp} = I$*
2. *If $a, b \in L$ and $a \leq b$ then $b^{\perp} \leq a^{\perp}$*
3. *If $a, b \in L$ and $a \leq b$ then $b = a \vee (a^{\perp} \wedge b)$ (orthomodular law)*

is called an orthomodular lattice (briefly an OML).

Definition 1. 2 *Let L be an OML. Then elements $a, b \in L$ will be called:*

- *orthogonal ($a \perp b$) if $a \leq b^{\perp}$;*
- *compatible ($a \leftrightarrow b$) if $a = (a \wedge b) \vee (a \wedge b^{\perp})$*

Definition 1. 3 *Let L be an OML. A map $m : L \to [0,1]$ such that*

(i) *$m(O) = 0$ and $m(I) = 1$*

(ii) *If $a \perp b$ then $m(a \vee b) = m(a) + m(b)$*

is called a state on L.

Definition 1. 4 *Let L be an OML. Then L will be called a quantum logic if there exists a state m on L.*

Definition 1. 5 (3) *Let L be a quantum logic. A map $p : L^2 \to [0, 1]$ will be called s-map if the following conditions hold:*

(s1) $p(I, I) = 1$;

(s2) if $a \perp b$ *then* $p(a, b) = 0$;

(s3) if $a \perp b$ *then for each* $c \in L$:

$$p(a \vee b, c) = p(a, c) + p(b, c)$$

$$p(c, a \vee b) = p(c, a) + p(c, b).$$

Definition 1. 6 (3) *Let $p : L^2 \to [0, 1]$ be s- map. Then p will be called commutative if $p(a, b) = p(b, a)$ for each $a, b \in L$*

The s-map has been introduced in [3] and there the following properties have been proved.

1. For each $a \in L$ a map $\nu(a) = p(a, a)$ is a state on L.
2. If $a \leftrightarrow b$, then $p(a, b) = \nu(a \wedge b) = p(b, a)$.
3. For each $a, b \in L$ $p(a, b) \leq p(b, b)$ and $p(a, b) \leq p(a, a)$.

In [3] can be found the example of the s-map p on OML L, which is not commutative.

2. S-map and convex combinations

Let us denote $\mathcal{P}$ the set of all s-maps on L, $\mathcal{P}_S$ the set of all commutative s-maps on L and $\mathcal{P}_N$ the set of all non commutative s-maps on L.

Definition 2. 1 *Let C be a subset of $\mathcal{P}$. C is said to be convex, if for all p_1 and $p_2 \in C$, $kp_1 + (1 - k)p_2 \in C$ for all $k \in [0, 1]$*

Proposition 2. 1 *Let L be a quantum logic. The sets $\mathcal{P}$ and $\mathcal{P}_S$ are convex.*

Proof. It follows directly from the definition of $\mathcal{P}$ and $\mathcal{P}_S$

Proposition 2. 2 *For any $p \in \mathcal{P}_N$ there exists $p_s \in \mathcal{P}_S$ such that for each $c \leftrightarrow d$ $p(c, d) = p_s(c, d)$*

Proof. It is enough to put

$$p_s(a, b) = \frac{1}{2}(p(a, b) + p(b, a)).$$

It is clear, that for $c \leftrightarrow d$, $p(c, d) = p_s(c, d)$ holds.

Proposition 2. 3 *Let $p \in \mathcal{P}$ and let $p_1(a,b) = kp(a,b) + (1-k)p(b,a)$. Then*

(i) $p_1 \in \mathcal{P}$

(ii) If $p \in \mathcal{P}_S$, *then* $p_1 \in \mathcal{P}_S$,

(iii) If $p \in \mathcal{P}_N$, *then* $p_1 \in \mathcal{P}_S$ *iff* $k = 0.5$.

Proof. The statements (i) a (ii) follow from the convexity of $\mathcal{P}$ and $\mathcal{P}_S$, respectively. (iii). It is clear, that $p_1(a,b) = \frac{1}{2}p(a,b) + \frac{1}{2}p(b,a) \in \mathcal{P}_S$, for any $p \in \mathcal{P}$.
Conversely, suppose that $p(a,b)$ is non commutative. Then there exist $a, b \in l$ such that $p(a,b) \neq p(b,a)$. Let $p_1(a,b) = kp(a,b) + (1-k)p(b,a)$ be commutative.
Then

$$p_1(a,b) = kp(a,b) + (1-k)p(b,a) = kp(b,a) + (1-k)p(a,b) = p_1(b,a).$$

$$kp(a,b) + p(b,a) - kp(b,a) = kp(b,a) + p(a,b) - kp(a,b).$$

$$(2k-1)p(a,b) = (2k-1)p(b,a)$$

$$(2k-1)(p(a,b) - p(b,a)) = 0$$

Since $p(a,b) - p(b,a) \neq 0,\ k = \frac{1}{2}$.

Definition 2. 2 *Let L be a quantum logic and let $p_n \in \mathcal{P}$ for $n = 1, 2, ...$ We say, that p_n converges to p ($p_n \to p$) if for each $a, b \in L$*

$$\lim_{n\to\infty} p_n(a,b) = p(a,b).$$

Remark 2. 1 *If $p_n \in \mathcal{P}$ ($p_n \in \mathcal{P}_S$) and $p_n \to p$ for $n = 1, 2, ...,$ then $p \in \mathcal{P}$ ($p \in \mathcal{P}_S$). It means, that $\mathcal{P}$ and $\mathcal{P}_S$ are closed.*

Theorem 2. 1 *Let $p \in \mathcal{P}_N$ and*
$p_n(a,b) = kp_{n-1}(a,b) + (1-k)p_{n-1}(b,a)$, *for $n = 1, 2, 3, ...$ and $k \in (0, 0.5) \cup (0.5, 1)$.*
Then
(i) $p_n \in \mathcal{P}_N$ for all $n = 1, 2, 3, ...$
(ii) $\lim_{n\to\infty} p_n(x,y) = \frac{1}{2}(p(a,b) + p(b,a)) \in \mathcal{P}_S$.

Proposition 2. 4 *Let X be a non empty set and let $f : X \times X \to R$ be a function. Let the following sequence of functions be defined f_n $n = 0, 1, 2...$ by the way*

$$f_0(x,y) = f(x,y)$$

and

$$f_{n+1}(x,y) = kf_n(x,y) + (1-k)f_n(y,x)$$

where $k \in (0,1)$. Then

$$\lim_{n\to\infty} f_n(x,y) = \frac{1}{2}(f(x,y) + f(y,x))$$

.

Proposition 2. 5 *Any function* $f_{n+1}(x,y) = kf_n(x,y) + (1-k)f_n(y,x)$ *can be written in the form*

$$f_n(x,y) = a_n f(x,y) + b_n f(y,x) \quad where \quad a_n + b_n = 1.$$

Proof. The proof is by induction on n. The case $n = 1$ the proposition is clear, since $a_1 = k$ and $b_1 = 1-k$. Suppose the proposition is true for some n, it means $a_n + b_n = 1$. By the definition

$$f_{n+1}(x,y) = f(x,y)[ka_n + (1-k)b_n] + f(y,x)[kb_n + (1-k)a_n].$$

It follows

$$a_{n+1} = ka_n + (1-k)b_n$$

and

$$b_{n+1} = kb_n + (1-k)a_n,$$

consequently

$$a_{n+1} + b_{n+1} = a_n(k+1-k) + b_n(1-k+k) = 1.$$

Proposition 2. 6 *Any function* $f_{n+1}(x,y) = kf_n(x,y) + (1-k)f_n(y,x)$ *can be written in the form*

$$f_n(x,y) = a_n f(x,y) + b_n f(y,x) \quad where \quad a_n - b_n = (2k-1)^n.$$

Proof. The proof is again by induction on n. The case $n = 1$ the proposition is clear, since $a_1 = k$ and $b_1 = 1-k$. Suppose the proposition true for some n, it means $a_n - b_n = (2k-1)^n$. By the definition $f_{n+1}(x,y) = a_{n+1}f_n(x,y) + b_{n+1}f_n(y,x)$ where $a_{n+1} = ka_n + (1-k)b_n$ *a* $b_{n+1} = kb_n + (1-k)a_n$. Consequently $a_{n+1} - b_{n+1} = (2k-1)^{n+1}$.

Proposition 2. 7 *Any function* $f_{n+1}(x,y) = kf_n(x,y) + (1-k)f_n(y,x)$ *where* $k \in (0,1)$ *can be written in the form*

$$f_n(x,y) = a_n f(x,y) + b_n f(y,x) \quad where \quad \lim_{n\to\infty}(a_n - b_n) = 0.$$

Proof. We want to show $\lim_{n\to\infty}(a_n - b_n) = 0$ for all $k \in (0,1)$. If $k \in (0,1)$, then $2k \in (0,2)$ and so $2k-1 \in (-1,1)$. Therefore $\{a_n - b_n\}_{n=0}^{\infty}$ is a geometrical sequence, where $|q| < 1$.

Proposition 2. 8 *Any function* $f_{n+1}(x,y) = kf_n(x,y) + (1-k)f_n(y,x)$ *where* $k \in (0,1)$ *can be written in the form*

$$f_n(x,y) = a_n f(x,y) + b_n f(y,x) \quad where \quad \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \frac{1}{2}.$$

Proof. Using Proposition 2.7 $\lim_{n\to\infty}(a_n - b_n) = 0$ and Proposition 2.5 $b_n = 1 - a_n$ we get

$$\lim_{n\to\infty}(a_n - 1 + a_n) = 0.$$

It means

$$\lim_{n\to\infty}(2a_n - 1) = 0.$$

Then

$$\lim_{n\to\infty}(2a_n - 1) + \lim_{n\to\infty} 1 = \lim_{n\to\infty} 1,$$

consequently $\lim_{n\to\infty}(2a_n) = 1$ and $\lim_{n\to\infty} a_n = \frac{1}{2}$. By Propositin 2.5 $a_n + b_n = 1$, so $\lim_{n\to\infty} b_n = \frac{1}{2}$.

Proof of Theorem 2.1. The statement (ii) follows from Proposition 2.3 (iii). Using Proposition 2.4 - 2.8, $\lim_{n\to\infty} p_n(x, y) = \frac{1}{2}(p(a, b) + p(b, a))$. From Proposition 2.3 it follows $\frac{1}{2}(p(a, b) + p(b, a)) \in \mathcal{P}_S$.

Remark 2. 2 *The set $\mathcal{P}_N$ of all non commutative s-maps is not closed.*

Acknowledgement

This work was supported by Science and Technology Assistance Agency under the contract No. APVV-0375-06, VEGA 1/4024/07, VEGA-1/3014/06

References

1. Khrennikov, A. Yu. (2003) *Contextual viewpoint to quantum stochastic*, J. Math. Phys., **44**, 2471 – 2478.

2. Khrennikov, A., Yu., Smolyanov, O., G., Truman, A., (2005) *Kolmogorov Probability Spaces Describing Accardi Models of Quantum Correlations* Open Sys. and Inf. Dyn. **12**, 371 – 384.

3. Nánásiová, O. (2003) *Map for Simultaneous Measurements for a Quantum Logic*, Int. Journ. of Theor. Phys., **42**, 1889 – 1903.

4. Khrennikov, A., Nánásiová, O.(2006) *Representation theorem of observables on a quantum system* . Int. Journ. of Theor. Phys.

5. Pták P., Pulmannová S., Quantum Logics, Kluwer Acad. Press, Bratislava (1991).

6. Riečan B., Neubrun T., Measure theory, Kluwer Acad. Press, Bratislava.

7. Varadarajan V., Geometry of quantum theory, Princeton, New Jersey, D. Van Nostrand, (1968)

Adresses:

Olga Nánásiová
Department of Mathematics, FCE Slovak University of Technology,
Radlinského 11, 813 68 Bratislava.
E-MAILS: olga@math.sk

Katarína Trokanvá
Department of Mathematics, FCE Slovak University of Technology,
Radlinského 11, 813 68 Bratislava.
E-MAILS: trokan@math.sk

Ivan Žembery
Department of Mathematics, FCE Slovak University of Technology,
Radlinského 11, 813 68 Bratislava.
E-MAILS: *zembery@math.sk*

Maps on a quantum logic

Ol'ga Nánásiová, L'ubica Valášková

Abstract

In this paper we will study functions G of two variables on a quantum logic L, such that for each compatible elements $a, b \in L \quad G(a,b) = m(a \vee b)$ or $G(a,b) = m(a \triangle b)$, where m is a state on L.

1. Introduction and the basic notions

Let $(\Omega, \mathcal{S}, P)$ be a probability space and let $A, B \in \mathcal{S}$. The measures of random events $A \cap B$, $A \cup B$ and $A \triangle B = (A \cap B^c) \cup (A^c \cap B)$ play an important role in statistical procedures. Indeed, we can study these measures as functions of two variables $A, B \in \mathcal{S}$.

In this paper we will assume an orthomodular lattice L with a state m as a basic model. This structure is also called a quantum logic, but we will not use this notion. We will study functions G of two variables on an orthomodular lattice L, such that for each $a, b \in B$, where B is a Boolean sub-algebra of L, $G(a,b) = m(a \vee b)$ (j-map [10]) or $G(a,b) = m(a \triangle b)$ (d-map [10]), where $a \triangle b = (a \wedge b^{\perp}) \vee (a^{\perp} \wedge b)$ and the function $p(a,b) = m(a \wedge b)$ (s-map [8]).

In this part we will briefly recall the notions as an orthomodular lattice, a state, s-map, j-map, d-map and their basic properties. These notions have been introduced and more details can be found in [3-10].

Definition 1. 1 *Let L be a lattice (a nonempty set endowed with a partial ordering $\leq$, the lattice operations supremum $\vee$ and infimum $\wedge$) with the greatest element I and the smallest element O. Let $\perp$: $L \to L$ be an unary operation on L. Then $(L, O, I, \vee, \wedge, \perp)$ with the following properties:*

1. *$\forall a \in L \; \exists! a^{\perp} \in L$ such that $(a^{\perp})^{\perp} = a$ and $a \vee a^{\perp} = I$*
2. *If $a, b \in L$ and $a \leq b$ then $b^{\perp} \leq a^{\perp}$*
3. *If $a, b \in L$ and $a \leq b$ then $b = a \vee (a^{\perp} \wedge b)$ (orthomodular law)*

is called an orthomodular lattice (briefly an OML).

Definition 1. 2 *Let L be an OML. Then elements $a, b \in L$ will be called:*

1. *orthogonal ($a \perp b$) if $a \leq b^{\perp}$;*
2. *compatible ($a \leftrightarrow b$) if $a = (a \wedge b) \vee (a \wedge b^{\perp})$.*

Definition 1. 3 *Let L be an OML. A map $m : L \to [0,1]$ such that*

1. *$m(O) = 0$ and $m(I) = 1$*
2. *If $a \perp b$ then $m(a \vee b) = m(a) + m(b)$*

is called a state on L.

Example 1. 1 *Let $\Omega = \{1, 2, 3, 4, 5\}$ and let $a = \{1, 2, 3\}$, $b = \{3, 4\}$, $I = \Omega$, $O = \emptyset$. Let $L = \{I, O, a, a^c, b, b^c\}$. Then (Ω, L) is not a measurable space, because $a \cap b = \{3\} \notin L$. On the other hand, there exist an additive set function m on L. For example $m : L \to [0, 1]$, $m(a) = 0.6$, $m(a^c) = 0.4$, $m(b) = 0.4$, $m(b^c) = 0.6$, $m(\emptyset) = 0$, $m(I) = 1$.*

Comment 1. 1 *We can see, that the set complement has the same properties as $\perp$ and so L from Example 1.1 is an OML and m is a state on L.*

Comment 1. 2 *The OML L from Example 1.1 can be described as an union of two Boolean algebras $\mathcal{B}_a = \{O, I, a, a^\perp\}$ and $\mathcal{B}_b = \{O, I, b, b^\perp\}$, but it is not a Boolean algebra. Indeed a, b are not compatible ($(a \wedge b) \vee (a \wedge b^\perp) = O < a$).*

.

2. Special non additive maps

The purpose of this part is to introduce a non additive maps G, which include j-maps and d-maps as a special cases. By means of this gentle generalization we would like to find more important properties of j-map and d-map.

Definition 2. 1 *Let L be an OML. A map $G : L^2 \to [0, 1]$ will be called a special non additive map (a SNAM) if the following conditions hold:*

(G1) $G(O, I) = G(I, O) = 1$;

(G2) $G(a, b) = G(a, O) + G(O, b)$ *for* $a \perp b$;

(G3) If $a \perp b$, $c \in L$ *then*

$$G(a \vee b, c) = G(a, c) + G(b, c) - G(O, c)$$

$$G(c, a \vee b) = G(c, a) + G(c, b) - G(c, O).$$

Comment 2. 1 *It is clear, that the maps $\nu_1(a) = G(a, O)$ and $\nu_2(a) = G(O, a)$ are states on L, it follows directly from the Definitions 1.3 and 2.1. In the next lemma we can see, that $\nu_1 = \nu_2$.*

Lemma 2. 1 *Let L be an OML and let G be a SNAM. Then the following statements are true:*

1. $G(a, a^\perp) = 1$ *for all* $a \in L$;

2. $G(O,a) = G(a,O)$ *for all* $a \in L$ *;*

3. $G(a,b) = G(b,a)$ *for all compatible* $a, b \in L$*.*

Proof. 1. Let $a \in L$. Then $I = a \vee a^{\perp}$ and from (G1)-(G3) we get

$$1 = G(I,O) = G(a \vee a^{\perp}, O) = G(a,O) + G(a^{\perp},O) - G(O,O) = G(a,O) + G(a^{\perp},O)$$

and analogically

$$1 = G(O,I) = G(O, a \vee a^{\perp}) = G(O,a) + G(O,a^{\perp}).$$

Hence $a \perp a^{\perp}$, then $G(a,O) + G(O,a^{\perp}) = G(a,a^{\perp})$. And so

$$2 = G(I,O) + G(O,I) = G(a,a^{\perp}) + G(a^{\perp},a).$$

And so

$$G(a,a^{\perp}) = G(a^{\perp},a) = 1.$$

2. Let $a \in L$. Then

$$G(I,O) = G(a,O) + G(a^{\perp},O),$$

then

$$G(a,O) = 1 - G(a^{\perp},O).$$

It follows from (2.)

$$1 = G(a^{\perp},a) = G(a^{\perp},O) + G(O,a),$$

then

$$G(O,a) = 1 - G(a^{\perp},O).$$

And so

$$G(a,O) = G(O,a).$$

3. Let $a \perp b$. Then $G(a,b) = G(a,O) + G(O,b) = G(O,a) + G(b,O)$. It follows from

$$G(a,b) = G(b,a).$$

Let $a \leq b$. Then

$$G(a,b) = G(a, a \vee (a^{\perp} \wedge b)) = G(a,a) + G(a, a^{\perp} \wedge b) - G(a,O).$$

Since $a \perp a^{\perp} \wedge b$ we get

$$G(a,b) = G(b,a).$$

Let $a \leftrightarrow b$. Then $a = (a \wedge b) \vee (a \wedge b^{\perp})$. Then

$$G(a,b) = G((a \wedge b) \vee (a \wedge b^{\perp}), b) = G(a \wedge b, b) + G(a \wedge b^{\perp}, b) - G(O,b).$$

Because $a \wedge b \leq b$ and $a \wedge b^{\perp} \perp b$, we get

$$G(a,b) = G(b,a).$$

(Q.E.D.)

Comment 2. 2 *As any $a \in L$ is compatible with I, we have $G(a, I) = G(I, a)$. But the last statement is not true in general. It means, that there exists a SNAM G and $a, b \in L$ such that $G(a, b) \neq G(b, a)$. If $G(a, b) \neq G(b, a)$ we know, that a, b are not compatible. This fact can be rewritten as follows: $a > (a \wedge b) \vee (a \wedge b^{\perp})$. On the other hand, if $G(a, b) = G(b, a)$, it does not imply $a \leftrightarrow b$.*

Some properties of SNAMs depend directly on value $G(I, I)$. These relations are described in the following propositions.

Proposition 2. 1 *Let L be an OML and let G be a SNAM. Then the following statements are equivalent:*

1. $G(I, I) = 1$
2. $G(a, I) = 1$ *for all* $a \in L$
3. $G(a, O) = G(a, a)$ *for all* $a \in L$.
4. $G(a, b) \leq G(a, O) + G(O, b)$ *for all* $a, b \in L$.

Proposition 2. 2 *Let L be an OML and let G be a SNAM. Then the following statements are equivalent:*

1. $G(I, I) = 0$
2. $G(a, a) = 0$ *for all* $a \in L$
3. $G(a, I) = G(O, a^{\perp})$ *for all* $a \in L$
4. $G(a, O) = G(a^{\perp}, I)$ *for all* $a \in L$
5. $G(a^{\perp}, b) = G(a, b^{\perp})$ *for all* $a, b \in L$.

Lemma 2. 2 *Let L be an OML and let G be a SNAM on it. Then $G(a, a) \leq G(a, b)$ and $G(b, b) \leq G(a, b)$.*

The following propositions result from our previous considerations.

Proposition 2. 3 *Let L be an OML and let G be a SNAM. Then G is a j-map if and only if $G(I, I) = 1$.*

Proposition 2. 4 *Let L be an OML and let G be a SNAM. Then G is a d-map if and only if $G(I, I) = 0$.*

Comment 2. 3 *If L is a Boolean algebra, the functions G can decribe an operations analogous the set's operations on a set algebra (an union and a symmetric difference).*

Let L be a Boolean algebra and let G be a SNAM. It is easy to show, that for all $a, b \in L$

$$G(a,b) = G(a \wedge b, a \wedge b) + G(a^{\perp} \wedge b, O) + G(O, a \wedge b^{\perp}).$$

If $G(I,I) = 1$, we get

$$G(a,b) = G(a \wedge b, a \wedge b) + G(a^{\perp} \wedge b, a^{\perp} \wedge b) + G(a \wedge b^{\perp}, a \wedge b^{\perp}).$$

Let us denote $G(a,a) = \nu(a)$, it is easy to see, that ν is a state on L. Then

$$G(a,b) = \nu(a \wedge b) + \nu(a^{\perp} \wedge b) + \nu(a \wedge b^{\perp}) = \nu(a \vee b) = q(a,b).$$

And if $G(I,I) = 0$, we get

$$G(a,b) = G(a^{\perp} \wedge b, O) + G(O, a \wedge b^{\perp}).$$

Now, let us denote $G(a,O) = \nu(a)$, again we can see, ν is a state on L. And so

$$G(a,b) = \nu(a \wedge b^{\perp}) + \nu(a^{\perp} \wedge b) = \nu(a \triangle b) = d(a,b),$$

where $a \triangle b$ is a symmetric difference of a, b. It has been shown in [9] and [10].

Acknowledgement

This work was supported by Science and Technology Assistance Agency under the contract No. APVV-0375-06, VEGA 1/4024/07.

References

1. Khrennikov, A. Yu. (2003) *Contextual viewpoint to quantum stochastic*, J. Math. Phys., **44**, 2471 – 2478.

2. Khrennikov, A., Yu., Smolyanov, O., G., Truman, A., (2005) *Kolmogorov Probability Spaces Describing Accardi Models of Quantum Correlations* Open Sys. and Inf. Dyn. **12**, 371 – 384.

