Blanka Sedlačíková

Andrej Andrejevič Markov a matematická lingvistika (2)

Učitel matematiky, Vol. 19 (2011), No. 2, 112–123

Persistent URL: http://dml.cz/dmlcz/150351

ANDREJ ANDREJEVIČ MARKOV A MATEMATICKÁ LINGVISTIKA (2)

Blanka Sedlačíková

Dokončení z minulého čísla

Vědecké dílo A. A. Markova

Rané práce akademika Markova, jednoho z nejvýraznějších představitelů tzv. petrohradské matematické školy (dále P. L. Čebyšev, A. M. Ljapunov), se týkaly především matematické analýzy (např. limity integrálů, teorie aproximací a konvergence řad, konstruktivní teorie funkcí, diferenciální rovnice) a teorie čísel. Ačkoliv je jeho prací z teorie čísel poměrně málo – 15, měly značný význam. Tyto práce byly věnovány především teorii neurčitých kvadratických forem, důkazům transcendentnosti čísel a a konečně studiu čistě kubických polí.

Nejvýznamnější je však jeho přínos do teorie pravděpodobnosti. Z přibližně 120 původních vědeckých prací se více než jedna třetina týká právě teorie pravděpodobnosti. Zde se stal pokračovatelem myšlenek svého velkého učitele P. L. Čebyševa (například po roce 1900 Markov aplikoval metodu řetězových zlomků, jejímž průkopníkem byl právě Čebyšev, na teorii pravděpodobnosti). Z korespondence A. A. Markova s A. A. Čuprovem (několik desítek dopisů věnovaných obecným otázkám matematické statistiky) se můžeme dočíst, jak se z nejprve negativního vztahu k této matematické disciplíně postupně vyvíjí vztah kladný. Práce z teorie pravděpodobnosti můžeme rozčlenit do tří hlavních skupin, a to: 1. zákon velkých čísel, 2. centrální limitní věty a 3. teorie markovských řetězců.

Markovovy myšlenky o závislých náhodných veličinách vedly ke vzniku zcela nového odvětví teorie pravděpodobnosti, k tzv. *teorii markovských řetězců.* Ta se rozvinula v rozsáhlou a velmi důležitou oblast – *teorii stochastických procesů.* Markovským řetězcem rozumíme posloupnost náhodných proměnných, v nichž je budoucí proměnná určena proměnnou současnou bez ohledu na způsob, jakým přítomný stav ze stavu předchozího vznikl. V roce 1923 se Norbert Wiener stal prvním, kdo přesně a souvisle pojednával o „markovských procesech". Vznik obecné teorie se připisuje Andreji Nikolajeviči Kolmogorovi v průběhu 30. let 20. století. Název *markovské řetězce* byl údajně poprvé použit francouzským vědcem Jacquesem Hadamardem (1865–1963) a toto označení se rozšířilo do odborné literatury. Vůbec poprvé Markov podstatu řetězové závislosti objasňuje ve své práci *Rozšíření zákona velkých čísel na veličiny navzájem závislé* (*Rasprostraněnije zakona bolšich čisel na veličiny, zavisjaščije drug ot druga*) z roku 1907, samozřejmě ještě bez užití termínu *řetězec.* Jednoduše a srozumitelně tu objasňuje princip toho, čemu my dnes říkáme *stejnorodé markovské řetězce* a dále zde dokazuje zákon velkých čísel pro posloupnost náhodných veličin spojených v řetězec za předpokladu existence konečných rozptylů u jednotlivých sčítanců.

Vedle vědecké práce věnoval Markov značnou pozornost vydávání svých přednášek (úvod do analýzy, sférická trigonometrie, diferenciální počet, teorie pravděpodobnosti aj.), v nichž lze vysledovat zvláštnosti osobnosti Markova, jeho chápání předmětu, jeho představy o úrovni znalostí, jež by měli studenti mít.

Co se týče stylu prací Markova je charakteristická jasnost a srozumitelnost jazyka, pečlivé zpracování detailů (dovedení řešení k číslu a algoritmu), ale i dovedení výsledků do možnosti praktického použití.

A. A. Markov a matematická lingvistika

Prvotní nápad provést statistický výzkum výskytu souhlásek a samohlásek v ruských textech vzešel z Markovovy korespondence s A. A. Čuprovem. Šetření prováděl Markov celkem na dvou textech, a to v románu A. S. Puškina *Evžen Oněgin* a v povídce

T. Aksakova *Detskije gody Bagrova-vnuka*, za účelem přiblížit teorii prostých markovských řetězců na takovém příkladu, který bude obecně srozumitelný. Ukažme si podrobně jeho rozbor textu Evžena Oněgina, který byl proveden jako první. Zajímavé je to, že ačkoliv Markov rozvinul svou teorii markovských řetězců jako čistě matematickou práci, aplikoval ji na literární text. S největší pravděpodobností se tedy jedná o první důslednou aplikaci matematiky v jazykovědě.