3. Pták P., Pulmannová S., Quantum Logics, Kluwer Acad. Press, Bratislava (1991).

4. Riečan B., Neubrun T., Measure theory, Kluwer Acad. Press, Bratislava.

5. Varadarajan V., Geometry of quantum theory, Princeton, New Jersey, D. Van Nostrand, (1968)

6. Nánásiová, O. (2004) *Principle conditioning*, Int. Jour. of Theor. Phys., **43**, 1383 – 1395.

7. Khrennikov, A., Nánásiová, O.: Representation theorem of observables on a quantum system. Int. Journ. of Theor. Phys. *45* (2006), pp. 481 – 494.

8. Nánásiová, O. (2003) *Map for Simultaneous Measurements for a Quantum Logic*, Int. Journ. of Theor. Phys., **42**, 1889 – 1903.

9. Bohdalová M., Minárová M., Nánásiová O.: A note to algebraic approache to uncertainty. Forum Stat. Slovacum, (2006),*3*, pp. 31-39.

10. Nánásiová O., Minárová M., Mohammed A.: Measure of "symmetric difference", Proc. Magia, (2006), ISBN 978-80-227-2583-5, ISBN 80-227-2583-8, pp. 55-60

Adresses:

Oľga Nánásiová (*olga@math.sk*), Ľubica Valášková (*luba@math.sk*)
Dept. of Math. FCE Slovak University of Technology, Radlinského 11, 813 68 Bratislava, Slovakia

Long memory process and fractional integration in hydrologic time series

RNDr. Danuša Szőkeová

Faculty of Civil Engineering
Slovak University of Technology
Department of mathematics and descriptive geometry
Bratislava
szoke@math.sk

Abstract

Brief introduction to long memory process and fractional integration is provided. Hypothesis testing for short-term memory is described and computed. The riverflow time series of Slovak rivers are tested for short memory and fractional integrated time series are modeled with ARFIMA, STAR and LM–STAR models.

Keywords: Fractional integration, long memory, ARFIMA model, STAR model, LM–STAR model.

1. Introduction

Adequately modelling of financial and environmental time series is crucial for understanding the economic and natural processes. Long memory processes constitute a broad class of models for stationary and nonstationary time series in economics and other fields (see [1, 2, 4]). Their key feature is persistence, high correlation between events that are remote in time. Long memory processes have in recent years obtained considerable interest from both theoretical and empirical researchers in time series .

In this paper the monthly average riverflow is investigated by means of fractional integration techniques. The structure of the paper is as follows: Section 2 describes a fractional integrated process. In Section 3, the test of short memory is presented and Hurst coefficient is computed for some riverflows. Section 4 proposes ARFIMA models, STAR models and LM–STAR models for standardised observed time series or residuals time series without systematic components in order to determine the best model specification for the time series. Section 5 analyses the selected models of hydrologic time series.

2. Fractional integrated processes.

In the analysis of stationary ARMA models autocorrelation function converges exponentially to zero. That means the events separated by increasing time lags are uncorrelated. In hydrologic practice there are often time series with strong dependence between events with an increasing time lags.

The process y_t is said to be integrated of order d, or I(d), if

$(1 - B)^d y_t = \varepsilon_t$,

where B is the lag operator and ε_t is a stationary and ergodic process with a bounded and positively valued spectrum at all frequencies. We denote the covariance stationary process I(0). The I(1) process requires the first order differencing to become stationary,

etc. If order d has fractional value, -0.5 < d < 0.5, the process $(1 - B)^d y_t = \varepsilon_t$ is called fractional integrated process. For 0 < d < 0.5, y_t is a long memory process, its autocorrelations are all positive and decay at a hyperbolic rate. For -0.5 < d < 0, the sum of absolute values of the process autocorelations tends to a constant, so that it has short memory (antipersistent).

3. Short memory testing, estimation of an integrate coefficient by Hurst coefficient

To discover long memory in time series we can use first empirical analysis by:
- time series plot,
- sample correlation function plot at long lags,
- sample spectrum plot.

An illustrative time series discussed in this paper is the univariate time series of monthly average Bebrava flow at station Biskupice, basin Vah.

The hydrologic time series which are analysed in this study are the standardised time series and the residuals time series from additive model with constant trend, seasonal and cyclical components.

The member of the standardised time series are computed by the form (see [2, 3])

$$m_{ij} = (y_{ij} - \bar{y}_i)/\hat{s}_i$$

where

y_{ij} is the monthly average of riverflow, i = 1,…,12, j = 1,…,N (N represents number of years),

$\bar{y}_i = \frac{1}{N}\sum_{j=1}^{N} y_{ij}$ is the monthly mean and $\hat{s}_i = \sqrt{\frac{\sum_{i=1}^{N}(y_{ij} - \bar{y}_i)^2}{(N-1)}}$ is the monthly standard deviation.

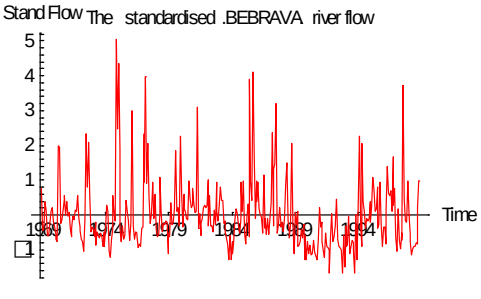


Fig. 1a: Standardised time series plot

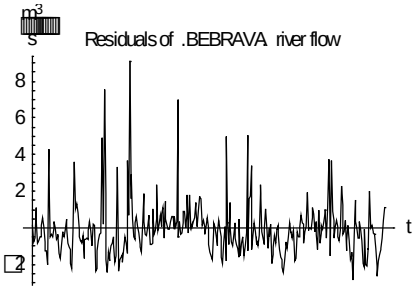


Fig. 1b: Residuals time series plot

To test the null hypothesis H_0 : there is only short memory in the time series, the asymptotic behavior of the rescaled range statistic (R/S) can be considered.

The rescaled range statistic R/S denoted $\tilde{Q}(m)$ is defined as

$$\tilde{Q}(m) = \frac{1}{S(m)}\left[\max_{1 \le k \le m}\sum_{t=1}^{k}(y_t - \bar{y}_m) - \min_{1 \le k \le m}\sum_{t=1}^{k}(y_t - \bar{y}_m)\right]$$

where $\bar{y}_m$ and S(m) are the sample mean and the standard deviation of the series y_t, t=1,…,m, $1 \le m \le N$.

As the number of observations increases, the rescaled range statistic normalised by the square root of the number of observations, $V(m) = \tilde{Q}(m) / \sqrt{m}$ (normalised rescaled range) converges to the distribution function of random variable V given as

$$F_V(v) = 1 + 2 \sum_{k=1}^{\infty} (1 - 4k^2 v^2) e^{-2(k v)^2}$$

Using normalised rescaled range statistic the null hypothesis can be rejected even when short memory is present in the time series. The rescaled range must be modified so that its statistical behavior is invariant over a general class of short memory process. To account for the short-term memory in a time series the rescaled range is accomplished by the modified rescaled range $\tilde{Q}(m,q)$:

$$\tilde{Q}(m,q) = \frac{1}{\hat{\sigma}_m(q)} \left[\max_{1 \le k \le m} \sum_{t=1}^{k} (y_t - \bar{y}_m) - \min_{1 \le k \le m} \sum_{t=1}^{k} (y_t - \bar{y}_m) \right],$$

where $\hat{\sigma}_m^2(q) = \hat{\sigma}_y^2(m) + 2 \sum_{j=1}^{q} w_j(q) \hat{\gamma}_j$, $\hat{\sigma}_y^2(m)$ and $\hat{\gamma}_j$ are estimates of the sample variance and lag j autocovariance of the series y_t, t =1,…m. The weights $w_j(q)$ are defined as

$$w_j(q) = 1 - \frac{j}{q+1}.$$

The modified rescaled range statistic $\tilde{Q}(m,q)$ and rescaled range statistic $\tilde{Q}(m)$ differ in how the range

$$R(m) = \left[\max_{1 \le k \le m} \sum_{t=1}^{k} (y_t - \bar{y}_m) - \min_{1 \le k \le m} \sum_{t=1}^{k} (y_t - \bar{y}_m) \right]$$

is normalised, $\tilde{Q}(m)$ is normalised by the sample variance S(m) and $\tilde{Q}(m,q)$ by a weighted average of the sample autocovariance up to lag q. The asymptotic behavior of the $\tilde{Q}(m)$ and $\tilde{Q}(m,q)$ will be the same only when $\hat{\sigma}_m(q)$ and the sample standard deviation S(m) converge to the same number as the sample size increase. The $\tilde{Q}(m)$ and $\tilde{Q}(m,q)$ will generally convergence to different limits in the presence of statistically significant autocorrelation.

Rejection of the null hypothesis at a particular significant level implies that a hydrologic time series may have statistically significant long memory. However, non-acceptance of the null hypothesis could only be the outcome of the test, even if only short memory is present in the time series.

Based on Hurst´s hydrological analysis rescaled range statistic was proposed method to estimate the Hurst exponent H. Then fractional parameter d is: d = H – 0.5.

The Hurst exponent H can be estimated by performing linear regression of (see [4]):

$\log(R/S)_m = H \log(m) + \log c.$

When $d \in (0, \frac{1}{2})$:

1. autocorrelations are all positive, decrease hyperbolically and have an infinite sum, $\sum_{k=0}^{n} \rho_k \to \infty$ as $n \to \infty$, such series are said to exhibit long memory because data in the distant past exerts small but non-negligible effects on the present,
2. spectral density is infinite at zero frequency.

River	V(m) =Q[a,m]/√m m=n/2	V(m) m=n/4	V(m,q) =Q1[a,m,q]/√m m=n/2 q=n/5	V(m,q) m=n/2 q=n/10	V(m,q) m=n/4 q=n/5	V(m,q) m=n/4 q=n/10	H	d
Bebrava	1.476	1.338	0.244	0.255	0.238	0.248	0.772	0.272
Bela	2.352	1.660	0.494	0.469	0.315	0.298	0.752	0.252
Boca	1.582	1.972	0.336	0.373	0.412	0.458	0.669	0.169
Čierny Hron	1.923	1.639	0.343	0.395	0.288	0.332	0.738	0.238
Ipel	1.750	0.916	0.263	0.311	0.287	0.340	0.738	0.238
Kysuca	1.192	1.051	0.431	0.399	0.387	0.360	0.748	0.248
Krupinica	1.859	1.876	0.517	0.504	0.513	0.500	0.738	0.238
Lubochnianka	2.927	2.429	0.613	0.612	0.531	0.530	0.734	0.234

Table 1 : Results of short memory testing of standardised time series

River	V(m) =Q[a,m]/√m m=n/2	V(m) m=n/4	V(m,q) =Q1[a,m,q]/√m m=n/2 q=n/5	V(m,q) m=n/2 q=n/10	V(m,q) m=n/4 q=n/5	V(m,q) m=n/4 q=n/10	H	d
Bebrava	1.949	1.090	0.341	0.328	0.191	0.184	0.678	0.178
Bela	1.897	1.454	0.310	0.316	0.231	0.236	0.509	0.009
Boca	1.428	1.513	0.450	0.407	0.482	0.441	0.551	0.051
Čierny Hron	1.620	1.168	0.229	0.225	0.170	0.168	0.624	0.124
Ipel	1.326	0.799	0.139	0.135	0.084	0.081	0.653	0.153
Kysuca	1.160	1.020	0.071	0.067	0.063	0.060	0.653	0.153
Krupinica	2.029	1.132	0.332	0.335	0.179	0.181	0.652	0.152
Lubochnianka	2.254	1.912	0.646	0.652	0.557	0.562	0.655	0.155

Table 2 : Results of short memory testing of residuals time series

The null hypothesis is rejected at 5% level of significance if the value of normalised rescaled range statistic V(m) or V(m,q) is not contained in the interval [0.809,1.862] (see [2]). If the value of normalised rescaled range statistic V(m) was not greater than 1.862 we have computed modified rescaled range statistic V(m,q).

4. ARFIMA and STAR models of fractional integrated time series

ARFIMA model

If we find that a differenced process is a stationary process, we can look for an ARMA model of that differenced process. In practice if differencing is used, usually d=1, or maybe d=2, is enough.

The process y_t is said to be *an autoregressive integrated moving average process* ARIMA(p,d,q) if its d–th difference $\Delta^d y_y$ is an ARMA(p,q) process. ARIMA(p,d,q) model has the form (see [5]):

$\Phi_p(B)\,\Delta^d\, y_t = \mu + \Theta_q(B)\,\varepsilon_t$,

where

p,d,q are non-negative integers $\Theta_q(z) = 1 + \Theta_1 z + \ldots + \Theta_q z^q$ and $\Phi_p(z) = 1 - \phi_1 z - \ldots - \phi_p z^p$ are polynomials of order p and q and ε_t is a zero mean sequence assumed to be Gaussian random variables. We suppose that the mean $\mu=0$.

ARFIMA(p,d,q) model (*an autoregressive fractional integrated moving average*) generalizes ARIMA(p,d,q) model where coefficient d takes fractional values.

The fractional integrated time series $\{ {}^d y_t \}$ from original series $\{y_t\}$ is:

$${}^d y_t = y_t + {}^d\alpha_1\, y_{t-1} + {}^d\alpha_2\, y_{t-2} + \ldots$$

according to *the fractional coefficient* d.

Coefficients ${}^d\alpha_i$ are:

$${}^d\alpha_1 = -d$$

$${}^d\alpha_i = {}^d\alpha_{i-1}\,(i-1-d)/i, \quad i = 2, 3\ldots$$

We have used ARFIMA estimation method follows two step approach. In the first step an estimate of fractional power d is obtained, in the second step ARMA estimation technique is applied to the transformed time series (filtered by the fractional differencing operator).

STAR model (see e. g. [7])

To capture nonlinear dynamics, STAR models *(the smooth transition autoregressive*) allow the model parameters to change according to the value of a *threshold variable* q_t. In each different regime, the time series y_t follows a different AR(p) model.

The standard two regimes STAR model of order p for a univariate time series y_t is given by

$$y_t = \Phi_1(B)\, y_t\,(1 - G(q_t\,;\, c) + \Phi_2(B)\, y_t\, G(q_t\,;\, c) + \varepsilon_t$$

where

$\Phi_1(B)$, $\Phi_2(B)$ are autoregressive polynomials with the shift operator B,

$G(q_t; c)$, is a so-called transition function, i.e. a non-decreasing surjective map of the values of a so-called threshold variable q_t to the interval [0, 1],

q_t is *the threshold variable* divide the domain into two different regimes concerning *the location parameters* c in each regime,

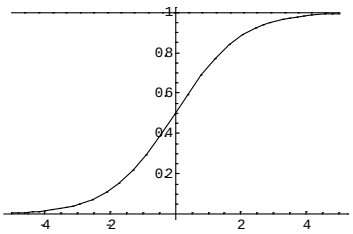
ε_t is i.i.d.$(0, \sigma^2)$.

Now the observations y_t switch between two regimes smoothly in the sense that the dynamics of y_t may be determined by both regimes.

As the transition variable q_t we usually use lagged variable y_{t-d} .

LSTAR model

The LSTAR models *(logistic smooth transition autoregressive)* are regime switching models in which the regimes switch happens gradually through the logistic function in a smooth fashion (see e. g. [7]):



$$G(q_t; \gamma; c) = \frac{1}{1 + \exp\left(-\gamma\,(q_t - c)\right)}$$

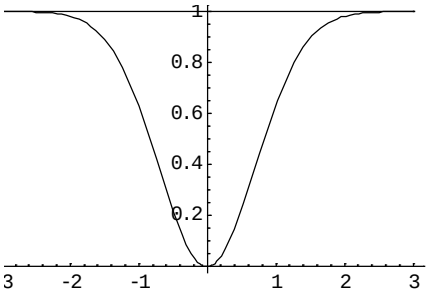
The location parameters c determines the location of the two regimes that correspond to extreme values of $G(q_t;\gamma; c)$, i.e. zero or one, (can be interpreted as the threshold), $G(c; \gamma, c) = ½$.

The smoothing parameter $\gamma > 0$ determines the speed and smoothness of transition between regimes with respect to changes in the transition variable q_t. For an LSTAR model, $G(q_t;\gamma; c) = 0$ and $G(q_t;\gamma; c) = 1$ correspond to "lower" and "upper" regimes.

ESTAR model

Another choice for $G(q_t;\gamma;c)$ is the exponential function given by (see e. g. [7]):

$$G(q_t; \gamma, c) = 1 - \exp(-\gamma(q_t - c)^2), \quad \gamma > 0,$$



which gives *an exponential* STAR (ESTAR) model .

For ESTAR model, $G(q_t;\gamma; c) = 0$ and $G(q_t\gamma; c) = 1$ correspond to "inner" and "outer" regimes. As in LSTAR models, c can be interpreted as the threshold, and γ determines the speed and the smoothness of transition.

LM–STAR model

When modeled time series is a long memory time series STAR model is called LM–STAR (*long memory* STAR).

In spite of the similarity between LSTAR and ESTAR models, they actually allow for different types of transitional behavior. We have constructed all these models for two types of datasets:

a) residuals without systematic components of observed riverflow data (models indexed with r),

b) standardised observed riverflow data (models indexed with o).

5. Analysis of selected models for hydrologic data

The gauging stations provide continuous records of the riverflow through time (see e. g. [6]). The monthly average flows of Slovak rivers observed from November 1968 to October 1999 (n=372) are examined by means of fractional integration techniques, used for analysis and forecasting with ARFIMA and LM–STAR models. The data have been divided into in-sample part, M = 360 observations and out-of-sample part, P = 12.

Fitting ARFIMA model

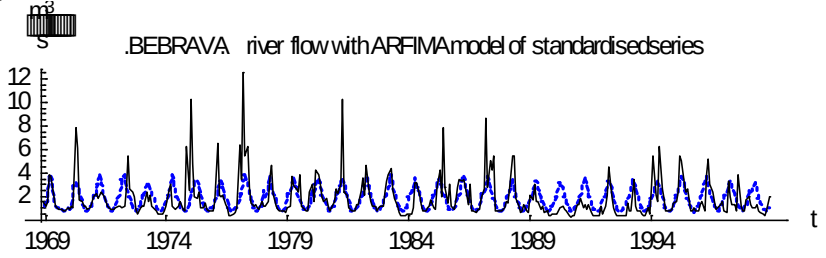


Fig.2a :Observed time series and ARFIMA$_o$(p=1,d=0.27,q=0) standardised series model

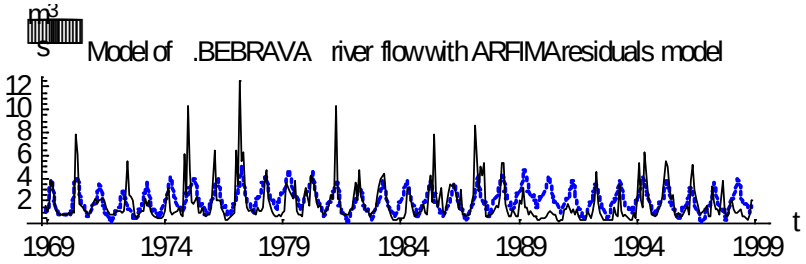


Fig.2b :Observed time series and ARFIMA$_r$(p=1,d=0.178,q=0) residuals series model

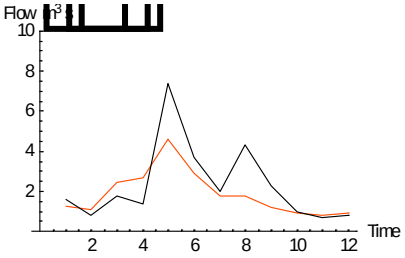


Fig. 3a:Observed out of sample time series and forecast with ARFIMA$_o$ model

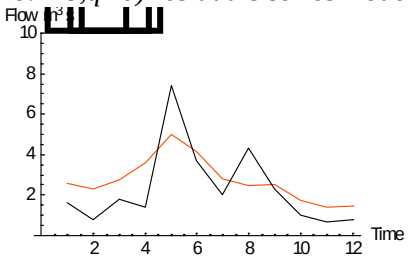


Fig. 3b:Observed out of sample time series and forecast with ARFIMA$_r$ model

Fitting LM–STAR model

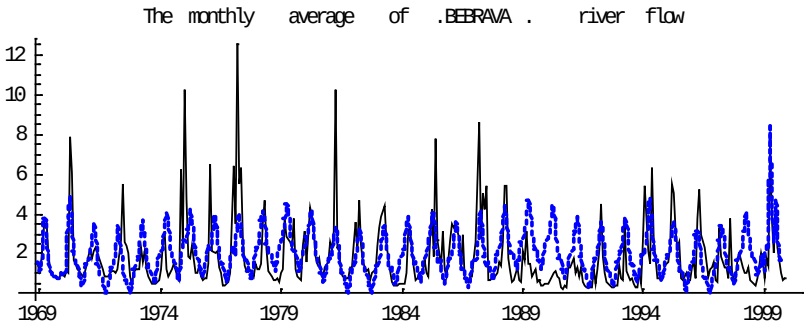


Fig.4 :Observed time series and LM–LSTAR$_r$ model of standardised diferenced series

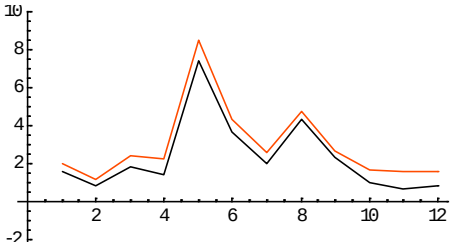


Fig. 5:Observed out of sample time series and forecast with LM–ESTAR$_r$ model

Residuals variance $\hat{\sigma}^2$ *(30 years)*					
Model	**Bebrava**	**Bela**	**Boca**	**Lubochnianka**	**Cierny Hron**
Original data	3.482	8.449	2.526	1.525	6.894
Syst. Components	2.449	3.098	1.589	0.931	4.924
$AR_r(1)$	2.462	3.368	2.973	1.773	9.905
$ARFIMA_o$	2.004	2.969	1.474	0.764	4.095
$ARFIMA_r$	2.154	3.365	1.423	0.835	4.500
$LSTAR_o$	**1.047**	5.680	1.889	1.086	6.185
$LSTAR_r$	1.094	3.365	**1.276**	**0.382**	2.462
$LM\text{-}LSTAR_o$	2.059	3.062	1.394	0.857	**1.394**
$LM\text{-}LSTAR_r$	2.017	4.563	1.607	0.619	4.805
$ESTAR_o$	1.638	7.464	2.155	0.986	2.531
$ESTAR_r$	1.744	3.369	1.749	0.392	2.428
$LM\text{-}ESTAR_o$	2.021	**2.873**	1.469	0.848	4.612
$LM\text{-}ESTAR_r$	2.171	3.724	1.560	0.903	4.768

Table 3 : Summary of ARMA, ARFIMA, STAR and LM–STAR models of some rivers

Forecast errors (12 months)										
	Bebrava		**Bela**		**Boca**		**Lubochnianka**		**Cierny Hron**	
Model	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**	**RMSE**	**MAE**
$AR_r(1)$	1.987	1.007	1.016	1.007	1.182	1.106	1.399	1.030	1.424	2.076
$ARFIMA_o$	1.224	0.851	0.720	0.605	0.634	0.501	0.797	0.710	1.182	0.784
$ARFIMA_r$	1.310	0.851	0.516	0.461	0.789	0.561	0.773	0.637	1.336	0.972
$LSTAR_o$	0.539	0.475	0.672	0.492	0.430	0.3621	0.494	0.431	1.025	0.9708
$LSTAR_r$	0.508	0.455	0.276	0.231	**0.093**	**0.268**	0.261	0.221	0.598	0.580
$LM\text{-}LSTAR_o$	1.574	1.085	0.961	0.778	0.886	0.702	0.89419	0.794	1.574	1.157
$LM\text{-}LSTAR_r$	0.779	0.987	1.009	0.889	0.555	0.435	0.407	0.356	0.840	0.619
$ESTAR_o$	0.536	0.453	0.840	0.751	0.662	0.559	0.487	0.446	0.910	0.803
$ESTAR_r$	**0.367**	**0.321**	**0.239**	**0.186**	0.367	0.301	**0.253**	**0.221**	0.555	0.500
$LM\text{-}ESTAR_o$	1.574	1.076	2.060	1.290	0.794	0.619	0.912	0.761	1.439	1.058
$LM\text{-}ESTAR_r$	0.681	0.640	0.928	0.914	0.530	0.466	0.420	0.369	**0.550**	**0.483**

Table 4 : Summary of ARMA, ARFIMA, STAR and LM–STAR forecast errors of some rivers

Conclusion

The results show the hydrologic time series can be specified in terms of fractional integrated statistical models with orders of integration significantly smaller than 0.5. Thus, the riverflows seem to be stationary.

We have tested for short memory riverflows time series from Slovakia regions. After hypothesis testing the most of riverflows series seems to be the long memory time series. In such case LM−STAR models may be considered for data description or data forecasting.

From results presenting in table 3 is evident that the LM−LSTAR$_o$ model is the best one for Cierny Hron riverflow and LM−ESTAR$_o$ model is the best one for Bela riverflow data description. For data forecasting only model LM-ESTAR$_r$ for Cierny Hron riverflow has the better results than STAR models. This leads to conclusion the long memory STAR models may be in some cases better than short time models.

References:

[1] Dueker,M., Asea,P.K. :Non-Monotonic Long Memory Dynamics in Black-Market Exchange Rates, Working paper, 1995

[2] Ramachandra,A.,R., Bhattacharya,D.: Hypothesis testing for long-term memory in hydrologic series, *Journal of Hydrology* 216 (1999)

[3] Fouskitakis,G.N., Fassois,S.D.: On the estimation of long-memory time series models, (from internet)

[4] Sakalauskiene,G.: The Hurst Phenomenon in Hydrology, *Environmental research, engineering and management*, 2003.No.3(25)

[5] Artl,J.,Artlová,M.: *Ekonomické časové rady*, GRADA Publishing, 2007

[6] Kohnová,S., Solín,Ľ., Szolgay,J.: Regional flood frequency analysis, *Environment*, Vol.27 (in Slovak)

[7] Franses, P. H., D. van Dijk (2000): *Non-linear time series models in empirical finance*. Cambridge University Press, Cambridge

Numerical simulation of heat and moisture propagation in building envelopes during climatic cycles

Stanislav Šťastník, Jiří Vala and Radek Steuer
Brno University of Technology, Faculty of Civil Engineering
602 00 Brno, Veveří 95, Czech Republic
vala.j@fce.vutbr.cz, stastnik.s@fce.vutbr.cz, steuer.r@fce.vutbr.cz

Introduction

Reliable simulation of heat propagation in the envelopes of buildings, determined (especially in exterior layers) by the moisture content in pores, under the day and year quasi-periodic climatic conditions, belongs to the crucial computational problems in civil engineering. Most software packages (both commercial and research ones) implement rather complicated algorithms with strange material characteristics that cannot be obtained easily for particular insulation materials in practice; on the other hand, they neglect some significant processes, as long- and short-wave radiation from and to the earth atmosphere totally. The fundamentals of the theory of heat propagation in building structures, based on the laws of conservation from classical physics, needed in this paper, are introduced and explained in [1] and [10]. The more detailed analysis of sources and mechanisms of radiation activities is contained in [7] where also the mathematical basis and the main ideas of numerical algorithms for the computational modelling and simulations, reported in this paper, are formulated. From the Central European view, even such complex computational systems as WUFI ("Wärme Und Feuchte Instationär") at the Frauenhofer Institute for Building Physics or DELPHIN at the Technical University in Dresden, as well as other HAM ("Heat, Air and Moisture") packages, referenced in [6], are not able to couple heat propagation including both types of radiation from the external environment with moisture redistribution properly.

The heat and moisture propagation in building envelopes is driven by the climatic changes, whose character is nearly periodic – they occur in day and year cycles, and is conditioned by both insulation and accumulation properties of particular layers of building structures. We shall demonstrate that the reconstruction of an old panel block of flats is able to prepare better conditions for the condensation of atmospheric vapour on and near the exterior surface of a building envelope; the logical consequence of the occurrence of such wet locations is the population of algae as a visible surface damage and a starting point for further destructive processes.

Moreover, the thermal technical material characteristics, namely in case of materials with high porosity, differ dramatically in the wet state and in the dry state; this highlights the significance of the theory and practice of the identification of basic material properties. However, the main result of this paper is the presentation of selected results of numerical simulations for one commonly used structure of the insulation system: firstly with neglected radiation, then respecting the radiation influence. Finally we shall mention some practical recommendations for the producer of building insulations, how to prevent or reduce the damage caused by radiation.

Observations and measurements

The effects of growing of algae thanks to the presence of liquid water can be easily observed on both new bulding structures and reconstructed bulding envelopes. Fig. 1 shows that the algae population if often more durable than the attempts to hide the damage by additional coatings, here as a part of some commercial advertisement.

Figure 1: The damaged wall with visible growing algae

To quantify the sources of such phenomena, it is important to know the development of the temperature field on the damaged surface. The Faculty of Civil Engineering of BUT develops new ecological insulation materials, based on the wood waste, and identifies their properties. Fig. 2 documents the thermo-visual measurements of several samples, provided by various coatings, installed on a roof under real climatic conditions; the temperature of outer environment decreases by 2.7 K in 43 minutes here.

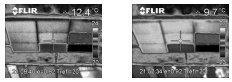


Figure 2: Two typical outputs from the thermo-visual measurements of nine samples

However, the inexpensive and precise measurement device for the reliable quantification of emissivity of various coatings, conditioning the process of thermal radiation, is not available. The surfaces of most coating materials are not perfect planar, they can be better characterized as rather ragged. Such insufficiency does not admit direct measurements, based on the monitoring of the intensity of rays, reflected from the surface. This difficulty can be avoided by the integral measurement of the reflected diffusive radiation from the surface, known as the "ball" method, implemented at the unique FTIR (Fourier transform infrared) spectrometer Bruker IFS-66/S. The gold-coated interior sphere of the measurement device handles the

wide spectral range from the near to the far infrared one, in practice the wavelengths from 5 to 20 μm. Some necessary measurements for selected coating materials have been done at the National Institute of Chemistry in Ljubljana (Slovenia), as documented in [4]. However, the above described measurements are expensive and not available for common use, thus the indirect method with the plate stationary measurement device Holometrix Lambda 2000, is used: it is based on the measurement of the thermal resistance of a plate which serves as a carrier of the analyzed coating.

Let us remind that more complex projects, analyzing climatic changes, as the SWERA project, mentioned in [7], makes use of various pyrheliometers, pyrgeometers and albedometers. Nevertheless, the quantitative outputs from particular measurements differ substantially; therefore some authors even try to avoid all deterministic settings of values of model parameters, using the correlation analysis, Monte Carlo simulations, etc.

Identification of material properties

The basic thermal technical characteristics, namely the heat conduction coefficient (decissive for the thermal insulation ability of materials) and the heat capacity (crucial for the thermal accumulation process), can be fortunately obtained by more reliable methods than the parameters from the preceding section. This characteristics are related to the dry material state; this is difficult to guarantee by most classical measurement methods, especially in case of the heat capacity. This was the motivation for the development of the original inexpensive measurement device, whose efficiency comes from the non-trivial numerical simulation of the heat propagation under artificially arranged simple initial and boundary conditions; the detailed description of such device has been presented in [2], the complete mathematical and numerical analysis, based on the Fourier method, will appear in [3].

The study of uncertainties in measurements is important also in this case. The approach of [3] and [2] involves such analysis as its integral part (some steps from the proposed algorithm of identification of two basic thermal technical characteristics can be used for the quantification of the uncertainties in measurements, too), in other cases this must be done separately by standard probabilistic and statistical methods – this is needed e.g. for accreditation of each technical laboratory.

A special methodology must be used for the experimental analysis of the process of moisture transfer in the pore space of material layers; this process is much slower than the heat transfer, depends on the number and size of macro- and micro-pores, capillary channels, etc., and under real climatic condidtions can be modified by phase changes, as from vaper to liquid water (as metioned above) or from liquid water to ice. Moisture (in all admissible phases) in porous system of building materials replaces air. Water molecules adsorb on surface of pores in an initial period of filling porous system by water and form mono-molecule layer and filling of pores by adsorbed water; this is accompanied by formation of local heat bridges and enormous increase of the heat conduction coefficient; macro-pores are then completely filled

by water during progressive penetration of moisture, while the speed of increasing of value of the coefficient is not so fast. The extensive experimental work in this research branch is connected with the development of WUFI, as documented in [5]; the methods of testing of insulations for extreme use are discussed in [9].

Numerical simulation

For the analysis of simultaneous processes of heat and moisture transfer, including the radiation influence, the original software code (written in Pascal), based on the finite difference and volume methods, was created at the Faculty of Civil Engineering at BUT. The necessary mathematical and numerical preliminaries are sketched in [7]; all inputs and outputs of this program are able to be compared and identified with long-time observations and laboratory measurements.

Fig. 3 shows the evolution of temperature and moisture during a time period of one month (March, only a short cut from all simulation results is presented here), obtained from the one-dimensional model – the heat propagation is active only in one direction, perpendicular to all material layers. The drawn curves correspond to various points of particular material layers. The first (left-hand side) part of Fig. 3 has been obtained under the assumption of zero radiation, its second (right-hand side) part includes radiation, because of the present lack of reliable data due to the rather simple recommendations of the Society of German Engineers by [8].

The following table presents the typical sample of both measured and simulated temperatures (in Celsius degrees) on the external surface:

time (hour)	19	20	21	22	23	0	1	2	3	4	5	6
environment	13.2	12.1	11.0	9.9	8.8	7.6	6.5	5.4	4.4	3.6	3.2	3.3
dewpoint	4.5	4.1	3.7	3.3	3.0	2.6	2.2	1.9	1.5	1.1	0.7	0.5
measured	15.8	13.9	12.0	10.1	8.2	6.4	4.7	3.2	2.2	1.7	1.9	3.0
radiation	15.9	14.0	12.0	10.1	8.1	6.3	4.6	3.2	2.1	1.7	2.0	3.1
no radiation	16.3	14.7	12.9	11.2	9.5	7.9	6.4	5.1	4.1	3.7	3.9	4.9

Although the condensation is not active here (the dewpoint has not been reached), this table is able to illustrate the fact that the simulation results coincide with practical observations much better for the computation with included influence of radiation than for that with neglected radiation. The moisture contect in all layers near the exterior surface is increased because of the low diffusive resistance of the thin outer porous silicate layer; the sorption effect can be highlighted by the presence of condensed liquid water. The consequential volume changes, bringing additional stresses into constructive and insulation materials, can be even more dangerous than the most visible population of algae.

Figure 3: Comparison of the temperature and moisture development without and with radiation

Conclusions

The computational modelling of heat propagation, incorporating moisture redistribution and influence of atmospheric radiation, based on the proper formulation of the conservation laws of classical physics, produces better approximation of real phenomena than the simplified formulae from (still valid) technical standards. However, the quantitative analysis of the process of slow moisture propagation, determined by the microstructural phenomena in the pore space, is not quite reliable and requires a lot of experimental and theoretical study in the near future; morever, this process can be modified significantly by radiation effects from external environment.

Nowadays the possibilities for repairs of damaged surface of constructions are limited – we are not allowed to solve such problems by changing nature. To reduce the thermal insulating ability of masonry is unreasonable. Chemical agents are the only mean of protection against algae; those agents do not have long-term effect and are not environmentally harmful. The choice of building materials must respect the emissivity, which should be low in the long-wave infra-red spectrum. Several Czech producers of building materials are therefore interested in practical outputs from the project, covering both the design of new insulation materials, the progress in the methodology of measurements of material properties and the development of the above demonstrated modelling and simulation software.

Acknowledgement. This research is supported by the Czech research project CEZ MSM 0021630511.

References

[1] Davies, M. G.: Building Heat Transfer, John Wiley & Sons, 2004.

[2] Šťastník, S., Vala, J., Kmínová, H.: Identification of thermal technical characteristics from the measurement of non-stationary heat propagation in porous materials, Forum Statisticum Slovacum 3 (2006), pp. 203-209.

[3] Šťastník, S., Vala, J., Kmínová, H.: Identification of thermal technical characteristics of building materials, Kybernetika (Acad. Sci. Czech Rep.), to appear.

[4] Šťastník, S., Vala, J., Steuer, R.: Evaluation of the emissivity of outer surfaces of building constructions, Abstr. Thermophysics 2006 in Kočovce (Slovak Rep.), p.9, CD-ROM Proc., 5 pp.

[5] Time, B., Kvande, T., Waldum, A. F., Oustad, M.: Rain penetration resistance of renders – laboratory testing and numerical calculations, research report of Climate 2000 Project, Norwegian Buliding Research Institute, 2004, available at http://www.byggforsk.no/prosjekter/klima2000/portals/0/Rain%20Penetration%20Resistance%20of%20Renders.pdf.

[6] Vala, J., Šťastník, S.: On the thermal stability in dwelling structures, Building Research Journal 52 (2004), pp. 31-56.

[7] Vala, J., Šťastník, S., Steuer, R.: Modelling of radiative heat transfer in building envelopes, Numerical simulation of heat and moisture propagation in building envelopes during climatic cycles, Forum Statisticum Slovacum 4 (2007), to appear.

[8] Verein Deutcher Ingenieure: Umweltsmeteorologie – Wechselwirkungen zwischen Atmosphäre und Oberflächen, Berechnung der kurz- und der landwelligen Strahlung, VDI Richtlinien 3879, Blatt 2, 1994.

[9] Wang, W.-T.: Characterization of syntactic foam pack-in-place and pour-in-place insulation for ultra-deepwater, research report of Cuming Corporation, Bayou Flow Technologies, 2002, available at http://www.cumingcorp.com/pdf/OPT2002.pdf.

[10] Wit, M. H.: Heat and Moisture in Building Envelopes, Technische Universiteit Eindhoven, 2006.

Metóda funkcie užitočnosti vo viackriteriálnom rozhodovaní

Milan Terek[1]

Summary. The paper deals with the normative approach to multiattribute decision making. In the first part the axioms of the expected utility and the principle of the maximization of expected utility are described. Then the using of the additive utility function is illustrated on the simple example from the area of the application of the Taguchi method.

1 Kľúčové slová: viackriteriálne rozhodovanie, aditívna funkcia užitočnosti, Taguchiho metóda

Úvod

V prvej časti príspevku uvedieme základné poznatky o teórii užitočnosti. Na základe niekoľkých axióm „rozumného“ správania v rozhodovacích situáciách v podmienkach neurčitosti možno dospieť k poznaniu, že rozhodovateľ, ktorý pri rozhodovaní rešpektuje tieto axiómy, sa rozhoduje na základe maximalizácie očakávanej užitočnosti.

V druhej časti si všimneme problém viackriteriálneho rozhodovania. Podrobnejšie si všimneme tzv. *normatívny prístup*, ktorý je založený na množine axióm o štruktúre preferencií rozhodovateľa[2].

1 Axiómy očakávanej užitočnosti a princíp maximalizácie očakávanej užitočnosti

Uvedieme najprv axiómy správania rozhodovateľa, ktoré vedú k rozhodovaniu na základe očakávanej užitočnosti[3].

2 *Úplnosť*. Pre ľubovoľné dva varianty a_1, a_2, rozhodovateľ buď preferuje a_1 pred a_2 ($a_1 \prec a_2$), alebo preferuje a_2 pred a_1 ($a_2 \prec a_1$), alebo je indiferentný medzi a_1 a a_2 ($a_1 \sim a_2$).

3 *Tranzitívnosť*. Pre ľubovoľné tri varianty a_1, a_2, a_3 platí: keď $a_1 \prec a_2$ a $a_2 \prec a_3$, potom $a_1 \prec a_3$.

---

[1] Tento príspevok vznikol s prispením grantovej agentúry VEGA v rámci projektu číslo 1/3182/06 Zlepšovanie kvality produkcie strojárskych výrobkov pomocou štatistických metód.

[2] Opisný prístup je založený na párových porovnávaniach variantov vzhľadom na všetky kritériá. Podrobnejšie napríklad BEROGGI (1999).

[3] V literatúre o teórii užitočnosti sa uvádzajú rozličné systémy axióm, aj keď rozdiely medzi nimi nie sú veľké. My sme sa pri formulácii systému axióm opierali najmä o publikácie CLEMEN – REILLY (2001) a BEROGGI (1999).

4*Spojitosť*. Rozhodovateľ je indiferentný medzi dôsledkom A a neurčitou udalosťou, ktorá obsahuje len výstupy A_1 a A_2, pričom $A_1 \sim A \sim A_2$. To znamená, že možno formulovať takú referenčnú hru s pravdepodobnosťou $p \in (0, 1)$, pre ktorú rozhodovateľ je indiferentný medzi touto referenčnou hrou a A.

5*Redukcia zložených neurčitých udalostí*. Rozhodovateľ je indiferentný medzi zloženou neurčitou udalosťou (komplikovaná kombinácia hier alebo lotérií) a jednoduchou neurčitou udalosťou, ktorá vznikla redukciou zloženej neurčitej udalosti pomocou štandardných operácií s pravdepodobnosťami.

6*Substitúcia.* Rozhodovateľ je indiferentný medzi pôvodnou neurčitou udalosťou, ktorá obsahuje výstup A, a neurčitou udalosťou, v ktorej je výstup A nahradený takou neurčitou udalosťou B, že rozhodovateľ je indiferentný medzi A a B.

7*Monotónnosť.* Keď existujú dve referenčné hry s rovnakými výstupmi, rozhodovateľ preferuje hru s väčšou pravdepodobnosťou získania preferovaného výstupu.

8*Invariancia.* Na určenie preferencií rozhodovateľa medzi neurčitými udalosťami sú potrebné výplaty (alebo dôsledky) a príslušné pravdepodobnosti.

9*Konečnosť.* Neuvažujeme o nekonečne dobrých alebo zlých dôsledkoch.

Možno asi súhlasiť s tým, že tieto axiómy predstavujú rozumný základ na rozhodovanie.

Keď akceptujete uvedené axiómy 1 až 8, mali by ste prijať aj tento návrh:
Majme dve neurčité udalosti B_1 a B_2. Keď axiómy 1 až 8 platia, existujú čísla $u_1, u_2, ..., u_n$, ktoré reprezentujú preferencie (alebo užitočnosti) spojené s dôsledkami tak, že celková preferencia medzi neurčitými udalosťami je vyjadrená očakávanými hodnotami u. Ak teda akceptujete axiómy 1 až 8, potom môžete nájsť funkciu užitočnosti, pomocou ktorej ohodnotíte dôsledky, a rozhodovať by ste sa mali na základe maximalizácie očakávanej užitočnosti.

2 Viackriteriálne rozhodovanie

Postup viackriteriálneho rozhodovania sa pokúsime ilustrovať na jednoduchom príklade z oblasti aplikácie Taguchiho metód.

Príklad. Uvažujeme 4 faktory: F_1, F_2, F_3, F_4 a dve ozvy K_1, K_2 – výkonové charakteristiky typu „the smaller the better“, ktoré považujeme za kritériá rozhodovania. Experiment sa realizoval podľa Taguchiho ortogonálneho návrhu $L_9\left(3^4\right)$. V tabuľke 1 je uvedený návrh $L_9\left(3^4\right)$ (namiesto číslic 1, 2, 3 sme na označenie úrovní faktorov použili číslice (– 1), 0 a 1) a výsledky experimentu – hodnoty kritérií – oziev K_1 *a* K_2, v posledných dvoch stĺpcoch

Tabuľka 1. *Návrh a výsledky experimentu*

Pokus číslo	F_1	F_2	F_3	F_4	K_1	K_2
1	– 1	–1	– 1	– 1	350,81	22,14
2	– 1	0	0	0	353,53	26,69
3	– 1	1	1	1	346,05	22,19
4	0	–1	0	1	348,51	26,63
5	0	0	1	–1	348,70	25,77
6	0	1	–1	0	351,57	25,89
7	1	–1	1	0	354,33	17,59
8	1	0	– 1	1	345,21	17,77
9	1	1	0	–1	350,00	18,59

V súlade s Taguchiho koncepciou nebudeme uvažovať interakcie. Pre každý faktor separovane odhadneme pre každú úroveň veľkosť straty[4], pomocou hodnoty výberového priemeru.

Výsledky sú v tabuľke 2. Výraznejším písmom sú v tabuľke 2 zapísané minimálne hodnoty, pre kombinácie kritérium – úroveň faktora. V tabuľke 2 vidieť, že ani jeden faktor nemá optimálnu tú istú úroveň pre jedno aj druhé kritérium.

Tabuľka 2. *Straty pre kombinácie kritérií a úrovní faktorov*

	K_1				K_2			
faktory úrovne	F_1	F_2	F_3	F_4	F_1	F_2	F_3	F_4
– 1	350,13	351,22	**349,20**	349,84	23,67	**22,12**	21,93	**22,17**
0	**349,59**	**349,15**	350,68	353,14	26,10	23,41	23,97	23,39
1	349,85	349,21	349,69	**346,59**	**17,98**	22,22	**21,85**	22,20

2**Aditívna funkcia užitočnosti**

Viackriteriálnu funkciu užitočnosti môžeme definovať ako aditívnu:

$$u_a = w_1 K_1 + w_2 K_2$$

kde w_1, w_2 sú váhy kritérií.

Sú známe viaceré metódy na ohodnotenie váh kritérií: metóda rovnakého oceňovania (pricing out), metóda výkyvného váženia (swing-weighting) a metóda porovnania lotérií (lottery weights)[5].

Pri aplikácii aditívnej funkcie užitočnosti sa v prvom kroku normujú hodnoty oboch kritérií pre jednotlivé pokusy podľa vzťahu[6]

[4] Čím väčšia je hodnota kritéria, tým väčšia je hodnota straty, pretože cieľová hodnota každého kritéria je 0.

[5] Pozri napr. CLEMEN (1991), CLEMEN – REILLY (2001), TEREK (2006).

[6] Pozri MILANI, A. S. – EL-LAHHAM, C. – NEMES, J. A. (2004).

$$NK_{il} = \frac{K_{il}}{\sum_{l=1}^{9} K_{il}} \qquad \text{pre } i = 1, 2 \ ; l = 1, 2, ..., 9$$

kde i je index kritéria a l je index pokusu.