Zkoumání bylo provedeno na 20 000 ruských hláskách (bez „tvrdého jeru" a „měkkého jeru") veršovaného románu A. S. Puškina *Evžen Oněgin*, které tvoří celou první kapitolu a 16 slok kapitoly druhé. Ukažme si, jak Markov ve své práci postupoval.

Nechť těchto 20 000 hlásek tvoří posloupnost 20 000 za sebou následujících pokusů, z nichž každý může nabývat dvou hodnot, a to hodnoty samohláska nebo souhláska. Připusťme existenci konstantní pravděpodobnosti p, která vyjadřuje pravděpodobnost toho, že daná hláska je samohláskou. Tu odhadneme z výsledků pozorování pomocí počtu samohlásek v textu. Stejně odhadneme i pravděpodobnosti p_1, p_0, $p_{1,0}$, $p_{0,1}$, $p_{0,0}$, $p_{1,1}$, kde:

p_1	je pravděpodobnost, že samohláska následuje za samohláskou
p_0	je pravděpodobnost, že samohláska následuje za souhláskou
$p_{1,0}$	je pravděpodobnost, že samohláska následuje za souhláskou, které předchází samohláska
$p_{0,1}$	je pravděpodobnost, že samohláska následuje za samohláskou, které předchází souhláska
$p_{0,0}$	je pravděpodobnost, že samohláska následuje za dvěma souhláskami
$p_{1,1}$	je pravděpodobnost, že samohláska následuje za dvěma samohláskami

Analogicky odhadneme z textu pravděpodobnosti q, q_1, q_0, $q_{1,0}$, $q_{0,1}$, $q_{0,0}$, $q_{1,1}$, kde:

q je pravděpodobnost, že daná hláska je souhláskou

q_1 je pravděpodobnost, že souhláska následuje za samohláskou

q_0 je pravděpodobnost, že souhláska následuje za souhláskou

$q_{1,0}$ je pravděpodobnost, že souhláska následuje za souhláskou, které předchází samohláska

$q_{0,1}$ je pravděpodobnost, že souhláska následuje za samohláskou, které předchází souhláska

$q_{0,0}$ je pravděpodobnost, že souhláska následuje za dvěma souhláskami

$q_{1,1}$ je pravděpodobnost, že souhláska následuje za dvěma samohláskami

Nyní porovnáme hodnoty získané z textu s předpokládanými teoretickými hodnotami, a to ve dvou případech:

1. výskyt hlásek odpovídá nezávislým pokusům,
2. hlásky jsou uspořádány v řetězec.

add 1)
Předpokládejme, že výskyt samohlásek a souhlásek v textu jsou nezávislé pokusy. To odpovídá situaci, kdy z osudí, v kterém jsou bílé a černé koule (v našem případě samohlásky a souhlásky), opakovaně táhneme s tím, že vytaženou kouli vždy vracíme zpět do osudí. 20 000 hlásek textu rozdělíme na 200 skupin po 100 hláskách (bez porušení pořadí) a spočítáme, kolik je v každé stovce samohlásek. Získáme tak následující tabulku:

Počet samohlásek ve skupině	37	38	39	40	41	42	43
Počet skupin	3	1	6	18	12	31	43

Počet samohlásek ve skupině	44	45	46	47	48	49
Počet skupin	29	25	17	12	2	1

Z tabulky snadno určíme aritmetický průměr, který je roven 43,19, a odtud pak odhadneme pravděpodobnost výskytu samohlásky, kde

$$p \cong 0{,}4319 \cong 0{,}432.$$

Nyní vypočteme součet druhých mocnin odchylek počtu samohlásek každé stovky od aritmetického průměru a dostaneme číslo 1 022,8 odkud po vydělení číslem 200 získáme hodnotu rozptylu (pro skupinu sta hlásek), tj. 5,114.

Nezávislost stanovených veličin (tj. počet samohlásek v stovce hlásek za sebou následujících) potvrzuje i ten fakt, že jestliže je spojíme po dvou, čtyřech a pěti (vytvoříme tedy skupiny po dvou stech, čtyřech stech a pěti stech hláskách po sobě následujících) a vypočítáme pro těchto sto, padesát a čtyřicet posloupností sumu kvadrátů jejich odchýlení od 86,4; 172,8 a 216 (aritmetický průměr počtu samohlásek mezi dvěma sty, čtyřmi sty a pěti sty hláskami), získáme hodnoty 827,6; 975,2; 1 004, které se mnoho neliší od výše získané hodnoty 1 022,8.