Potom možno hodnoty aditívnej funkcie užitočnosti pre jednotlivé pokusy vypočítať podľa vzťahu

$$Nu_l = w_1 NK_{1l} + w_2 NK_{2l} \qquad \text{pre } l = 1, 2, ..., 9$$

Predpokladajme, že podľa niektorej z uvedených metód stanovenia váh kritérií ste určili $w_1 = 0{,}8$, $w_2 = 0{,}2$. V tabuľke 3 sú pre jednotlivé pokusy normované hodnoty kritérií a hodnoty aditívnej funkcie užitočnosti.

Tabuľka 3. *Normované hodnoty kritérií a hodnoty aditívnej funkcie užitočnosti*

pokus číslo	NK_{1l}	NK_{2l}	Nu_l pre $w_1 = 0{,}8$, $w_2 = 0{,}2$
1	0,1114	0,1089	0,1109
2	0,1123	0,1313	0,1161
3	0,1099	0,1092	0,1098
4	0,1107	0,1310	0,1148
5	0,1107	0,1268	0,1139
6	0,1117	0,1274	0,1148
7	0,1125	0,0865	0,1073
8	0,1096	0,0874	0,1052
9	0,1112	0,0915	0,1073

Potom sa pre každý faktor vypočíta priemerná hodnota aditívnej funkcie užitočnosti pre každú úroveň faktora a nájdu sa optimálne úrovne faktorov. Výsledky výpočtov sú v tabuľke 4.

V tabuľke 4 sú výraznejším písmom uvedené minimálne hodnoty priemerov aditívnej funkcie užitočnosti v jednotlivých stĺpcoch, ktoré určujú optimálne úrovne faktorov.

Tabuľka 4. *Hodnoty priemerov aditívnej funkcie užitočnosti pre kombinácie úrovní faktorov*

pre váhy	$w_1 = 0{,}8$, $w_2 = 0{,}2$			
faktory úrovne	F_1	F_2	F_3	F_4
– 1	0,1123	0,1110	**0,1103**	0,1107
0	0,1145	0,1117	0,1127	0,1127
1	**0,1066**	**0,1106**	0,11033	**0,1099**

Optimálna kombinácia úrovní faktorov je

$(F_1, F_2, F_3, F_4)^o = (1, 1, -1, 1)$

Nutnou podmienkou použitia aditívnej (ordinálnej) funkcie užitočnosti v podmienkach určitosti je vzájomná nezávislosť kritérií vzhľadom na preferencie[7]. Keď rozhodujeme v podmienkach neurčitosti, je nevyhnutné splniť silnejšiu podmienku.

3 Určitosť, neurčitosť a funkcia užitočnosti

Keď ide o model rozhodovania v podmienkach určitosti, používa sa niekedy namiesto termínu funkcia užitočnosti termín *hodnotová funkcia* (*value function*) alebo *ordinálna funkcia užitočnosti* (*ordinal utility function*)[8]. Hodnotová funkcia vyžaduje len usporiadanie určitých (deterministických) výstupov podľa preferencie rozhodovateľa. Neuvažuje o lotériách alebo neurčitých výstupoch.

Keď ide o rozhodovanie v podmienkach neurčitosti, mal by sa používať skôr termín *kardinálna funkcia užitočnosti*[9]. Kardinálna funkcia užitočnosti zohľadňuje vzťah rozhodovateľa k riziku tak, že lotérie sú usporiadané spôsobom, ktorý je konzistentný so vzťahom rozhodovateľa k riziku. Všeobecne ide vlastne o porovnávanie rozdelení pravdepodobnosti alebo lotérií.

V našom príklade ide len o usporiadanie deterministických výstupov[10].

Všetky metódy na ohodnotenie váh, ktoré sme spomenuli sú vhodné v podmienkach určitosti aj neurčitosti.

Záver

Zaoberali sme sa len aditívnou funkciou užitočnosti, ktorá je aditívnou kombináciou preferencií pre individuálne funkcie užitočnosti. Možno uvažovať aj o *multiplikatívnej funkciu užitočnosti* (*multiplicative utility function*), prípadne iných formách funkcie užitočnosti[11]

Literatúra

[7] Hovoríme, že kritérium *Y* je nezávislé od kritéria *X* vzhľadom na preferencie, keď preferencie pre špecifikované výstupy *Y* nezávisia od hodnoty kritéria *X*.

[8] Pozri napr. CLEMEN – REILLY (2001), BEROGGI (1999).

[9] CLEMEN – REILLY (2001), s.620.

[10] Niekedy možno pri odhadovaní výstupov využiť aj tolerančné intervaly (pozri GARAJ – JANIGA (2005)).

[11] Pozri napr. MILANI, A. S. – EL-LAHHAM, C. – NEMES, J. A. (2004), MIETTINEN, K. M. (1998), KEENEY, R. L. – RAIFFA, H. (1976).

BEROGGI, G. E. G.: *Decision Modeling in Policy Management. An Introduction to the Analytic Concepts.* Norwell, Massachusetts: Kluwer Academic Publishers. 1999.

CLEMEN, R. T.: *Making Hard Decisions: An Introduction to Decision Analysis.* Boston: PWS – Kent, 1991.

CLEMEN, R. T. – REILLY, T.: *Making Hard Decisions with Decision Tools*. USA: Duxbury Thomson Learning, 2001.

GARAJ, I. – JANIGA, I.: *Jednostranné tolerančné medze normálneho rozdelenia s neznámou strednou hodnotou a rozptylom. One sided tolerance limits of normal distribution with unknown mean and variability.* Bratislava: STU, 2005.

KEENEY, R. L. – RAIFFA, H.: *Decisions with Multiple Objectives.* New York: J. Wiley and Sons, 1976.

KEENEY, R. L. – RAIFFA, H.: *Decisions with Multiple Objectives: Preferences and Value Tradeoffs.* New York: J. Wiley and Sons, 1993.

MIETTINEN, K. M.: *Nonlinear Multiobjective Optimization.* Norwell Massachusetts: Kluwer Academic Publishers, 1998.

MILANI, A. S. – EL-LAHHAM, C. – NEMES, J. A.: *Different Approaches in Multiple-Criteria Optimization using the Taguchi Method: A Case Study in a Cold Heading Process*. Journal of Advanced Manufacturing Systems, Vol. 3, No 1 (2004), 53 – 68.

TEREK, M.: *Úvod do analýzy rozhodovania a bayesovskej indukcie*. Bratislava: Ekonóm, 2000.

TEREK, M.: *Viackriteriálne rozhodovanie v Taguchiho metóde*. Slovenská štatistika a demografia 2/2006. s. 21 – 39.

Kontakt: doc. Ing. Milan Terek, PhD., Katedra štatistiky, Ekonomická univerzita, Dolnozemská cesta 1, 852 35 Bratislava. tel.: 02-67295713, e-mail: terek@dec.euba.sk

Skúsenosti s výučbou štatistických predmetov s podporou SPSS na EF UMB

Vladimír Úradníček, Gabriela Nedelová

Abstract: Since the academic year 2006/2007 we have started to teach statistics and related courses at our faculty following new study programmes. Lectures, exercises and seminars of these courses were runned with a support of newly implemented statistical software SPSS version 13.0 (Statistical Package for Social Sciences). Totally 1444 students were taught following the reformed courses. The reform got a financial support by a grant No. 11230100060 of the European Social Fund. In the present article we briefly dissucuss outputs of the reform.

Key words: statistics, teaching, SPSS

1. Úvod

Od akademického roku 2005/2006 začalo zrekonštruované oddelenie štatistiky a ekonomickej analytiky Katedry kvantitatívnych metód a informatiky etapu modernizácie a transformácie výučby aj predmetov štatistického zamerania na Ekonomickej fakulte UMB. Cieľom tejto modernizácie a transformácie je v období do septembra 2009 najmä :

- modernizácia výučby základných a derivovaných kurzov štatistiky s využitím novej softvérovej podpory (SPSS) a
- reforma obsahovej náplne jednotlivých predmetov.

2. Východiská reformy a modernizácie výučby štatistických predmetov

Hlavným cieľom reformy obsahu vyučovaných štatistických predmetov na EF UMB je zabezpečiť vyššiu kvalitu získaných poznatkov vo väzbe na moderné ponímanie podstaty ekonomickej vedy (pozri napr. články a najmä diskusiu odbornej verejnosti na www.etrend.sk – Bieda slovenskej ekonómie (november 2006) - Bieda kritikov slovenskej ekonómie (Diskusia) (december 2006); Fórum: Čo (nie) je ekonómia, kto (nie) sú ekonómovia; Posledná utópia (Diskusia) (december 2006).
Medzi základné parciálne ciele reformy patria:

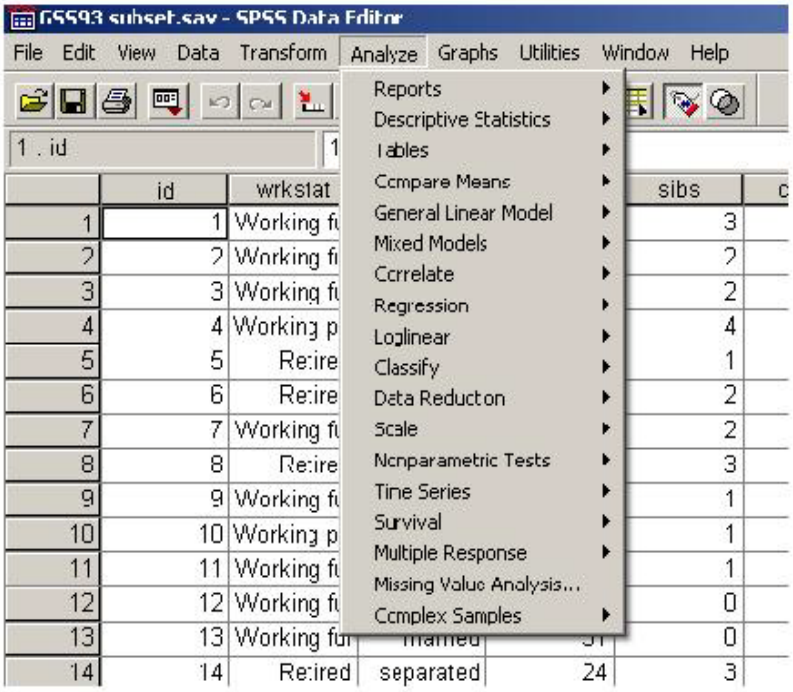
- primerané zavedenie novej softvérovej podpory (SPSS) do výučby štatistických predmetov;
- štandardizovanie učiva pri zachovaní špecifík osobnosti jednotlivých vyučujúcich (aby absolvovanie týchto kurzov u učiteľov Ekonomickej fakulty UMB bolo atraktívne aj pre študentov z prostredia mimo EF UMB);
- prehodnotenie spôsobu výkladu učiva – snaha o maximálne priblíženie možnosti využitia štatistiky ako nástroja v ekonomickej praxi, vo väzbe na profil absolventa príslušného študijného programu a ambície fakulty obstáť pri nadchádzajúcej komplexnej akreditácii ako pracovisko univerzitného typu. S tým úzko súvisí aj nová koncepcia preverovania vedomostí a zručností, ktoré študent pri štúdiu štatistiky získava (nejde len o memorovanie, ale o schopnosť študenta tvorivo riešiť nastolované problémy pomocou poznatkov a zručností, ktoré získa v procese štúdia);
- vedenie študenta k náročnosti voči vyučujúcemu, ale aj voči sebe samému; priviesť študenta k poznaniu, že úspech je vo veľkej miere determinovaný aj primeraným

stupňom samoštúdia a aktívneho využívania moderných metód vyučovania; všestranné reálne (nielen verbálne) naplňovanie jednotlivých atribútov kreditového systému štúdia (napr. aj priebežná práca počas celého semestra);
- snaha o vyvolanie záujmu študenta o preniknutie do hĺbky, objavenie krásy ekonómie ako vedy, nájdenie a porozumenie súvislostí a nadväzností, zapájanie študentov do vedeckých projektov oddelenia, ich účasť na pravidelne organizovaných vedeckých seminároch oddelenia štatistiky, prezentácia študentov EF UMB na odborných seminároch a konferenciách štatistického zamerania doma a v zahraničí (príkladom je aktívna účasť študentov odboru Financie, bankovníctvo a investovanie na Prehliadke prác mladých demografov a štatistikov v rámci Medzinárodného semináru Výpočtová štatistika v Bratislave v rokoch 2005 a 2006; zapojenie študentov do prípravy medzinárodnej konferencie FernStat organizovanej oddelením štatistiky v spolupráci so Slovenskou štatistickou a demografickou spoločnosťou, aktívne zapájanie študentov do projektov VUGA a od roku 2007 aj projektov typu VEGA);
- zabezpečenie všetkých základných kurzov štatistiky vlastnou literatúrou prednášajúcich (literatúru ponúkať prostredníctvom štandardnej celorepublikovej knižnej distribučnej siete, a nielen v predajniach s lokálnym významom) v slovenskom jazyku (do roku 2007), vo svetovom jazyku (do roku 2009),
- priebežné prenášanie skúseností z ekonomicko-finančnej praxe do vyučovacieho procesu – zohľadňovanie potrieb praxe na základné vedomosti a zručnosti z oblasti štatistických metód. Viacerí členovia oddelenia štatistiky a ekonomickej analytiky KKMI majú konkrétne skúsenosti zo spolupráce s odbornou praxou – školenie zamestnancov daňových úradov, realizácia analýz škodovej kvóty komerčnej poisťovne, školenia úverových analytikov slovenských komerčných bánk, realizácia analýz a modelovania rizika priemyselných havárií a i.
- jednoznačná profilácia vedeckej činnosti členov oddelenia – osobitne prostredníctvom jednotlivých vedeckých projektov (snaha o získanie projektov VEGA, KEGA a najmä medzinárodných vedeckých projektov), hosťujúce prednášanie na iných vysokých školách najmä v zahraničí, publikovanie výstupov vedeckej práce (kvantita zabezpečená v štandardných typoch publikácií – zborníky a odborná tlač; kvalita – aspoň jeden článok v karentovaných časopisoch za obdobie dvoch – troch rokov – v závislosti od typu, regionálnej príslušnosti a jazykovej mutácie karentovaného časopisu), aktivizácia členov oddelenia pri svojej popularizácii v odbornej domácej a zahraničnej štatistickej komunite, organizovanie pravidelnej medzinárodnej vedeckej konferencie (FernStat). Zvyšovanie kvalifikácie učiteľov by malo mať odraz aj vo zvyšovaní kvality nielen ich vedeckej, ale aj pedagogickej činnosti.

3. Výučba štatistických predmetov s podporou SPSS

V rámci projektu Európskeho sociálneho fondu: Modernizácia vzdelávania učiteľov prírodných vied a ekonómov, na ktorého realizácii v období rokov 2005 – 2007 participujú aj zamestnanci oddelenia štatistiky a ekonomickej analytiky Katedry kvantitatívnych metód a informatiky EF UMB, bol za účelom modernizácie výučby štatistických predmetov zakúpený softvérový produkt SPSS. Ako uvádza spoločnosť SPSS CR, spol. s r.o. na svojej webovskej stránke (www.spss.cz) „Spoločnosť SPSS CR, spol. s r.o. bola založená v roku 1995 a je výhradným distribútorom softvéru SPSS, vytvoreným nadnárodnou spoločnosťou SPSS Inc., a poskytovateľom analytických a štatistických služieb v Českej a Slovenskej republike. Medzi základné produktové rady patrí systém SPSS pre štatistické aplikácie, ktorý je modulárny a umožňuje napojiť ďalšie externé programové doplnky.“ Od tohto akademického roku bol systém SPSS 13.0 prirodzeným spôsobom využívaný vo výučbe

viacerých predmetov kvantitatívneho charakteru. Celkovo bol uvedený softvérový produkt využitý v tomto akademickom roku pri výučbe 1433 študentov dennej a externej formy štúdia v predmetoch Štatistika (1134 študentov), Ekonometria (45), Finančná analytika (36), Finančno-ekonomická analýza podniku – (168), Kvantitatívny manažment (40) a Statistics (10). Z základného modulu, ktorého základný nástroj *Analyse* znázorňuje obrázok 1, boli využívané pre jednotlivé predmety najmä nasledovné položky ponuky Analyse:
v predmetoch Štatistika a Statistics – Descriptive Statistics, Tables, Compare Means, Correlate, Regression, Time Series;
v Ekonometrii, Finančnej analytike, Finančno-ekonomickej analýze podniku a v Kvantitatívnom manažmente – Descriptive Statistics, Compare Means, General Linear Model, Correlate, Regression, Loglinear, Classify, Data Reduction, Nonparametric Tests, Complex Samples.



Obrázok 1 Dialógové okno SPSS Data Editor
Prameň: Vlastné spracovanie

Využívanie nástrojov produktu SPSS našlo pozitívnu odozvu medzi študentami. Študenti ocenili, že pri ich výučbe sa využíva jeden z celosvetovo najrozšírenejších štatistických softvérov pre stolové PC.

4. Analýza výsledkov úspešnosti študentov z predmetov Štatistika 1 a Štatistika 2

Tabuľka 1 udáva prehľad dosiahnutých výsledkov za obdobie rokov 2002/2003 až 2006/2007 študentmi EF UMB. Podľa študijného poriadku fakulty študent získa kredity za predmet („vyhovel“), ak z cekového počtu bodov za jednotlivé formy hodnotenia študijných výsledkov v rámci predmetu dosiahne aspoň 65 precent. Tabuľky 1 – 4 uvádzajú len výsledky hodnotenia študentov dennej formy štúdia.

Tabuľka 1 Absolútne vyjadrenia výsledkov z predmetu Štatistika 1

	Počet študentov			
Školský rok	vyhoveli	nevyhoveli	nezúčastnili sa	spolu
2002/2003	117	47	39	203
2003/2004	150	72	34	256
2004/2005	128	66	84	278
2005/2006	232	152	153	537
2006/2007	379	216	153	748

Prameň: Vlastné spracovanie

Tabuľka 2 Percentuálne vyjadrenia výsledkov z predmetu Štatistika 1

	Podiel študentov			
Školský rok	vyhoveli	nevyhoveli	nezúčastnili sa	spolu
2002/2003	57.64	23.15	19.21	100
2003/2004	58.59	28.13	13.28	100
2004/2005	46.04	23.74	30.22	100
2005/2006	43.20	28.31	28.49	100
2006/2007	50.67	28.88	20.45	100

Prameň: Vlastné spracovanie

Tabuľka 3 Absolútne vyjadrenia výsledkov z predmetu Štatistika 2

	Počet študentov			
Školský rok	vyhoveli	nevyhoveli	nezúčastnili sa	spolu
2002/2003	76	22	30	128
2003/2004	56	26	11	93
2004/2005	33	37	27	97
2005/2006	128	54	14	196
2006/2007	228	70	14	312

Prameň: Vlastné spracovanie

Tabuľka 4 Percentuálne vyjadrenia výsledkov z predmetu Štatistika 2

	Podiel študentov			
Školský rok	vyhoveli	nevyhoveli	nezúčastnili sa	spolu
2002/2003	59.38	17.19	23.44	100.00
2003/2004	60.22	27.96	11.83	100.00
2004/2005	34.02	38.14	27.84	100.00
2005/2006	65.31	27.55	7.14	100.00
2006/2007	73.08	22.44	4.49	100.00

Prameň: Vlastné spracovanie

Vzhľadom na skutočnosť, že od akademického roku 2005/2006 zahŕňa analýza aj študentov študijného odboru Financie, bankovníctvo a investovanie (po zlúčení s Fakultou financií),

zaujímala nás štatistická signifikancia rozdielov v dosiahnutých výsledkoch práve medzi akademickými rokmi 2005/2006 a 2006/2007 (zavedenie SPSS do výučby).

Tabuľka 5 Pomocné výpočty

Predmet Štatistika 1	Počet študentov			Podiel študentov		
v škol. roku	vyhoveli	nevyhoveli	spolu	vyhoveli %	nevyhoveli %	spolu %
2005/2006	360	337	697	51.65	48.35	100.00
2006/2007	379	216	595	63.70	36.30	100.00
Predmet Štatistika 2						
2005/2006	161	216	377	42.71	57.29	100.00
2006/2007	228	70	298	76.51	23.49	100.00

Prameň: Vlastné spracovanie

Teoretický základ:

Nech X je náhodná premenná z binomického rozdelenia Bi (m_1, π_1) a Y je náhodná premenná z binomického rozdelenia Bi (m_2, π_2); X a Y sú nezávislé náhodné premenné.

Odhadom π_1 je $p_1 = \frac{X}{m_1}$; odhadom π_2 je $p_2 = \frac{Y}{m_2}$.

Ďalej platí:

$$E[p_1 - p_2] = \pi_1 - \pi_2; \quad D[p_1 - p_2] = \frac{\pi_1(1-\pi_1)}{m_1} + \frac{\pi_2(1-\pi_2)}{m_2}.$$

Ak predpokladáme, že parametre binomických rozdelení sú rovnaké, $\pi_1 = \pi_2 = \pi$, potom testujeme nasledovné hypotézy:

Hypotézy: 1. .H_0: $\pi_1 = \pi_2$ 2. H_0: $\pi_1 = \pi_2$ 3. H_0: $\pi_1 = \pi_2$

H_A: $\pi_1 \neq \pi_2$ H_A: $\pi_1 > \pi_2$ H_A: $\pi_1 < \pi_2$

Odhadom π je $p = \frac{X+Y}{m_1+m_2}$.

Za predpokladu platnosti H_0 je $E[p_1 - p_2] = 0$ a

$$D[p_1 - p_2] = \frac{\pi(1-\pi)}{m_1} + \frac{\pi(1-\pi)}{m_2} = \pi(1-\pi).\left(\frac{1}{m_1} + \frac{1}{m_2}\right).$$

Testovacia štatistika

$$U = \frac{p_1 - p_2}{\sqrt{p(1-p)\left(\frac{1}{m_1} + \frac{1}{m_2}\right)}}$$

má za predpokladu platnosti H_0 asymptoticky (pre dostatočne veľké m_1 a m_2) N(0,1) rozdelenie.

Nakoľko v našom prípade sú splnené požiadavky na korektnosť testu o zhode parametrov π_1 a π_2 dvoch binomických rozdelení (o zhode dvoch relatívnych početností) testovali sme hypotézu, že po modifikovaní výučby predmetov Štatistika 1 a Štatistika 2 sa vo všeobecnosti zvýšil podiel úspešných študentov v danom predmete (t.j. tých, ktorí z predmetu vyhoveli).

H_0: $\pi_1 = \pi_2$ vs. H_A: $\pi_1 < \pi_2$, kde π_1 zodpovedá „starému“ systému výučby (reprezentujú ho výsledky akademického roku 2005/2006) a π_2 zodpovedá „novému“ systému výučby (reprezentujú ho výsledky akademického roku 2006/2007). Výsledky testu uvádza tabuľka 6.

Tabuľka 6 Výsledky testu

Štatistika 1	p_1	p_2	p	u	p-hodnota
	0.52	0.64	0.57	-4.36	0
Štatistika 2	p_1	p_2	p	u	p-hodnota
	0.43	0.77	0.58	-8.83	0

Prameň: Vlastné spracovanie

Z výsledkov vyplýva, že v oboch prípadoch môžeme na všetkých bežných hladinách významnosti zamietnuť nulovú hypotézu o rovnosti oboch parametrov π_1 a π_2, t.j. potvrdil sa predpoklad, že inovácia výučby predmetov (aj pomocou zavedenia softvérového produktu SPSS do výučby) viedla vo všeobecnosti k zvýšeniu podielu úspešných študentov v danom predmete.

5. Záver

Z percentuálnej analýzy výsledkov úspešnosti študentov z predmetu Štatistika 1 (tabuľka 2), resp. Štatistika 2 (tabuľka 4) vyplýva, že po realizácii prvých zmien vo výučbe uvedených predmetov sa podarilo zastaviť nepriaznivý klesajúci trend počtu študentov, ktorí vyhoveli z daných predmetov. Štatistickú signifikanciu rozdielu v podiele úspešných študentov „pred“ a „po“ uskutočnených zmenách potvrdili aj výsledky testu v tabuľke 6. Tieto výsledky ešte umocňuje skutočnosť, že napriek zhoršujúcej sa kvalite prijímaných študentov na štúdium, došlo pri nezmenených požiadavkách na získanie príslušného hodnotenia, k zvýšeniu podielu úspešných študentov v sledovaných predmetoch.

6. Literatúra

KAŠČÁKOVÁ. A. – KRÁĽ, P. – KULČÁR., L. – NEDELOVÁ, G. – BALCIAR, A. 2005. *Štatistika na EF UMB v Banskej Bystrici*. In: Forum Statisticum Slovacum 3/2005, s. 26 – 35. ISSN 1336-7420.

KRÁĽ, P. – BALCIAR, A. 2006. *Dištančné vzdelávanie štatistických predmetov na EF UMB*. In: Forum Statisticum Slovacum 3/2006, s. 175 – 187. ISSN 1336-7420.

KANDEROVÁ, M. – ÚRADNÍČEK, V. 2007. *Štatistika a pravdepodobnosť pre ekonómov. 2. časť*. Banská Bystrica : OZ Financ, 2007. 190 s. ISBN 978-80-969535-1-6.

Adresa autorov:
Ing. Vladimír Úradníček, PhD. – RNDr. Gabriela Nedelová, PhD.
Ekonomická fakulta UMB
Katedra kvantitatívnych metód a informatiky
Oddelenie štatistiky a ekonomickej analytiky
Tajovského 10
975 90 Banská Bystrica
vladimir.uradnicek@umb.sk
gabriela.nedelova@umb.sk

Aplikácie numerického integrovania s podporou vybraných programových prostriedkov

Alena Vagaská

Fakulta výrobných technológií

Technická univerzita v Košiciach

Bayerova 1

080 01 Prešov

vagaska.alena@fvt.sk

Abstrakt

Po zavedení štruktúrovaného štúdia registrujeme na fakultách technických univerzít rapídny pokles počtu hodín určených pre matematické predmety. Keďže návrat vyššieho počtu hodín matematiky nie je reálny, autorka článku uvádza konkrétne možnosti počítačovej podpory výučby numerickej matematiky (s využitím Matlabu a Excelu), čo prispieva k riešeniu načrtnutej problematiky.

Kľúčové slová: numerické integrovanie, Matlab, MS Excel, blended learning

Úvod

Vzhľadom na náročnosť matematických výpočtov v technických aplikáciách sa využívanie výpočtovej a informačnej techniky v edukácii matematiky na technických univerzitách stáva nevyhnutnosťou a samozrejmosťou. V súlade so zavádzaným systémom manažérstva kvality je potrebné priebežne aktualizovať obsah aj formy vzdelávania a do výučby zavádzať nové, moderné prvky.

Pre uľahčenie štúdia matematiky v bakalárskych študijných programoch na mnohých fakultách technických univerzít zaznamenávame zvýšené úsilie učiteľov matematiky implementovať využívanie počítačov do vyučovacieho procesu. S ohľadom na kladný postoj novoprijatých študentov k počítačom sa predpokladá, že zavádzanie kombinovanej formy výučby - blended learning (kombinácia klasickej prezentačnej formy vyučovania – prednáška a cvičenia a e-learningovej podpory – elektronické materiály prístupné na internete či počítačovej sieti) urobí matematiku pre mnohých z nich atraktívnejšou, ale hlavne názornejšou a pochopiteľnejšou. Využívanie informačno-komunikačných technológií (IKT) totiž umožňuje sprostredkovať vizualizáciu predstáv o abstraktných matematických pojmoch – čím sa skracuje samotný proces učenia sa [2]. Pozitívne skúsenosti s implementáciou IKT v edukácii

numerickej matematiky majú učitelia na mnohých fakultách technických univerzít v Čechách a na Slovensku. Ich skúsenosti potvrdzujú, že sa osvedčila výučba metód numerickej matematiky s podporou vhodného počítačového programu (napr. Matlab, MS Excel, ...) na tzv. počítačových cvičeniach [1].

2 Numerické integrovanie (kvadratúra)

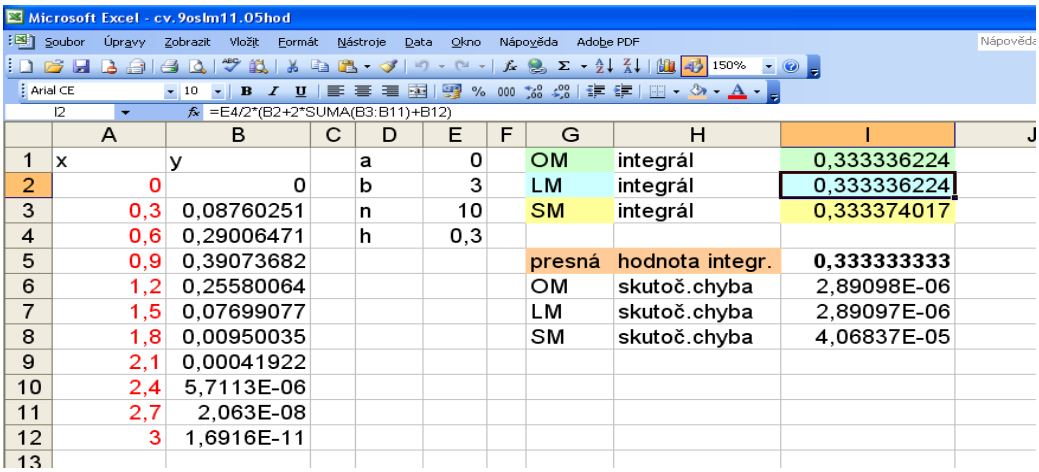
Keďže hodnotu určitého integrálu nezápornej funkcie f môžeme interpretovať ako obsah plochy pod jej grafom, numerická kvadratúra a numerické integrovanie označujú to isté: aproximovať určitý integrál $\int_a^b f(x)\,dx$ funkcie f na intervale $[a,b]$ na základe hodnôt funkcie f v konečnom počte bodov x_i intervalu $[a,b]$. Integrovanie pomocou známych Newton-Cotesových vzorcov spočíva v myšlienke rozdelenia intervalu $[a,b]$ sieťou bodov (x_i), $i=1,2,...,n$ pričom $a=x_1<x_2<...<x_n=b$ na jednotlivé podintervaly $[x_i, x_{i+1}]$, na ktorých potom následne aproximujeme funkciu f interpolačným polynómom P^i. Vďaka aditivite integrovania aproximáciu integrálu I na $[a,b]$ dostaneme sčítaním aproximácií získaných na podintervaloch $[x_i, x_{i+1}]$:

$$I=\int_a^b f(x)\,dx=\int_{x_1}^{x_2} f(x)\,dx+\int_{x_2}^{x_3} f(x)\,dx+...\int_{x_{n-1}}^{x_n} f(x)\,dx\approx\sum_{i=1}^{n-1}\int_{x_i}^{x_{i+1}} P^i(x)\,dx. \qquad (1)$$

Ak P^i je „dobrá“ aproximácia funkcie f na $[x_i, x_{i+1}]$, tak $\int_{x_i}^{x_{i+1}} P^i(x)\,dx$ je „dobrá“ aproximácia $\int_{x_i}^{x_{i+1}} f(x)\,dx$ a chyba integračného N-C vzorca je menšia. Pre rovnomernú sieť bodov (x_i) je $x_{i+1}-x_i=h=konšt.$, uzly interpolácie sú rozložené rovnomerne. V závislosti od stupňa s interpolačného polynómu P^i (písaného v Newtonovom či Lagrangeovom tvare) získavame odpovedajúce elementárne N-C vzorce: obdĺžnikový, lichobežnikový, Simpsonov, Booleov a Milneho vzorec (písané postupne pre $s=0,1,2,3,4$, kde $s=m-1$, m je počet uzlov interpolácie). Pre elementárne N-C vzorce môžeme písať všeobecný kvadratúrny vzorec $I=(b-a).\mathbf{w}.\mathbf{f}$ v ktorom $\mathbf{w}$ je riadkový vektor váh a $\mathbf{f}$ je stĺpcový vektor funkčných hodnôt (f-hodnôt) v uzloch interpolácie, t.j. $\mathbf{f}=[f_1, f_2, ..., f_m]^T$. Pre riadkový vektor váh platí: ak $m=2$, $w=\left[\frac{1}{2},\frac{1}{2}\right]$; ak $m=3$, $w=\left[\frac{1}{6},\frac{4}{6},\frac{1}{6}\right]$; ak $m=4$ $w=\left[\frac{1}{8},\frac{3}{8},\frac{3}{8},\frac{1}{8}\right]$, ak $m=5$, $w=\left[\frac{7}{90},\frac{32}{90},\frac{12}{90},\frac{32}{90},\frac{7}{90}\right]$. Aplikovaním elementárnych N-C vzorcov na podintervaly $[x_i, x_{i+1}]$ využitím vzťahu (1) získavame zložené N-C vzorce, uvedené napr. v [3].

3 Kvadratúra s využitím Excelu a Matlabu

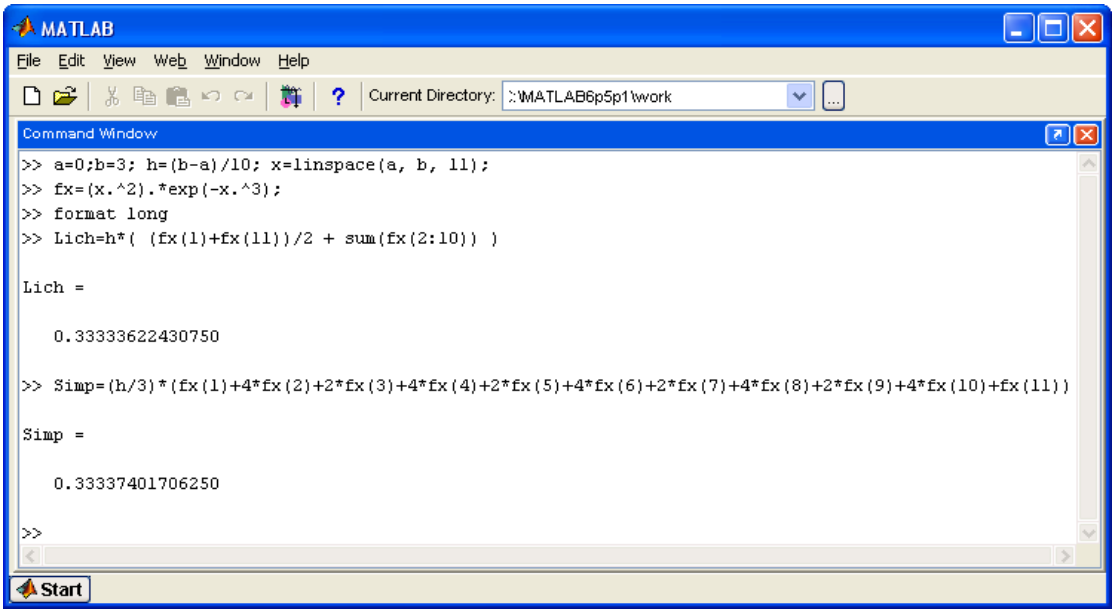
Vypočítajme aproximáciu integrálu funkcie $f(x)=x^2.e^{-x^3}$ na intervale $[a,b]=[0,3]$, keď použijeme rovnomernú sieť uzlov s krokom $h=0,3$. Zistíme, že pracujeme s desiatimi podintervalmi, pretože pre ich počet n platí: $n=\frac{(b-a)}{h}$. Ak sa rozhodneme použiť Excel, ktorého nepopierateľnou výhodou je jeho dostupnosť, do príslušných buniek Excelu (podľa obr. 1) si zapíšeme známe hodnoty a,b,h. Keďže Excel nepracuje s premennými, ale s odkazmi na bunky, do bunky E3 zapíšeme vzťah pre n v tvare =(E2-E1)/E4, čím získame počet podintervalov n. V stĺpci A si vytvoríme číselnú postupnosť s diferenciou určenou krokom h. Ak do bunky B2 vpíšeme vzťah =(A2^2)*EXP((-A2)^3), po odklepnutí cez Enter a kopírovaním obsahu tejto bunky (relatívny odkaz) dostaneme výpočet potrebných f-hodnôt v uzlových bodoch interpolácie. V bunkách I1, I2 a I3 sú v danom poradí vpísané zložené N-C vzorce: obdĺžnikový, ktorého zápis v Exceli je =E4*(SUMA(B3:B12)), lichobežníkový určený predpisom =E4/2*(B2+2*SUMA(B3:B11)+B12) a Simpsonov zložený vzorec v tvare =E4/3*(B2+4*B3+2*B4+4*B5+2*B6+4*B7+2*B8+4*B9+2*B10+4*B11+B12). Po potvrdení cez Enter získame potrebné výsledky. Tento integrál sa dá vypočítať aj analyticky, jeho presná hodnota je $(1-e^{-27})/3=0,\overline{3}$. Vieme teda zistiť, akej skutočnej chyby sa dopúšťame pri aproximácii podľa jednotlivých metód (obr. 1).



	A	B	C	D	E	F	G	H	I	J
1	x	y		a	0		OM	integrál	0,333336224	
2	0	0		b	3		LM	integrál	0,333336224	
3	0,3	0,08760251		n	10		SM	integrál	0,333374017	
4	0,6	0,29006471		h	0,3					
5	0,9	0,39073682					presná	hodnota integr.	**0,333333333**	
6	1,2	0,25580064					OM	skutoč.chyba	2,89098E-06	
7	1,5	0,07699077					LM	skutoč.chyba	2,89097E-06	
8	1,8	0,00950035					SM	skutoč.chyba	4,06837E-05	
9	2,1	0,00041922								
10	2,4	5,7113E-06								
11	2,7	2,063E-08								
12	3	1,6916E-11								
13										

Obr. 1 *Numerická integrácia v Exceli s daným krokom delenia, resp. počtom podintervalov*

K tým istým výsledkom sa dopracujeme využitím Matlabu, kde uvádzam lichobežníkovú a Simpsonovu metódu. Po naeditovaní príkazov (zrejmých z obr. 2) do jednotlivých promptov (príkazových riadkov) a odklepnutí cez Enter dostávame tie isté výsledky ako v Exceli.



Obr. 1 *Numerická integrácia v Matlabe s daným krokom delenia, resp. počtom podintervalov*

4 Kvadratúra s požadovanou presnosťou

Niektoré integrály nevieme vypočítať analyticky, napr. aj integrál z funkcie $f(x)=e^{-x^2}$, o ktorej je známe, že jej primitívnu funkciu nie je možné vyjadriť ako kombináciu elementárnych funkcií. V dôsledku toho nespoznáme skutočnú chybu, ktorej sa aproximáciou dopúšťame, môžeme získať len odhad chyby danej aproximácie. Vypočítajme obdĺžnikovou, lichobežníkovou a Simpsonovou metódou aproximáciu integrálu $\int_0^{1,4} e^{-x^2}\,dx$ v Matlabe, keď použijeme rovnomernú sieť s krokom $h=0,2$. Poznamenajme, že v obdĺžnikovej metóde využívame stredy podintervalov $z_1, z_2, \ldots, z_7$, čomu zodpovedá uvedený príkaz v prompte, ako aj hodnoty z a $f(z)$ po potvrdení.

```
>> z=0.1:0.2:1.4, fz=exp(-z.^2)
z =
    0.1000    0.3000    0.5000    0.7000    0.9000    1.1000    1.3000
fz =
    0.9900    0.9139    0.7788    0.6126    0.4449    0.2982    0.1845
>> format long
>> obd=0.2*sum(fz)
obd =
   0.84459661318441

>> format short
>> a=0; b=1.4; h=(b-a)/6; x=linspace(a, b, 7), fx=exp(-x.^2)
x =
         0    0.2333    0.4667    0.7000    0.9333    1.1667    1.4000
fx =
    1.0000    0.9470    0.8043    0.6126    0.4185    0.2564    0.1409
>> format long
>> Lich=h*( (fx(1) + fx(7))/2 + sum(fx(2:6)) )
Lich =
  0.84215435326511
```

```
>> a=0;b=1.4; h=(b-a)/7; x=linspace(a, b, 8);
>> fx=exp(-x.^2);
>> format long
>> Lich=h*( (fx(1)+fx(8))/2 + sum(fx(2:7)) )
Lich =
   0.84262767770934
```

V Simpsonovej metóde je potrebný párny počet podintervalov, príkazy sú následovne:

```
>> a=0; b=1.4; h=(b-a)/6; x=linspace(a, b, 7), fx=exp(-x.^2)
x =
     0   0.2333   0.4667   0.7000   0.9333   1.1667   1.4000
fx =
    1.0000   0.9470   0.8043   0.6126   0.4185   0.2564   0.1409
>> format long
>> Simp=(h/3)*(fx(1) + 4*fx(2)+2*fx(3)+4*fx(4)+2*fx(5)+4*fx(6)+fx(7))
Simp =
    0.84392718862057
```

Ak by sme chceli porovnať presnosť získaných výsledkov, môžeme daný integrál vypočítať s vopred stanovenou presnosťou ε, čo prezentuje ukážka výpočtov v Exceli (obr. 3.), kde sme získali výsledky s presnosťou $\varepsilon = 5.10^{-3}$ a $\varepsilon = 5.10^{-5}$. Nevýhodou kvadratúry so stanovenou presnosťou je nutnosť istej analytickej práce pri určovaní počtu podintervalov n zo vzťahov pre odhad chyby daného N-C vzorca:

Lichobežníková metóda: $$n \geq \sqrt{\frac{M_2 (b-a)^3}{12\varepsilon}}, \text{ kde } M_2 = \max_{x\in[a,b]} |f''(x)| \qquad (2)$$

Simpsonova metóda: $$n \geq \sqrt[4]{\frac{M_4 (b-a)^5}{180\varepsilon}}, \text{ kde } M_4 = \max_{x\in[a,b]} |f^{(4)}(x)| \qquad (3)$$

Potrebujeme totiž vyhodnotiť monotónnosť a ohraničenosť funkcie $|f^{(i)}(x)|$, avšak aj tu si môžeme pomôcť využitím uvedených programových prostriedkov.

Microsoft Excel - cv.9oslm11.05hod

Soubor Úpravy Zobrazit Vložit Formát Nástroje Data Okno Nápověda Adobe PDF

Arial CE 10

D53

	A	B	C	D	E	F	G	H
1	a	0						
2	b	1,4	LM	M2	2	n>=	9,563820715	
3	presnosť 1	0,005	SM 1	M4	12	n>=	2,910011657	
4	presnosť 2	0,00005	SM 2	M4	12	n>=	9,202264855	
5	LM	n=	10	h=	0,14			
6	SM 1	n=	4	h=	0,35			
7	SM 2	n=	10	h=	0,14			
8								
9	**presnosť 0,005**		**presnosť 0,005**		**presnosť 0,00005**			
10	LM		SM 1		SM 2			
11	x	y	x	y	x	y		
12	0	1	0	1	0	1		
13	0,14	0,980590831	0,35	0,884705905	0,14	0,980590831		
14	0,28	0,924594515	0,7	0,612626394	0,28	0,924594515		
15	0,42	0,838282603	1,05	0,332039945	0,42	0,838282603		
16	0,56	0,730811294	1,4	0,140858421	0,56	0,730811294		
17	0,7	0,612626394	**integrál**	0,843861038	0,7	0,612626394		
18	0,84	0,493812198			0,84	0,493812198		
19	0,98	0,382739759			0,98	0,382739759		
20	1,12	0,285246945			1,12	0,285246945		
21	1,26	0,204415621			1,26	0,204415621		
22	1,4	0,140858421			1,4	0,140858421		
23	**integrál**	0,843296912			**integrál**	0,843939094		

Obr. 3 *Numerická integrácia v Exceli s požadovanou presnosťou*

Po určení hodnôt M_2 a M_4 prepíšeme do bunky G2 vzťah (2) v tvare =(E2/(12*B3)*(B2-B1)^3)^(1/2) a do buniek G3 a G4 vzťah (3) v tvare =(((B2-B1)^5*E3)/(180*B3))^(1/4). Keďže chceme celočíselný počet podintervalov, pri Simpsonovej metóde dokonca párny, pri určovaní podintervalov využívame zaokruhľovanie, t.j. napr. pre určenie počtu podintervalov v lichobežníkovej metóde zapíšeme do bunky C5 formulu =ZAOKR.NAHORU(G2;1). V bunkách B23, D17 a F23 sú zapísané zložené N-C vzorce pre danú metódu, čím získame výsledky s požadovanou presnosťou.

Ak by sme sa vrátili k výsledkom aproximácie integrálu $f(x) = x^2.e^{-x^3}$ získaným v Matlabe, môžeme prehlásiť, že pri danom počte podintervalov dal obdĺžnikový vzorec presnejšiu hodnotu ako lichobežníkový, avšak nie ako Simpsonov vzorec.

Záver

Na základe skúsenosti s výučbou numerickej matematiky v počítačovej učebni môžeme vyjadriť presvedčenie, že kombinovanie klasických metód vyučovania s novými modernými metódami využivajúcimi vhodné programové systémy vedie k zefektívneniu vyučovacieho procesu.

Literatúra

[1] BAŠTINEC, J.- NOVÁK, M.: *Výuka numerických metód v navazujícím magisterském studiu na FEKT VUT Brno*. In.: Sborník 29. konference o matematice na VŠTEZ, Mutěnice, 4.- 8. 9.2006, Zlín: UTB ve Zlíně, 2006, s.17 –22. ISBN 80-7318-450-8

[2] BEISETZER, P. Samostatná práca edukanta a počítač. Prešov: FHPV v Prešove, 2005, 107 s., ISBN 80-8068-428-6

[3] VOLAUF, P. *Numerické a štatistické výpočty v Matlabe*. Bratislava: STU, 2005.ISBN 80-227-2259-6

Kontaktná adresa:

Vagaská Alena, PaedDr., PhD.
Katedra matematiky, informatiky a kybernetiky
Fakulta výrobných technológií TU v Košiciach, Bayerova 1, 080 01
Prešov, SR, tel. 00421 517 723 931, fax 00421 517 733 453, e-mail vagaska.alena@fvt.sk

Príklad aplikácie DoE pri hodnotení kvality povrchu pri delení nehrdzavejúcej ocele AISI 304 vysokorýchlostným hydroabrazívnym prúdom

Alena Vagaská, Sergej Hloch, Miroslav Gombár

Fakulta výrobných technológií

Technická univerzita v Košiciach

Bayerova 1

080 01 Prešov

vagaska.alena@fvt.sk, hloch.sergej@fvt.sk, mirek@unipo.sk

Abstrakt

Cieľom príspevku je overenie vhodnosti využitia metódy DoE v oblasti delenia materiálov vysokorýchlostným hydroabrazívnym prúdom, ktorá je výhodná pre firmy ktoré používajú ISO 9000, ktorá je v novom znení zameraná na vykonanie činností firiem založených na procesoch a ich zlepšovaní. Práve vo výskume sa ponúkajú plánované pokusy ako prostriedok k zlepšovaniu výskumnej činnosti.