Přejdeme-li však od stovek pokusů k jednotlivým pokusům, pozorujeme, že číslo

$$\frac{5{,}114}{100} = 0{,}051\,14$$

(zjištěná hodnota rozptylu) se výrazně liší od hodnoty

$$p(1-p) = 0{,}432 \cdot 0{,}568 = 0{,}245\,376$$

(teoretická hodnota rozptylu za předpokladu, že pokusy jsou nezávislé). Koeficient disperze je v tomto případě roven

$$\frac{5\,114}{24\,537{,}6} \cong 0{,}208,$$

což naznačuje „vázanost“ našich pokusů (tato hodnota se výrazně liší od hodnoty 1, která odpovídá nezávislým pokusům). Vidíme tedy, že sumy vázaných veličin (po sobě následující hlásky textu) mohou tvořit téměř nezávislé veličiny (skupiny po sto hláskách). Později se ukázalo, že tento příklad je značně přínosný pro teorii spojů.

Vytvoříme nová spojení po 100 hláskách. Nejprve si uspořádáme každou původní stovku hlásek do čtvercové matice takto:

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23			...	...	...	...	
			...	...	...	...	...		
91	92	93	94	95	96	97	98	99	100

Z každé pětistovky hlásek vytvoříme pět nových stovek hlásek takto: první stovku sestavíme ze všech prvních a šestých sloupců, druhou z druhých a sedmých sloupců atd. Nyní se vedle sebe nevyskytují žádné dvě sousední hlásky z původních stovek. Dále spočteme, kolik je samohlásek v jednotlivých sloupcích. Tato čísla vždy po dvou sečteme takto: počet samohlásek v prvním a šestém sloupci, v druhém a sedmém sloupci, třetím a osmém sloupci, čtvrtém a devátém sloupci, pátém a desátém sloupci. Pro každou stovku písmen tak získáme pětici čísel, označenou symboly $(1,6)$, $(2,7)$, $(3,8)$, $(4,9)$, $(5,10)$, jejichž součet nám dá počet samohlásek dané stovky. Počet samohlásek v nových stovkách pak odpovídá sumám

$$\sum(1,6), \qquad \sum(2,7), \qquad \sum(3,8), \qquad \sum(4,9), \qquad \sum(5,10),$$

které jsou složeny z odpovídajících pěti sčítanců. (Všechny tyto výsledky uvádí Markov ve 40 přehledných tabulkách – každá tabulka odpovídá pětistovce hlásek. V prvním řádku tabulky je pět čísel $(1,6)$ a jejich suma, v druhém řádku pět čísel $(2,7)$ a jejich součet atd., poslední řádek uvádí počet samohlásek v první stovce, počet samohlásek v druhé stovce atd., až na posledním místě počet samohlásek ve všech pěti stovkách zmenšený o dvě stě.) Sestavíme tabulku četností samohlásek v jednotlivých stovkách podobnou té předešlé:

Počet samohlásek ve skupině	26	27	28	29	30	31	32
Počet skupin	1	0	0	0	1	2	1

33	34	35	36	37	38	39	40	41	42	43	44	45
3	5	1	2	9	13	12	13	11	17	16	15	10

46	47	48	49	50	51	52	53	54	55	56	57
10	16	10	10	5	5	3	3	3	0	1	2

Aritmetický průměr je stejný, a to 43,19. Součet kvadrátů jejich odchylek od 43,2 je značně větší než předešlý, a to 5 788,8. Odtud získáme hodnotu rozptylu

$$\frac{5\,788{,}8}{200} \cong 28{,}944$$

pro těchto 200 posloupností.

Přejdeme-li nyní od stovek hlásek k jednotlivým hláskám, dostaneme hodnotu 0,289 44, která se mnoho neliší od teoretické hodnoty rozptylu pro nezávislé náhodné veličiny

$$0{,}432 \cdot 0{,}568 \cong 0{,}245\,376$$

Koeficient disperze je pak roven

$$\frac{28\,944}{24\,537{,}6} \cong 1{,}18,$$

což se rovněž příliš neodlišuje od teoretické hodnoty koeficientu disperze, která je pro nezávislé náhodné veličiny rovna 1.

Na závěr ještě tyto nové stovky hlásek také pospojujeme po dvou, čtyřech a pěti. Jestliže pro sto, padesát a čtyřicet skupin vypočítáme postupně hodnotu sumy druhých mocnin jejich odchýlení od aritmetických průměrů počtu samohlásek v dané skupině, tj. od 86,4; 172,8; 216, získáme místo čísla 5 788,8 hodnoty 3 551,6; 3 089,2; 1 004, z nichž poslední je téměř šestkrát menší než číslo 5 788,8, což poukazuje na původní vázanost v každé pětistovce hlásek.

add 2)