Kľúčové slová: kvalita, plánované experimenty, vysokorýchlostný hydroabrazívny prúd

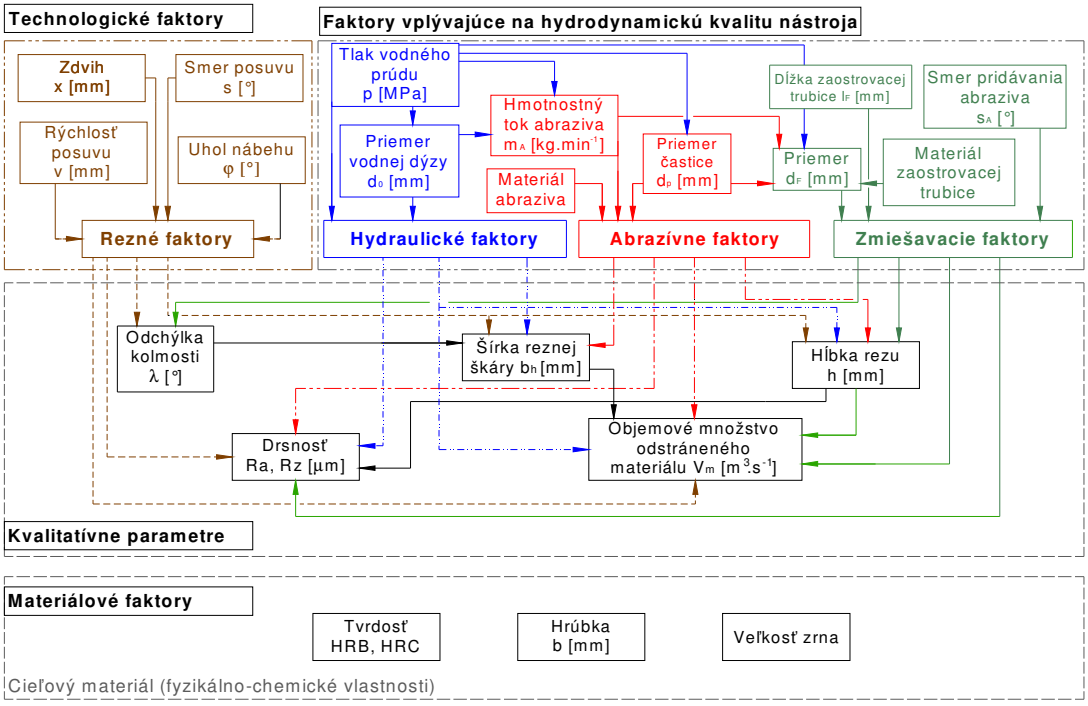
Úvod

Technologický proces prebiehajúci pri delení materiálov vysokorýchlostným hydroabrazívnym prúdom je v reálnych objektoch vo väčšine prípadov veľmi zložitý dynamický a stochastický proces. Analytický spôsob identifikácie tohto procesu sa javí ako neefektívny a málo praktický [9,10]. Jeho aplikovaním nie je možné získať úplný model procesu, pričom sa zanedbáva vplyv určitých veličín, nie sú známe presné hodnoty niektorých vstupných faktorov, ktoré sa navyše časom menia a v mnohých prípadoch riadiaci pracovník uplatňuje intuíciu, ktorú je možné nadobudnúť len pri mnohonásobnom vykonávaní danej činnosti, čo je z časového hľadiska neefektívne [11, 12]. Matematicko - štatistické metódy umožňujú zostavenie štatistických modelov aj z relatívne veľkého počtu vstupných údajov [13]. Hodnotenie kvality delenia vysokorýchlostným hydroabrazívnym prúdom v súčasnosti nie je uspokojivo rozpracované, hlavne čo sa týka aplikácie viacrozmerných štatistických metód, keďže súčasné analytické riešenia progresívnych technologických procesov sú postavené na intuitívnej a empirickej platforme riadiaceho technológa [2, 3, 5]. Pri delení materiálov vysokorýchlostným hydroabrazívnym prúdom je potrebné zabezpečiť požadované kvalitatívne charakteristiky obrobenej plochy, čo je podmienené znalosťou priebehu funkčných závislostí medzi parametrami kvality výrobku a procesnými faktormi výrobného systému čo má výrazný vplyv na optimalizáciu, zlepšovanie kvality

prevádzky progresívnych technologických systémov delenia. Tieto modely reálnych situácií naplnené konkrétnymi číslami a ich kauzalitu umožňujú analyzovať a vyhodnotiť plánované experimenty (Design of Experiments – DoE).

Charakteristika technologického procesu

Technologický proces delenia hydroabrazívnou eróziou sa uskutočňuje na výrobnom zariadení pomocou nástroja – vysokorýchlostného hydroabrazívneho prúdu, ktorého vlastnosti nie sú degradované v čase prevádzky na rozdiel od klasického obrábacieho rezného noža. Pomocou faktorov (obr. 1) sa vytvára obrobená plocha a tvorí ako obalová plocha trajektórie pracovného pohybu vysokorýchlostného hydroabrazívneho prúdu. Ide o zložitý špecifický spôsob obrábania charakteristický tým, že sa používa mnohoklinový nástroj, ktorého rezné kliny sú tvorené brúsnymi zrnami (abrazivom), ktoré sú v kvapaline náhodne orientované. Vysokorýchlostný hydroabrazívny prúd studeným spôsobom vytvára na obrobku reliéf, s dvoma zreteľnými oblasťami pozdĺž steny rezu, ktoré sú charakterizované odlišnou textúrou povrchu. Vytvorený reliéf sa z hľadiska kvality (posudzovanej pomocou parametra profilu drsnosti *Ra* povrchu vo zvislom smere) rozdeľuje na hornú eróznu zónu, ktorá sa vyznačuje nižšími číselnými hodnotami parametra profilu drsnosti *Ra* a na dolnú eróznu zónu, ktorá sa vyznačuje vyššími číselnými hodnotami parametra profilu drsnosti *Ra* [4, 6].



Obr. 1 Grafické znázornenie vzájomného pôsobenia faktorov na kvalitu obrobenej plochy.

Rozsah a veľkosť týchto zón závisí od faktorov AWJ [4]. Na kvalitatívne charakteristiky novovytvorenej plochy (obr. 1) a na úber materiálu vplýva množstvo faktorov, prostredníctvom pracovného nástroja a jeho kvalitatívnych charakteristík, ktoré sa rozdeľujú do dvoch skupín (priamych a nepriamych procesných faktorov)

ovplyvňujúcich aj okrem iných parametrov (odchýlka kolmosti, aj parameter profilu drsnosti Ra.

Experimentálna procedúra

Pri delení materiálov nepôsobia faktory AWJ samostatne (obr. 1), ale spoločne vo vzájomnej interakcii. Takéto procesy umožňujú analyzovať faktorové experimenty, pri ktorých sa zrealizujú pokusy pre všetky kombinácie úrovní uvažovaných faktorov [8]. Pre analýzu faktorov, a ich vzájomné závažnosti ich vplyvu na kvalitu delenej plochy sa využil plánovaný experiment, pri ktorom sa vykonali pokusy pre všetky kombinácie dvoch úrovní uvažovaných faktorov (tab. 1) pričom pri hodnotení uvedených faktorov sa použil kvalitatívny parameter profilu drsnosti Ra.

Tab.1 Zakódovanie faktorov do úrovní s priradenými skutočnými hodnotami.

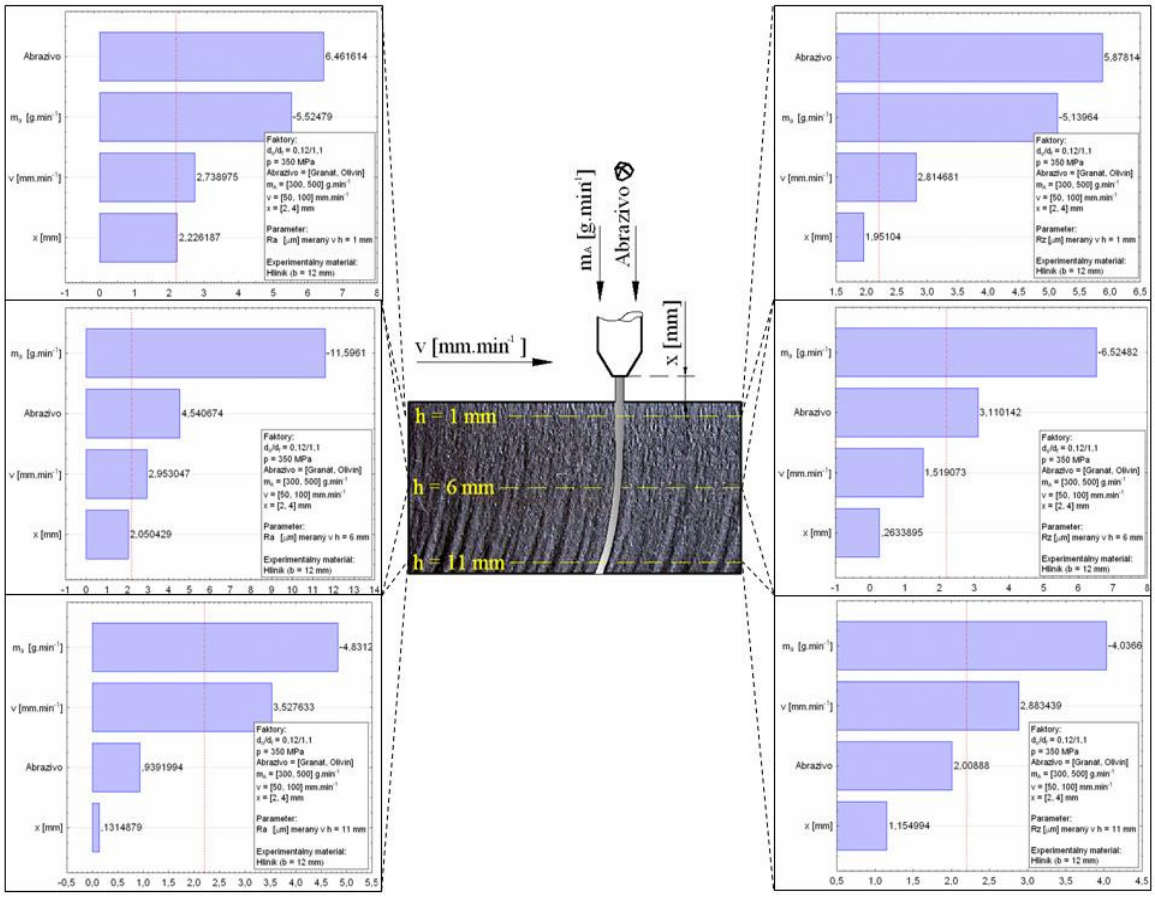
Č.	Faktory			Interval	
	Označenie	Názvoslovie	Rozmer	-1	+1
1	x_1	Tvar abraziva		Olivín	Granát
2	x_2	Hmotnostný tok	$g.min^{-1}$	300	500
3	x_3	Rýchlosť posuvu	$mm.min^{-1}$	50	100
4	x_4	Zdvih	mm	2	4

K štúdiu vplyvu procesných faktorov na kvalitu obrobenej plochy bol použitý presný dvojpolohovací stôl od firmy Wating, určený pre rovinné aplikácie delenia technológiou AWJ. Tlak vody bol vytváraný čerpadlom Stream Line SL II od firmy Ingersoll Rand. Ako technologická hlavica bola použitá hlavica firmy Ingersoll Rand AUTOLINE^{TM}. Všetky vzorky boli vyrobené s konštantným nastavením nasledujúcich faktorov: tlak čerpadla $p = 350\,\mathrm{MPa}$, MESH 80. Vzorky boli vyrobené z hliníka, ktorého hrúbka bola $12\,\mathrm{mm}$. Merania parametra profilu strednej aritmetickej drsnosti Ra boli vykonané na mobilnom prístroji SURFTEST – 301 v hĺbke $1\,\mathrm{mm}$ a $11\,\mathrm{mm}$.

Vyhodnotenie faktorového experimentu

Normalita rozloženia súboru opakovaných meraní parametra profilu - strednej aritmetickej odchýlky Ra a reziduí bola testovaná pomocou parametrického Shapiro – Wilksonovho testovacieho kritéria s použitím výpočtového systému STATISTICA, keď výsledkom bola hodnota W a hodnota p (dosiahnutá hladina pravdepodobnosti, teda pravdepodobnosť, že výberová hodnota odhadu testovaného parametra bude aspoň taká veľká ako pozorovaná, ak H_0 v skutočnosti platí.) Z uvedenej analýzy vyplynulo, že všetky opakované merania majú normálne Gaussove rozdelenie, čo umožnilo využitie parametrického Grubssov testu odľahlosti meraní a ich homoskedasticity. Pre vyhodnotenie skúmaných závislostí boli zvolené lineárne a linearizované 3D závislosti. Výpočet regresných koeficientov bol realizovaný v programe Excel, Matlab a Statistika. Pre jednotlivé typy regresných závislostí boli vypracované výpočtové programy, založené na metóde najmenších štvorcov a maticových operáciách. Pre posúdenie štatistickej významnosti regresných koeficientov bol použitý Studentov t – test. Testovacie kritérium t sa pre každý z regresných koeficientov porovnal s tabelovanou hodnotou kritéria na zvolenej hladine významnosti α. Významnosť sledovaných nezávislých premenných - faktorov je graficky znázornená v Paretových diagramoch

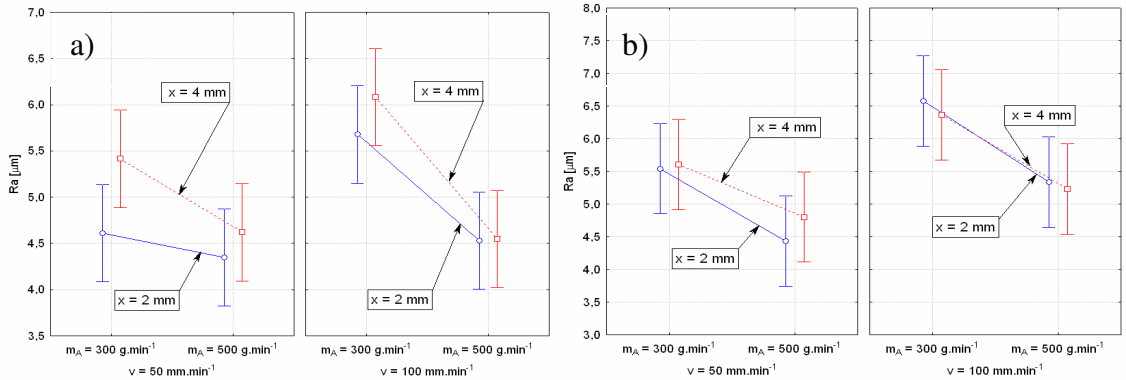
(obr. 2). V hĺbke 1 mm, kde sa zisťoval parameter profilu drsnosti Ra, sú dominantné podľa Studentovho kritéria všetky štyri faktory, po poradí: tvar abraziva, hmotnostný tok abraziva, rýchlosť posuvu a zdvih, vrátane významnej interakcie rýchlosti posuvu a hmotnostného toku abraziva.



Obr. 2. Vplyv faktorov na parametre Ra a Rz v posudzovaných hĺbkach pre materiál hliník pri šírke $b = 12\,\text{mm}$ vyjadrený pomocou Paretových grafov.

Podľa paretovej analýzy faktory tvar abraziva a rýchlosť posuvu majú kladný efekt, čo znamená, že so zvyšujúcou sa rýchlosťou posuvu a zvyšujúcou sa členitosťou abraziva číselné hodnoty drsnosti v meranej hĺbke $h = 1\,\text{mm}$ sa budú zvyšovať. Hmotnostný tok, so záporným efektom naznačuje, že so znižujúcim sa hmotnostným tokom abraziva, teda ak znížime množstvo abraziva do vysokorýchlostného prúdu vody, hodnoty drsnosti obrobenej plochy budú vyššie. V hĺbke $h = 11\,\text{mm}$ sa poradie a významnosť hodnotených faktorov zmenilo (Obr. 3). Najvýznamnejšími faktormi sú podľa Paretovho grafu 2 hmotnostný tok, rýchlosť posuvu, pričom zdvih a tvar abraziva sú podľa štatistického ohodnotenia nevýznamné, teda neovplyvňujú číselné hodnoty Ra vplyv na parameter profilu drsnosti Ra. Ako vidno významnosť faktorov, hmotnostného toku abraziva a rýchlosti posuvu, rastie s pribúdajúcou hrúbkou materiálu. Z paretovho grafu 3 vidno, že zlepšenie kvality povrchu možno dosiahnuť zo zväčšujúcim sa hmotnostným tokom abraziva m_A a znižujúcou sa rýchlosťou posuvu v. Vysoká variabilita drsnosti pri hodnotení troch faktorov (m_A, x, druh abraziva) bola

zistená v hĺbke $1\,\mathrm{mm}$ najmä pri rýchlosti posuvu $100\,\mathrm{mm.min^{-1}}$ v spojení s faktorom m_A čo je potvrdené aj paretovou analýzou, kde vidno aj pomerne veľký vplyv druhu abraziva na drsnosť povrchu. Pre porovnanie tej istej situácie v hĺbke $11\,\mathrm{mm}$ tvrdosť abraziva nemá významný vplyv na povrchovú drsnosť deleného materiálu. Významnosť hmotnostného toku rastie podľa grafu v nižších hĺbkach so zvyšujúcou sa rýchlosťou posuvu a zvyšujúcou sa hĺbkou. Čo sa týka druhu použitého abraziva ako hodnoteného faktora, za pozornosť stojí vývoj drsnosti v hĺbke $h = 11\,\mathrm{mm}$ pri rýchlosti posuvu $v = 100\,\mathrm{mm.min^{-1}}$. Ako vidno z danej situácie na obr. 3 vplyv druhu použitého abraziva na kvalitu vytvoreného povrchu je prakticky rovnaký.



Obr. 3. Vplyv zdvihu, hmotnostného toku abraziva a rýchlosti posuvu na parameter profilu drsnosti Ra v hĺbke a) $h = 1\,\mathrm{mm}$, b) $h = 11\,\mathrm{mm}$. Grafy vytvorené pri strednej hodnote sledovaného intervalu tvaru abraziva (olivín – Barton Garnet).

V nasledujúcich grafoch (Obr. 3a, 3b) je znázornený vplyv kombinácie faktorov (v, m_A, x). O vplyve vzdialenosti ústia zaostrovacej trubice a povrchu deleného materiálu sa popísalo v dielach [9], [10], [11], avšak doposiaľ nie je známa kvantifikácia tohto faktora s spolu s interakciami s inými faktormi technológie AWJ, pričom taktiež nie je známy jeho celkový podiel na celkovej významnosti pri hodnotení faktorov tejto technológie a jeho vplyv na kvalitu deleného materiálu. Nasledujúce grafy tento vplyv dokumentujú. Ako vidno tento faktor je významný v iniciačnej zóne pri rýchlosti posuvu $v = 50\,\mathrm{mm.min^{-1}}$ a $v = 100\,\mathrm{mm.min^{-1}}$ v hĺbke $1\,\mathrm{mm}$. V hĺbke $11\,\mathrm{mm}$ je významnosť tohto faktora vyššia pri rýchlosti posuvu $v = 50\,\mathrm{mm.min^{-1}}$ avšak číselné hodnoty Ra sú porovnateľné s číselnými hodnotami Ra v meranej hĺbke $1\,\mathrm{mm}$. Pri rýchlosti posuvu $v = 100\,\mathrm{mm.min^{-1}}$ v hĺbke $11\,\mathrm{mm}$ je vplyv tohto faktora nevýznamný čo dokumentuje graf, kde sa tieto lineárne priebehy takmer prekrývajú.

Záver

Na voľbu optimálneho typu faktorov, ktoré majú podstatný vplyv na kvalitatívne a výkonnostné hodnoty, môže v nových aplikáciách prispieť aj faktorová analýza. Predmetom článku bol príklad aplikácie plánovaných experimentov pri hodnotení vplyvu faktorov vysokorýchlostného hydroabrazívneho prúdu na parametre profilu drsnosti Ra a Rz povrchu hliníka. Plný faktorový experiment bol použitý za účelom zistenia vplyvu nezávislých premenných a ich interakcií na závislú premennú Ra a Rz v závislosti od hrúbky deleného materiálu. Kvalita povrchu je citlivejšia na nastavení

všetkých hodnotených faktorov, najmä od tvaru abraziva. Jedným z vysvetlení môže byť, že práve v tzv. iniciačnej zóne dochádza k prvému kontaktu hydrodynamického nástroja s deleným materiálom, v ktorej sú plytšie zárezy. Preto sú tam zreteľnejšie stopy po abrazívnych časticiach, najmä ich tvaru. V dôsledku už samotnej degradácie vysokorýchlostného hydroabrazívneho prúdu s narastajúcou hrúbkou materiálu sa významnosť faktorov zmenila. Ďalej je potrebné poukázať na to, že okrem zložitosti a prítomnosti množstva stochastických vplyvov v technológii delenia AWJ pre samotné hodnotenie povrchov vytvorených touto technológiou nestačí iba charakteristika parametra profilu drsnosti *Ra* teda iba len výšková charakteristika, pretože môžu existovať dva isté povrchy ktoré môžu mať rovnaké *Ra* , ale z hľadiska ich funkčnosti sú rozdielne [9]. Pretože okrem výškových parametrov, ktoré sú normované musíme uvažovať aj s parametrami pozdĺžnymi (rozstup) – frekvenčnými a aj uhlami sklonov. Preto podľa názoru autorov by bolo vhodné uvažovať o doplnení normovaných parametrov pre hodnotenie povrchov najmä o parametre pozdĺžnych deformácií (rozstup výškových nerovností) a priestorové frekvenčné charakteristiky a súčasne taktiež vypracovať normované postupy merania týchto parametrov.

Literatúra

[1] Martinec, P., Foldyna, J., Sitek, L., Ščučka, J., Vašek, J.: Abrasives for AWJ cutting. 2002,Academy of Sciences, Czech Republic. p. 80.

[2] Kovacevic, R. Erosion by single particle impact, http://engr.smu.edu/rcam/ research/waterjet/wj2.html.

[3] Brandt, S., Maros, Z., Monno, M.: AWJ parameters selection – a technical and economical evaluation. Jetting Technology, BHR Group, 2000.

[4] Gombár, M. Tvorba štatistického modelu drsnosti obrobeného povrchu s využitím Matlab. In Výrobné inžinierstvo. (s. 14-17) 2006.

[5] Lupták, M. Study of the measuring and scanning methods in the processes of the cutting and machining with waterjet. In Proc. Metalurgia Junior 05, Deň doktorandov HF Košice, 2005.

[6] Monka, P.: Teoretické vzťahy najvyššej nerovnosti profilu, Výrobné inžinierstvo, 2-3 / 2003, II. Ročník, pp. 20-21, FVT TU v Košiciach, Prešov,ISSN 1335-7972-01.

[7] Monková, K., Hatala, M. Využitie lasera pri povrchovom spracovaní zamerané na zmeny v štruktúre povlakovaného materiálu. In: Acta Mechanica Slovaca. roč. 10, č. 4-a/2006 (2006), s. 178-182. ISSN 1335-2393.

[8] Montgomery, D., C. Design and analysis of experiments. 5th. edition, Hamilton Printing Company, 2001. ISBN 0-471-31649-0.

[9] Jurisevic, B., Coray, P. S., Heineger, K. C., Junkar, M.: Tool Formation Process in Abrasive Water Jet Machining. In: Proceedings of the 6th International Conference on Management of Innovative Technologies MIT'2003 – post-conference edition. Ljubljana: University of Ljubljana. 2003. s. 73–85.

[10] Jurisevic B., Kuzman K., Junkar, M. Water jetting technology: an alternative in incremental sheet metal forming. The International Journal of Advanced Manufacturing Technology. Springer. pp. 18-23. ISSN 1433-3015. 2006.

[11] Ramulu, M., Hashish, M., Kunaporn, S., Posinasetti, P. Abrasive waterjet machining of aerospace materials. http://www.sampe.org/store/paper. aspx?pid=2046.

Poďakovanie

Článok vznikol za podpory projektu VEGA 1/4157/07 „Nelineárne matematické modelovanie a vibrodiagnostika progresívnych technologických procesov pri delení ťažkoobrábateľných materiálov pomocou DoE a Taguchiho dizajnu".

Modelling of radiative heat transfer in building envelopes

Jiří Vala, Stanislav Šťastník and Radek Steuer
Brno University of Technology, Faculty of Civil Engineering
602 00 Brno, Veveří 95, Czech Republic
vala.j@fce.vutbr.cz, stastnik.s@fce.vutbr.cz, steuer.r@fce.vutbr.cz

Introduction

The improvement of thermal technical properties of insulation layers is a crucial point of reconstruction of all building objects, with the aim to extend their lifetime. A typical phenomenon of such reconstruction is the presence of algae and the consequent surface degradation of new insulation systems; this can be explained by the night condensation of vapour from porous material structures, driven by the thermal radiation from building envelopes to the clear sky. The similar problems occur in the design of new buildings, especially in the optimization of the choice of materials for their envelopes. The standard physical, mathematical and numerical models of stationary heat transfer (for the validation of insulation properties), as well as of non-stationary one (for the validation of accumulation properties), are not sufficient because both the short-wave radiative fluxes, representing the effect of sun rays directed to the earth surface, and the long-wave one, emitted from the earth surface to the atmosphere and partially through the atmosphere to the free space, can influence the heat balance substantially.

The qualitative physical analysis of these processes is seemingly clear; however, reliable quantitative calculations are typically not available in standard heat propagation algorithms and corresponding software packages. At least three reasons are evident: i) some analyzed processes have a strong non-deterministic character: e. g.the angle between the sun ray and the earth surface (at every location and time) can be calculated exactly (although the resulting formula is not quite simple), taking into account the motion of the earth reference ellipsoid around the sun, but all radiative processes are affected by the weather, ii) the available data (in confrontation with those from classical thermal conduction and convection) are unprecise, uncertain and insufficient, iii) the mathematical model is usually able to analyse the thermal transfer only in particular layers of a bulding (not in the atmosphere where the radiative transfer equation should be formulated) and is much more complicated than the standard one – the tricky transfomations of nonlinear radiative transfer phenomena to some "equivalent" (quasi-)linear convection process cannot be expected to give good results.

For the first information let us come out from the re-calculated diagrams of [6]. The observations and measurements verify the assumption that the short- and long-wave radiation are strictly separated (by the dependence of the irradiance on the wavelength). The average energy budget, recorded in $\mathrm{W/m^2}$, can be expressed in the following way: i) *short-wave radiation – from space to surface*: $+342$ emitted from space, -67 absorbed by atmosphere (see later), -77 reflected by clouds and gases in atmosphere, -30 reflected from surface directly, totally: 168 absorbed by surface, 107 returned to space, ii) *long-wave radiation – from surface to space*: $+390$ emitted from surface, -350 absorbed by greenhouse gases and clouds (see later), totally: 40 reaching space, iii) *contained in atmosphere*: 67 contributed from short-wave part, 350 contributed from long-wave part, 24 added by direct heating (sensible heat), 78, from the total sum 519: 195 radiated to space, 324 counter-radiated to surface. Then the space balance can be expressed as $-342 + 107 + 40 + 195 = 0$, similarly the surface balance is $168 - 390 - 24 - 78 + 324 = 0$.

The physical processes, sketched in this table, should be respected in the computational model. The base of this model, taking into account the heat propagation by conduction and convection with respect to the moisture redistribution, can be adopted from [8]; this will be presented in the proceeding section. Then we shall come back to the analysis of radiative processes to demonstrate how the corresponding equations can be coupled with the discussed conduction-dominated model; the enqineering practice requires namely to study the phenomena on and near the interface between the building envelope and the atmosphere.

Heat convection and conduction with moisture

We shall assume that all material characteristics can be described at their macroscopic level by some "effective values, in general dependent on some unknowns, but this dependence can be expressed by certain deterministic functions. Moreover, we shall suppose that the macroscopic isotropy is valid. Let t be the positive time, starting from zero ($t = 0$), and let $x = (x_1, x_2, x_3)$ denote the Cartesian coordinate in some domain Ω or in its arbitrary subdomain ω belonging to the 3-dimensional Euclidean space. We shall use the standard notation $\nabla = (\partial/\partial x_1, \partial/\partial x_2, \partial/\partial x_3)^{\mathrm{T}}$ for gradients of functions. Let $\lambda(u)$ be the heat conduction factor and $\zeta(u)$ the product of the heat capacity and of the material density; u here is the dimensionless moisture content, related to the unite material volume. The redistribution of two unknown fields, prescribed for $t = 0$, will be studied in time – of the temperature $T(x, t)$ and of the above introduced moisture content $u(x, t)$. Moreover, in some considerations we shall use the decomposition $u = u_v + u_l$ where u_v will refer to vapour and u_l to liquid water. The balance of heat fluxes on ω can be expressed as

$$\int_\omega [\zeta(u)\partial T/\partial t - \partial L(u, u_l)/\partial t] \, \mathrm{d}x + \int_{\partial\omega} q \cdot \nu \, \mathrm{d}s(x) = 0 \qquad (1)$$

where $\partial\omega$ to the boundary of a subdomain ω, supplied by the surface measure s and by an outward unit normal ν (almost) everywhere, and the classical Fourier law (see [3], p. 191) evaluates the heat flux q as

$$-q = \lambda(u)\nabla T\,. \tag{2}$$

Let us remind that the symbol ∇ applied to a vector field means the divergence, the same applied to a scalar field means the gradient. The only non-standard term in (1), not included in most basic courses on heat propagation, is the second additive one on the left-hand side; $L(u, u_l)$ here characterizes the amount of heat energy lost by the phase change, taking into account the latent heat (no presence of ice is allowed). An arbitrary choice of ω allowed; thus, inserting (2) into (1), using the divergence theorem, we receive

$$\zeta(u)\partial T/\partial t + \nabla \cdot (\lambda(u)\nabla T) = \partial L(u, u_l)/\partial t \tag{3}$$

which is the differential form of our governing equation for heat conduction. For the heat transfer between two material layers, following [3], p. 203, we have

$$-q \cdot \nu = \alpha(T - T^*) \tag{4}$$

where T^* refers to the temperature from the adjacent layer (or, alternatively, of the external environment) and α is the heat convection factor. The same can be done even in case that T^* refers to the external environment.

The equation (3), supplied by all needed boundary conditions (4), is a partial differential equation for the evolution of T. It cannot be solved directly because we do not know the development of the moisture content u in advance. The proceeding approach is based on the moisture balancing. Let u^* be the maximal admissible moisture content for the material, whose porosity is P. Let p be the partial pressure of vapour and p^* the partial pressure of saturated vapour, depending on the temperature T. Evidently $0 \le u \le u^* \le P$, $p \le p^*(T)$. Then the diffusive flux can be evaluated for the vapour and liquid water separately from the formulae

$$q_v = \eta\nabla p\,, \qquad q_l = \xi(u_l)\nabla u_l\,; \tag{5}$$

here η denotes the diffusive ratio for air (the diffusive transfer coefficient, divided by the diffusive resistance) and ξ (a prescribed function of u_l) the factor of capillary diffusion. The conservation of mass implies

$$\int_\omega (\partial u/\partial t)\,\mathrm{d}x + \int_{\partial\omega} (q_l + q_v)\,\mathrm{d}s(x) = 0\,. \tag{6}$$

Following [8], for fixed vectors q_v and q_l from (5), we would be able to calculate u; this forms a good basis for the construction of an iterative numerical procedure. Let us suppose that some monotone dependence between u and p/p^*, nearly independent of T, called sorption curve; the detailed micromechanical analysis can be found in [12], p. 22. Then the univeral gas equation can be adopted to determine u_v from p; this is the last missing information, needed for the evaluation of $u = u_v + u_l$. Applying the method in discretization in time and working (for simplicity) with equidistant time steps h, we are able in any discrete time $t = jh$ with an integer j to evaluate T, taking all values of u and u_l from the preceding time step $(j - 1)h$, in practice from the numerical analysis of the equation (3) with corresponding boundary conditions, making use of the finite difference technique. Consequently

we can evaluate $u = u_v + u_l$, using the above sketched internal iteration loop. If the redistribution of u or u_l is rather significant, we are able to repeat the calculation of T in the same time step with corrected values of u and u_l , etc. Let us notice that for the evaluation of u from the integral equation (6) the discrete mesh from the finite difference analysis of (3) is always prepared. Let us also remark that our simplified approach can be generalized in several directions, leading to the HAM ("heat, air and moisture") modelling – for more references see [10].

Long- and short-wave radiation

In the last decade numerous authors have published a large variety of ideas, methods and algorithms how to bridge the gap between the data uncertainties and the need of reliable (nearly) deterministic simulation of thermal behaviour of buildings, forced by day and year climatic cycles. In the 1-dimensional modelling the problem of radiation can be solved as a nonlinear integral equation, making use of the theory of Green functions. In the 2- or 3-dimensional geometry most authors solve numerically certain equation of type (7) and apply some version of the "discrete ordinance method" (DOM – see [2]), based on the discrete representation of directional dependence of the radiation intensity, needed for numerical integration, frequently combined with other methods for the numerical analysis of differential and integral equations, namely with the finite volume method, with various finite difference schemes or with the discontinuous Galerkin finite element method; the more extensive overview of such approaches is being prepared by the authors of this paper. Parallel to these non-trivial techniques simple recommendations and formulae for technical calculations, as [11], occur. Since the long-wave radiation from a surface is conditioned by the visibility of the free space and, conversely, by the presence of other surfaces and obstacles in the vicinity of an analyzed building object, it can be important (by [7]) to find a relatively simple algorithm for the construction of so-called view factors, taking these phenomenon into account – cf. the differential formula (8). Other approaches start with the criticism of the quasi-deterministic evaluation of the scattering in the atmosphere and develop effective (hardware and software supported) algorithms of Monte Carlo simulations, or implement Gaussian probability functions into the DOM, or apply the advanced correlation analysis. Even the formal mathematical existence result for a model deterministic problem, based on the theory of Sobolev spaces, has been derived in [1].

The basic equation for the long-wave radiation, taking into account an emitting, absorbing and scattering gray medium, can be written in the form of [2], p. 376, as

$$v \cdot \nabla I(v) = -(\kappa + \sigma) I(v) + \kappa I_b(v) + \sigma \int_\Theta I(\widetilde{v}) \Psi(\widetilde{v}, v) \, \mathrm{d}s(\widetilde{v}) \tag{7}$$

where $I(v)$ (a function of x by default) is the radiation intensity along a path with a unit vector v, I_b (a function of x, too) is the black-body radiation intensity, $\Psi(\widetilde{v}, v)$ is the scattering phase function representing the probability that a ray from the direction $\widetilde{v}$ is scattered to the direction v and κ, σ and $\kappa+\sigma$ are absorption, scattering and extinction coefficients of the medium, respectively. Since the year 1879 we

know also the formula for the "emissive power" of blackbody radiation $I_b = ST^4$ with the Stefan-Boltzmann constant $S \approx 5.6697$ W/(m^{2}K^4), at most modified by some refractive index. Such formula can be obtained by the integration of the blackbody radiation intensity over the entire spectrum of wavelengths, using the spectral distribution by Planck; for all details see and [4], p. 323, and [3], p. 116, where also other important characteristics of are introduced – the surface emissivity ε and reflexivity ρ.

In most application in civil engineering, the crucial point of calculations is not to solve (7), but to set the radiative heat fluxes at the surface of an enclosure, i. e.to adopt (4) to handle radiation. Fortunately, we are allowed to join new additive terms to the right-hand side of (4), especially the incoming and outgoing radiative flux q_- and q_+. In general, we have to evaluate the energy interchange between two or more surfaces; cf. an illustrative example in [4], p. 322. To support this the evaluation of q_- and q_+, it is useful to detect all surfaces, identified by integer indices j and k, and for any couple of such indices to introduce the view factor F_{jk} as the fraction of energy leaving surface k that is incident on surface j; the non-trivial calculation in [7], p. 5, results

$$\mathrm{d}F_{jk} = \cos\theta_j \cos\theta_k / (\pi R_{jk}^2)\, \mathrm{d}A_j\, \mathrm{d}A_k \tag{8}$$

where R_{jk} means the representative distance between two differential areas $\mathrm{d}A_j$ and $\mathrm{d}A_k$ and θ_j and θ_k are two angles between the normals to the surfaces in C_j and C_k and the line joining C_j, C_k; practical calculations need some numerical integration technique.

Apart from the long-vawe radiation, the analysis in heat transfer in buildings cannot neglect even the short-wave radiation. Nevertheless, the irradiation of the earth surface depends strongly on the (non-deterministic) climatic conditions; the quantitative relationa by various authors have no unified form and give sometimes quite different results; thus the simplified approach of [11], p. 16 with its step-by-step evaluation of direct, global, diffuse and reflected solar radiation cannot be rejected. A more complicated Brasil-SR model, discussed in [5], calculates

$$q_* = I((\tau - \overline{\tau})(1 - c) + \overline{\tau}) \tag{9}$$

where I is the irradiation at the top of the atmosphere and the radiation flow reaching the surface q_* should be inserted from (9) to the extended right-hand side of (4). In (9) the factor τ, corresponding to a clear sky condition, is a function of i) the scattering on the surface, ii) the solar zenital angle, iii) the optical thickness of all atmospheric constituents, the similar factor $\overline{\tau}$ refers to the complete overcast condition; moreover, the cloud cover coefficient c is needed.

A practical application

The analysis of insulation and accumulation properties of building materials belongs to the research directions of the Department of Building Materials and Components at the Faculty of Civil Engineering of the Brno University in Technology. For the

modelling of heat conduction and convection processes, conditioned by the moisture redistribution, the original software code in Pascal was developped; a numerical simulation for a real wall with and without the influence of radiation is contained in [9].

Another serious (and not closed) problem, discussed in [9], too, is the reliable identification of material characteristics for simplified formulae from preceding sections. The relevant study involves both laboratory experiments and numerical simulations and refers to the analysis of uncertainties in measurements, physical models, numerical algorithms and calculation schemes.

Acknowledgement. This research is supported by the Czech research project CEZ MSM 0021630511.

References

[1] Amosov, A. A.: Global solvability of a nonlinear nonstationary problem with a nonlocal boundary condition of radiative heat transfer type, Differential Equations 41 (2005), pp. 96-109.

[2] Coelho, P. J.: A modified version of discrete ordinates method for radiative heat transfer modelling, Computational Mechanics 33 (2004), pp. 375-388.

[3] Davies, M. G.: Building Heat Transfer, John Wiley & Sons, 2004.

[4] Li, B. Q.: Discontinuous Finite Elements in Fluid Dynamics and Heat Transfer, Part 8: External Radiative Heat Transfer, pp. 319-361, Springer, 2006.

[5] Martins, F. R., Pereira, E. B., Levantamento dos recursos de energia solar no Brasil com o emprego de satélite geostacionário – o Projeto SWERA, Revista Brasileira de Ensino de Físcia 26 (2004), pp. 145-159.

[6] Mills, G.: Radiation Climatology, Encyclopedia of World Climatology (Oliver, J. E., ed.), Part 18, Springer, 2005.

[7] Minnich, A., Turner, J., Hall, M.: A viewfactor-based radiative heat transfer model for Telluride, Telluride Team Report, Los Alamos National Laboratory, 2002.

[8] Šťastník, S., Vala, J., Steuer, R.: Effect of thermal radiation on the surface degradation of materials of building claddings, Abstr. Thermophysics 2006 in Kočovce (Slovak Rep.), p.7, CD-ROM Proc., 10 pp.

[9] Šťastník, S., Vala, J., Steuer, R.: Numerical simulation of heat and moisture propagation in building envelopes during climatic cycles, Forum Statisticum Slovacum 4 (2007), to appear.

[10] Vala, J., Šťastník, S.: On the thermal stability in dwelling structures, Building Research Journal 52 (2004), pp. 31-56.

[11] Verein Deutcher Ingenieure: Umweltsmeteorologie – Wechselwirkungen zwischen Atmosphäre und Oberflächen, Berechnung der kurz- und der landwelligen Strahlung, VDI Richtlinien 3879, Blatt 2, 1994.

[12] Wit, M. H.: Heat and Moisture in Building Envelopes, Technische Universiteit Eindhoven, 2006.

Neural networks versus nonparametric regression*

Štefan Varga, Margaréta Hladíková

Department of Mathematics, Faculty of Chemical and Food Technology

Slovak University of Technology

Radlinského 9

812 34 Bratislava

e-mail: stefan.varga@stuba.sk

Abstrakt

A dependence of an observed variable on one or more predictors can be investigated both by neural networks and by regression analysis. The results of predictions in parametric regression models are usually better than the results of predictions in neural networks. Advantage of the parametric regression is the understanding of the model (type of the regression function). It is reasonable to compare prediction methods of neural networks and prediction methods of nonparametric regression because the inputs in the both methods are the same (unknown models). The comparison of the methods is the aim of the paper.

Key words: neural networks, neurons, predictions, smoothing

1. Neural networks

Neural networks are parallel distributed systems that are competent to place and repeatedly to use experiences acquired in the process of learning. The experiences are put in the parameters of the neural network – weight coefficients w_{ij} and threshold coefficients ϑ_i.

Formally, the neural network is an acyclic oriented graph. The neurons are the vertices of the graph and connections among the neurons are oriented edges of the graph.

A signal in the neural network is distributed in concord of the oriented edges. Input neurons receive the signal from the external surround. On the other hand, output neurons hand over the signal to the surround. The third type of neurons are hidden neurons.

Activities of the hidden and output neurons are determined by the formula

* This paper was supported by the grant VEGA 1 / 2005 / 05, by the project Kniha.sk, by the APVV 0375-06

$$x_i = \sigma\left(\sum_j w_{ij} x_j - \vartheta_i \right),$$

where x_i, x_j are the activities of the *i*-th and *j*-th neuron, w_{ij} is the weight coefficient (evaluating the edge connecting the *j*-th neuron with the *i*-th neuron), ϑ_i is the threshold coefficient of the *i*-th neuron and $\sigma(x)$ is a sigmoid realization of the transfer function

$$\sigma(x) = \frac{b + a\, e^{-\lambda x}}{1 + e^{-\lambda x}}.$$

Here *a*, *b* are constants and λ is a coefficient of the increase. It is evident that

$$\lim_{x \to -\infty} \sigma(x) = a, \quad \lim_{x \to \infty} \sigma(x) = b.$$

Let us consider a set of objects $O = \{o_1, o_2, ...\}$. Each object is determined by the descriptor $x(o)$ and the property $y(o)$. The relationship between the descriptor and the property is usually given by a hypothetical function

$$y = F(x).$$

In most cases of interest, an analytical form of this function *F* is unknown. The aim is to predict the property of the objects (elements of the set *O*). Nonparametric regression (second part of the contribution) deals with the same problem.

Adaptation process (learning) of neural networks is realized on a training set (regression table). The training set is a set of order pairs of k vectors $\boldsymbol{x}^{(k)}$ and $\boldsymbol{y}_{reg}^{(k)}$, where $\boldsymbol{x}^{(k)} = \left(x_1^{(k)}, x_2^{(k)}, ..., x_N^{(k)}\right)$ is a vector of *N* inputs and $\boldsymbol{y}_{reg}^{(k)} = \left(y_{1,reg}^{(k)}, y_{2,reg}^{(k)}, ..., y_{M,reg}^{(k)}\right)$ is a vector of *M* requisite outputs.

The partial objective function $E^{(k)}$ is the sum of squares of differences between the computed $\boldsymbol{y}^{(k)}$ and the requisite output $\boldsymbol{y}_{reg}^{(k)}$ of the network

$$E^{(k)}(\boldsymbol{w}, \vartheta) = \sum_{i=1}^{M} \left(y_i^{(k)} - y_{i,reg}^{(k)} \right)^2.$$

The objective function $E(\boldsymbol{w}, \vartheta)$ for all pairs of the training set is

$$E(\boldsymbol{w}, \vartheta) = \sum_k E^{(k)}(\boldsymbol{w}, \vartheta).$$

The adaptation process of the neural network means to find the weighted and threshold coefficients such that the objective function $E(\boldsymbol{w}, \vartheta)$ be minimal

$$\left(\boldsymbol{w}_{opt}, \vartheta_{opt}\right) = \arg\min_{(w,\vartheta)} \; E(\boldsymbol{w}, \vartheta).$$

We used own software of the gradient method of the biggest declivity. The weighted and threshold coefficients are repeatedly computed using the formulas

$$w_{i+1} = w_i - \alpha \frac{\partial E}{\partial w_i}$$
$$\vartheta_{i+1} = \vartheta_i - \alpha \frac{\partial E}{\partial \vartheta_i} \; .$$

2. Nonparametric regression models

Parametric regression models are usually studied in the form

$$y = f(x_1, x_2, ..., x_p, a_1, a_2, ..., a_m) = f(\boldsymbol{x}, a_1, a_2, ..., a_m),$$

where y is an observed variable, $\boldsymbol{x} = x_1, x_2, ..., x_p$ is a vector of p- predictors, $a_1, a_2, ..., a_m$ are unknown regression parameters and f is known function. First must be estimated the unknown regression parameters ($\hat{a}_i = est\, a_i$) and then we can predict the value (expectation) of the observed variable y in the point $\boldsymbol{x}_i$ ($\boldsymbol{x}_i = (x_{1i}, x_{2i}, ..., x_{pi})$

$$\hat{y}_i = f(\boldsymbol{x}_i, \hat{a}_1, \hat{a}_2, ..., \hat{a}_m).$$

On the other hand, nonparametric regression model is studied in the form

$$y = f(x_1, x_2, ..., x_p) = f(\boldsymbol{x}),$$

where y and $\boldsymbol{x} = (x_1, x_2, ..., x_p)$ are define as in the parametric regression model, but there are not any regression parameters and the function $f(x)$ is completely unknown (unknown type, unknown number of parameters) and we can only assume that it is a smooth continuous function with first and second derivatives. The observation is

$$y_i = f(x_i) + e_i$$

$(i = 1,2,...,n)$. The aim is to fit the model. The fitting curve $\hat{f}(x)$ is called smoother and methods for its finding are called smoothing methods. There are a lot of smoothers. In the following, we introduce some of them.

- **Spline smoothing.** The smoother $\hat{f}(x)$ minimizes the compromise between the fit and the degree of smoothness of the form

$$\sum_{i=1}^{n} \left(y_i - f(x_i)\right)^2 + h \int_{x_{\min}}^{x_{\max}} \left(f''(x)\right)^2 dx$$

over all twice-differentiable functions f. The parameter h, $h>0$ is a smoothing parameter. The bigger h, the more smooth is $\hat{f}(x)$.

- **Kernel smoothing.** The kernel smoother calculates a weighted average of the observations in a neighborhood of the target point x_i

$$\hat{f}(x_i) = \sum_{j=1}^{n} y_i \, K\left(\frac{x_i - x_j}{b}\right) \Big/ \sum_{j=1}^{n} K\left(\frac{x_i - x_j}{b}\right),$$

where b is a bandwidth parameter, which determines how large the neighborhood of the target point is used to calculate the local average. The function K used to calculate the weights is called a kernel function which has the following properties

$$K(t) \geq 0, \quad K(-t) = K(t), \quad \int_{-\infty}^{\infty} K(t)\, dt = 1, \quad \lim_{t\to\infty} t\, K(t) = 0.$$

One of the parametric sets of the kernel functions ($n \in (0,\infty), m \in (0,\infty)$, c is a constant, for $n = 1$ & $m = 1$ it is a triangle) is

$$K(t) = \begin{cases} c\,(1-|t|^m)^n & t \in [-1,1] \\ 0 & elsewhere \end{cases} \qquad c = \frac{\Gamma(1+n+1/m)}{2\,\Gamma(1+1/m)\,\Gamma(1+n)}.$$

If the parameter $m = 1$, we have more simple one parametric sets of kernel functions with the constant $c = (1+n)/2$.

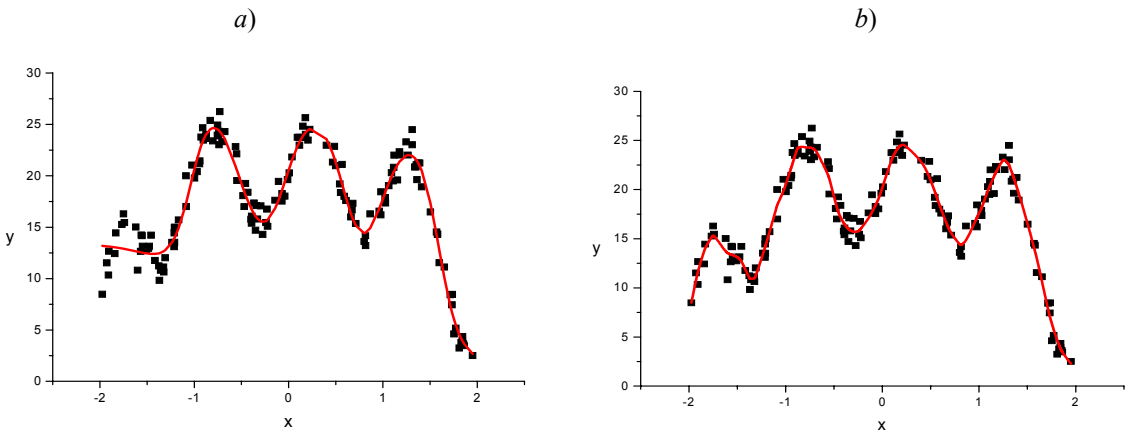
- **Mallows's smoothing.** This is a very simple robust smoother defined by the formula

$$\hat{f}(x_i) = \frac{S(v,i-2)}{4} + \frac{S(v,i-1)}{2} + \frac{S(v,i)}{4},$$

where $S(v,j) = med\,\{y_{j-u},...,y_j,...,y_{j+u}\}$ and $u = (v-1)/2$. In most cases $v = 5$ is used.

3. Comparison

In this part we compare the prediction abilities of neural networks and nonparametric regression. We used two hundred pairs of artificial data generated by the software S-plus. The relationship between the elements of the pairs is rather complicated, a combination of rational and goniometric functions. The understanding of the model $y = F(x)$, we did not use in predictions by neural networks and nonparametric regression.



Picture 1. Fitted values using the neural network (*a*)) and the nonparametric regression (*b*)).

The best result we received using the feed-forward neural network with six hidden neurons (Picture 1, *a*)). Mean quadratic error $S_{nn} = 1.298$. From the nonparametric regression method the best result we obtained using the kernel smoother with the bandwidth parameter $b = 0.4$ (Picture 1, *b*)). Mean quadratic error $S_{nr} = 1.155$.

References

[1] Härdle, W. : Smoothing Techniques with Implementation in S. New York, Springer – Verlag 1981.

[2] Krishnaiah, P.R., Sen, P.K. :Handbook of statistics 4, Nonparametric methods. Amsterdam, North Holland 1984.

[3] Kvasnička, V a kol.: Úvod do teórie neurónových sietí. IRIS, Bratislava, 1997.

[4] Mařík, V., Štepánková, O., Lažanský, J.: Umelá intelligence 1. Academia, Praha, 2001.

[5] Venables, W.N., Ripley, B.D. : Modern Applied Statistics with S-plus. New York, Springer - Verlag 1994.

[6] Varga, Š.: Nonparametric regression and smoothing. In *Proc.* PRASTAN 2004, Kočovce, 2004, 105-108.

[7] Varga, Š., Grančič, P., Katuščák, S., Szitás, A. : Predikcie v neurónových sieťach a logistickej regresii. Forum Statisticum Slovacum 3/2005, 138-145.

Fuzzy regression models with asymmetric fuzzy numbers*

Štefan Varga, Ildikó Horváthová

Faculty of Chemical and Food Technology

Slovak University of Technology

Radlinského 9

812 34 Bratislava

e-mail: stefan.varga@stuba.sk

Abstrakt

We consider asymmetric fuzzy observations in the fuzzy regression model with asymmetric unknown fuzzy parameters. Then we present one classical and one robust estimator of the parameters and show that in the case of the observations and unknown parameters are symmetric fuzzy numbers the proposed estimators are equal to the analogous estimators in the fuzzy regression models with symmetric fuzzy numbers.

Key words: Fuzzy regression model, robust estimations, predictions

1. Fuzzy regression models

We consider the fuzzy regression model in the form

$$Y = A_1 f_1(x) + A_2 f_2(x) + ... + A_m f_m(x)$$

where the input variable x (predictor) is a crisp (real) variable, $f_i(x)$ $(i = 1, 2, ..., m)$ are known real functions of the variable x, Y is an output fuzzy variable (response) and $A = (A_1, A_2, ..., A_m)^T$ is the vector of unknown fuzzy parameters ([7], [8]). It is easy to see that if the fuzzy numbers Y, A_i $(i = 1, 2, ..., m)$ are crisp (real number is a special case of fuzzy number), then the fuzzy regression model is equal to the analogous classical regression model. The uncertainty of an observation Y_i in the point x_i $(i = 1, 2, ..., n)$ is expressed by a membership function μ_{Y_i} of the fuzzy number Y_i. We do not have any probability distribution, any expectation and any variance of the observed value Y_i $(i = 1, 2, ..., n)$.

* This paper was supported by the grant VEGA 1 / 2005 / 05, by the project Kniha.sk, by the APVV 0375-06

The principle question is how to estimate the vector of unknown fuzzy parameters in the fuzzy regression model and how to define a quality of the estimator. One eventuality could be to generalize not only model, but to generalize the estimators defined in the classical regression model to the estimators in the fuzzy regression model. What does it mean? It means that, for example, the least squares (*LS*) estimator of the vector of unknown fuzzy parameters in the fuzzy regression model will be equal to the least squares estimator in the classical regression model in the case that the observations and unknown parameters are crisp numbers as special case of fuzzy numbers with the membership functions

$$\mu_{Y_i}(x) = \begin{cases} 1 & x = Y_i \\ 0 & x \neq Y_i \end{cases}.$$

Because the difference of two fuzzy numbers is a fuzzy number, we will not minimize a sum of squares or other function of differences $Y_i - est\ Y_i$ between observed and estimated fuzzy values, but distances between them that can be defined as crisp numbers.

2. Estimations in fuzzy regression models with symmetric fuzzy numbers

The membership function of a symmetric triangular fuzzy number $A = \langle a, s \rangle$ is

$$\mu_A(x) = \begin{cases} 1 - \dfrac{|x-a|}{s} & a - s \le x \le a + s \\ 0 & otherwise \end{cases}.$$

where a is a center and s is a spread ([4]).
For addition of two symmetric triangular fuzzy numbers $A = \langle a, s_1 \rangle$, $B = \langle b, s_2 \rangle$ and for multiplication with the real number *k*, we can use ([5], [6], [7], [8])

$$A + B = \left\langle a + b,\ \sqrt[w]{{s_1}^w + {s_2}^w} \right\rangle \qquad kA = \left\langle ka,\ \sqrt[w]{|k|}\ s_1 \right\rangle$$

where the parameter $w \in [1, \infty]$. We have the set of arithmetic, but the most interesting are the limit situations. For $w = 1$

$$A + B = \langle a + b,\ s_1 + s_2 \rangle, \qquad kA = \langle ka,\ |k| s_1 \rangle$$

and for $w = \infty$

$$A + B = \langle a + b,\ \max\{s_1, s_2\} \rangle, \qquad kA = \langle ka,\ s_1 \rangle.$$

The distance of two symmetric triangular fuzzy numbers $A = \langle a, s_1 \rangle$, $B = \langle b, s_2 \rangle$, that is a generalization of the Euclidean distance of two real numbers, is the Diamond distance ([1])

$$d^2(A, B) = (a - b)^2 + \frac{2}{3}(s_1 - s_2)^2.$$

Now when we have defined arithmetic and distance of symmetric triangular fuzzy numbers we can specify the studied fuzzy regression model and define the least squares and the robust M estimators of unknown fuzzy parameters.

The studied fuzzy regression model is

$$Y = A_1 f_1(x) + A_2 f_2(x) + ... + A_m f_m(x)$$

where the input variable x (predictor) is a crisp variable, $f_i(x)$ $(i = 1, 2, ..., m)$ are known real functions of the variable x, Y is an output fuzzy variable (response), the observation Y_i $(i = 1, 2, ..., n)$ is a symmetric triangular fuzzy number (y_i is a center and u_i is a spread)

$$Y_i = \langle y_i, u_i \rangle$$

($y_i \in R, u_i \in R^+$) and $A = (A_1, A_2, ..., A_m)^T$ is the vector of unknown symmetric triangular fuzzy parameters **(** $a_i \in R, s_i \in R^+$**)**

$$A_i = \langle a_i, s_i \rangle .$$

To estimate the vector of unknown fuzzy parameters $A = (A_1, A_2, ..., A_m)^T$ means, to estimate the vector of all centers $a = (a_1, a_2, ..., a_m)^T$ and the vector of all spreads $s = (s_1, s_2, ..., s_m)^T$ of the parameters.

- **Least squares (*LS*) estimator** of the unknown fuzzy parameters $A = (A_1, A_2, ..., A_m)^T$ is ([7])

$$est_{LS}\ A = est_{LS}(A_1, A_2, ..., A_m)^T = \arg\min_{a \in R^m, s \in R^{m+}} \sum_{i=1}^{n} d_i^2 =$$

$$= \arg\min_{a \in R^n\ s \in R^{n+}} \sum_{i=1}^{n} \left[(y_i - a^T f_i)^2 + \frac{2}{3}(u_i - s^T |f_i|)^2 \right]$$

where $f_i = (\ f_1(x_i), ..., f_m(x_i)\ )^T$, $|f_i| = (\ |f_1(x_i)|, ..., |f_m(x_i)|\ )^T$, $s^T = (s_1, ..., s_m)$ and d_i is the Diamond distance

$$d_i^2 = d^2(Y_i, est\ Y_i)$$

of the fuzzy observation Y_i and the fitted value $est\ Y_i$ $(i = 1, 2, \ldots, n)$.

- **Robust M estimator** of the unknown fuzzy parameters $A = (A_1, A_2, ..., A_m)^T$ is ([8])

$$est_M\ A = est_M(A_1, A_2, ..., A_m)^T = \arg\min_{a \in R^m, s \in R^{m+}} \sum_{i=1}^{n} \rho(d_i) =$$

$$= \arg\min_{a \in R^n\ s \in R^{n+}} \sum_{i=1}^{n} \rho\left(\sqrt{(y_i - a^T f_i)^2 + \frac{2}{3}(u_i - s^T |f_i|)^2} \right)$$

where the function $\rho(d_i)$ is a function ([2], [3], [7]) that more slowly increase to the infinity that the function d_i^2.

3. Estimations in fuzzy regression models with asymmetric fuzzy numbers

The membership function of an asymmetric triangular fuzzy number $A = \langle a, s, z \rangle$, where a is a center, s a left spread and z a right spread is

$$\mu_A(x) = \begin{cases} \dfrac{s-a+x}{s} & x \in [a-s, a] \\ \dfrac{z+a-x}{z} & x \in [a, a+z] \\ 0 & otherwise \end{cases}.$$

We can generalize addition of two symmetric triangular fuzzy numbers to the case when the fuzzy numbers are asymmetric. For addition (w = 1) of two asymmetric triangular fuzzy numbers $A = \langle a, s_1, z_1 \rangle$, $B = \langle b, s_2, z_2 \rangle$ we have

$$A + B = \langle a+b, s_1 + s_2, z_1 + z_2 \rangle$$

and similarly for multiplication of the asymmetric triangular fuzzy number $A = \langle a, s_1, z_1 \rangle$ with the real number k we have

$$kA = \langle ka, |k|s_1, |k|z_1 \rangle$$

The distance of two asymmetric triangular fuzzy numbers $A = \langle a, s_1, z_1 \rangle$, $B = \langle b, s_2, z_2 \rangle$ we can define by the formula

$$d^2(A,B) = \frac{1}{3}\left[(a-b)^2 + (a-b+z_1-z_2)^2 + (a-b-s_1+s_2)^2\right]$$

because when the fuzzy numbers A, B are symmetric ($s_1 = z_1$, $s_2 = z_2$) we have

$$d^2(A,B) = \frac{1}{3}\left[(a-b)^2 + (a-b+z_1-z_2)^2 + (a-b-s_1+s_2)^2\right] =$$

$$= \frac{1}{3}\left[(a-b)^2 + \big((a-b)-(s_1-s_2)\big)^2 + \big((a-b)+(s_1-s_2)\big)^2\right] = (a-b)^2 + \frac{2}{3}(s_1-s_2)^2$$

what is the Diamond distance of two symmetric triangular fuzzy numbers.

Now when we have defined arithmetic and distance for asymmetric triangular fuzzy numbers we can define the least squares estimator and the robust M estimator of the vector of unknown asymmetric triangular fuzzy parameters in the studied fuzzy regression model.
Recall that the studied fuzzy regression model is

$$Y = A_1 f_1(x) + A_2 f_2(x) + \ldots + A_m f_m(x)$$

where now the input variable x (predictor) is a crisp variable, $f_i(x)$ $(i = 1, 2, ..., m)$ are known real functions of the variable x, Y is an output fuzzy variable (response), the observation Y_i $(i = 1, 2, ..., n)$ is an asymmetric triangular fuzzy number (y_i is a center, u_i is a left spread and v_i is a right spread)

$$Y_i = \langle y_i, u_i, v_i \rangle$$

$(y_i \in R, u_i, v_i \in R^+)$ and $A = (A_1, A_2, ..., A_m)^T$ is the vector of unknown asymmetric triangular fuzzy parameters

$$A_i = \langle a_i, s_i, z_i \rangle$$

$(a_i \in R, s_i, z_i \in R^+)$.

To estimate the vector of unknown fuzzy parameters $A = (A_1, A_2, ..., A_m)^T$ means, to estimate the vector of all centers $a = (a_1, a_2, ..., a_m)^T$, the vector of all left spreads $s = (s_1, s_2, ..., s_m)^T$ and the vector of all right spreads $z = (z_1, z_2, ..., z_m)^T$.

Theorem 1. Least squares (*LS*) estimator of the unknown fuzzy parameters $A = (A_1, A_2, ..., A_m)^T$ in the fuzzy regression model with asymmetric fuzzy numbers is

$$est_{LS}\ A = est_{LS}(A_1, A_2, ..., A_m)^T = \arg\min_{a \in R^m,\, s \in R^{m+}} \sum_{i=1}^{n} d_i^2 =$$

$$= \arg\min_{a \in R^n\ s \in R^{n+}} \sum_{i=1}^{n} \frac{1}{3}\left[(y_i - a^T f_i)^2 + (y_i - a^T f_i - u_i + s^T |f_i|)^2 + (y_i - a^T f_i + v_i - z^T |f_i|)^2\right]$$

where $f_i = (\ f_1(x_i), ..., f_m(x_i)\ )^T$, $|f_i| = (\ |f_1(x_i)|, ..., |f_m(x_i)|\ )^T$, $s^T = (s_1, ..., s_m)$ and $z^T = (z_1, ..., z_m)$.

Proof. First of all let us calculate the fitted value *est* Y_i and then the square d_i^2 of the distance between the observed value $Y_i = \langle y_i, u_i, v_i \rangle$ and the fitted value *est* Y_i. Using the above introduced arithmetic and distance of asymmetric fuzzy numbers, we have

$$\begin{aligned} est\ Y_i &= A_1 f_1(x_i) + ... + A_m f_m(x_i) = \langle a_1, s_1, z_1 \rangle f_1(x_i) + ... + \langle a_m, s_m, z_m \rangle f_m(x_i) = \\ &= \langle a_1 f_1(x_i),\ s_1 |f_1(x_i)|,\ z_1 |f_1(x_i)| \rangle + ... + \langle a_m f_m(x_i),\ s_m |f_m(x_i)|,\ z_m |f_m(x_i)| \rangle = \\ &= \langle a_1 f_1(x_i) + ... + a_m f_m(x_i),\ s_1 |f_1(x_i)| + ... + s_m |f_m(x_i)|,\ z_1 |f_1(x_i)| + ... + z_m |f_m(x_i)| \rangle = \\ &= \langle a^T f_i,\ s^T |f_i|,\ z^T |f_i| \rangle \end{aligned}$$

and while $Y_i = \langle y_i, u_i, v_i \rangle$ the distance

$$d_i^2 = d^2(Y_i, est\ Y_i) = \frac{1}{3}\left[(y_i - a^T f_i)^2 + (y_i - a^T f_i - u_i + s^T |f_i|)^2 + (y_i - a^T f_i + v_i - z^T |f_i|)^2\right]$$

Theorem 2. Robust M estimator of the unknown fuzzy parameters $A = (A_1, A_2, ..., A_m)^T$ in the fuzzy regression model with asymmetric fuzzy numbers is

$$est_M\ A = est_M(A_1, A_2, ..., A_m)^T = \arg\min_{a \in R^m, s \in R^{m+}} \sum_{i=1}^{n} \rho(d_i) =$$

$$= \arg\min_{a \in R^n\ s \in R^{n+}} \sum_{i=1}^{n} \rho\left(\sqrt{\frac{1}{3}\left[(y_i - a^T f_i)^2 + (y_i - a^T f_i - u_i + s^T |f_i|)^2 + (y_i - a^T f_i + v_i - z^T |f_i|)^2 \right]} \right)$$

where the function $\rho(d_i)$ is a function ([2], [3], [7]) that more slowly increase to the infinity that the function d_i^2.

Proof. The proof of this theorem is a simple modification of the proof of theorem 1.

Theorem 3. If $u_i = v_i$ $(i = 1,2,...,n)$ in the fuzzy observation $Y_i = \langle y_i, u_i, v_i \rangle$ and $s_i = z_i$ $(i = 1,2,...,m)$ in the triangular fuzzy parameter $A_i = \langle a_i, s_i, z_i \rangle$ then the estimators from the theorem 1 and the theorem 2 are equal to the analogous estimators in fuzzy regression models with symmetric fuzzy numbers.

Proof. We have already proved that the presented distance of two asymmetric triangular fuzzy numbers is a generalization of the Diamond distance of two symmetric triangular fuzzy numbers and therefore we can use

$$d_i^2 = (y_i - a^T f_i)^2 + \frac{2}{3}(u_i - s^T |f_i|)^2$$

in the estimators in the theorem 1 and the theorem 2.

Conclusion. We generalized the classical estimators and the robust estimators in the fuzzy regression model with symmetric fuzzy numbers ([6], [7], [8]) to the analogous estimators in the fuzzy regression model with asymmetric fuzzy numbers (Theorem 1, Theorem 2).

References

[1] Bárdossy, A., Duckstein, L.: Fuzzy Rule – Based Modeling with Applications to Geophysical, Biological and Eng. Systems, CRC Press, Boca Raton, 1995.

[2] Huber, P.J.: Robust estimation of a location parameter. Ann. Math. Statist. 35, 1964, 73 - 101.

[3] Huber, P.J.: Robust statistics. New York, Wiley 1981.

[4] Klir, G. J., Yuan, B.: Fuzzy Sets and Fuzzy Logic - Theory and Applications, Prentice Hall PTR, Upper Saddle River, NJ, USA, 1995.

[5] Šabo, M.: On T-reverse of T-norms, Tatra Mt. Math. Pub. 12 (1997), 35 – 40.

[6] Varga, Š.: Classical regression model versus fuzzy regression models. Journal of Applied Mathematics, Statistics and Informatics, 1 (2005), No. 2, 95 – 101.

[7] Varga, Š.: Robust estimations in fuzzy linear regression models. Quo Vadis Computational Intelligence. Physica-Verlag, Heidelberg 2000, 239 – 246.

[8] Varga, Š. , Šabo, M.: Linear regression with fuzzy variables. The State of the Art in Computational Intelligence. Physica-Verlag, Heidelberg 2000, 99 – 103.

Včela vyrobí za svoj život dvanástinu lyžičky medu

Nevidí ako sa jej tvrdá práca pretaví na niečo veľké a aký má skutočne význam.
Vy to však vidieť môžete. S osvedčeným softvérom pre vzdelávanie od spoločnosti SAS.

Pridajte sa i vy k vzdelávacím inštitúciám, ktoré využívajú SAS:
- Fakulta managementu Univerzity Komenského
- Ekonomická univerzita v Bratislave
- Slovenská Poľnohospodárska Univerzita
- Prírodovedecká fakulta Univerzity Komenského

www.sas.com/slovakia/uni
education@svk.sas.com

sas | THE POWER TO KNOW®