Uvažujme nyní závislost jednotlivých pokusů. To odpovídá situaci, kdy jsou dána dvě osudí, jedno bílé a druhé černé. V každém z těchto osudí jsou bílé a černé koule (samohlásky a souhlásky). Postupně konáme tahy následujícím způsobem. Vytáhneme-li při n-tém tahu kouli bílé barvy, konáme $(n+1)$-ní tah z bílého osudí, vytáhneme-li při n-tém tahu kouli černé barvy, konáme $(n+1)$-ní tah z černého osudí. Vytažené koule se vždy vloží zpět do osudí. K vysvětlení závislosti nám může posloužit přibližné vyjádření pravděpodobností p_1, p_0. Mezi 20 000 hláskami zkoumaného textu se objeví posloupnost „samohláska – samohláska" celkem 1 104krát. Po vydělení počtem všech samohlásek v textu získáme

$$p_1 = \frac{1\,104}{8\,638} \cong 0{,}128.$$

Podobně najdeme

$$p_0 = \frac{7\,534}{11\,361} \cong \frac{7\,534}{11\,362} \cong 0{,}663.$$

Všimněme si, že pravděpodobnost toho, že hláska je samohláskou, se výrazně liší podle toho, zda jí předchází souhláska ($p_0 = 0{,}663$) nebo samohláska ($p_1 = 0{,}128$). Už tento fakt naznačuje jistou závislost. Označme si dále

$$\delta = p_1 - p_0 = 0{,}128 - 0{,}663 = -0{,}535.$$

Připusťme, že naše posloupnost 20 000 hlásek tvoří prostý řetězec, pak je teoretický koeficient disperze pro $\delta = -0{,}535$ (v souladu s Markovovou prací *Issledovanije zamečatelnogo slučaja zavisimych ispytanij*) roven $\frac{1+\delta}{1-\delta} = \frac{465}{1\,535} \cong 0{,}3$; které se sice neshoduje úplně s dříve získaným číslem 0,208, ale přibližuje se mu podstatně více, než číslo 1, odpovídající případu nezávislých pokusů.

Nyní připusťme, že posloupnost hlásek tvoří „složitý" řetězec a použijeme výsledky práce *Ob odnom slučaje ispytanij svjazanych v složnuju cep*. Sečteme, kolikrát se v posloupnosti hlásek objevuje posloupnost „samohláska – samohláska – samohláska" a „souhláska – souhláska – souhláska". Odtud dostáváme:

$$p_{1,1} = \frac{115}{1\,104} \cong 0{,}104$$

$$q_{0,0} = \frac{505}{3\,827} \cong 0{,}132$$

Abychom mohli aplikovat výsledky uvedené Markovovy práce, předpokládejme, že:

$$\begin{aligned} p &\cong 0{,}432 \\ q &\cong 0{,}568 \\ p_1 &\cong 0{,}128 \\ q_1 &\cong 0{,}872 \\ p_0 &\cong 0{,}663 \\ q_0 &\cong 0{,}337 \\ p_{1,1} &\cong 0{,}104 \\ q_{0,0} &\cong 0{,}132 \end{aligned}$$

a na základě těchto hodnot najdeme

$$\begin{aligned} \delta &= p_1 - p_0 \cong -0{,}535 \\ \varepsilon &= \frac{p_{1,1} - p_1}{q_1} = \frac{-24}{872} \cong -0{,}027 \\ \eta &= \frac{q_{0,0} - q_0}{p_0} = \frac{-205}{663} \cong -0{,}309 \end{aligned}$$

Zjištěné hodnoty dosadíme do vzorce pro výpočet koeficientu disperze

$$\frac{\{q(1-3\varepsilon)(1-\eta) + p(1-3\eta)(1-\varepsilon) - 2(1-\varepsilon)(1-\eta)\}\,(1-\delta) + 2(1-\varepsilon\,\eta)}{(1-\delta)(1-\varepsilon)(1-\eta)} =$$

$$= \frac{1+\delta}{1-\delta}\left\{\frac{1+\varepsilon}{2(1-\varepsilon)} + \frac{1+\eta}{2(1-\eta)} + \frac{1+\eta}{2(1-\eta)}\right\} + \frac{(q-p)(\eta-\varepsilon)}{(1-\epsilon)(1-\eta)}$$

a dostaneme hodnotu 0,195, která je již velmi blízká nalezené hodnotě 0,208.

Nemůžeme tvrdit, že příklad vyhovuje zcela teoretickým podmínkám, na druhou stranu je zřejmé, že tato výrazná shoda teoretické hodnoty koeficientu disperze s hodnotou získanou z výsledků měření není náhodná.

V tomtéž roce Markov rozšířil svá pozorování týkající se střídání souhlásek a samohlásek v ruské literatuře na 100 000 hlásek pověsti S. T. Aksakova *Detskije gody Bagrova-vnuka.* Získal tyto hodnoty: $p_1 = 0{,}552$, $p_0 = 0{,}365$, $\delta = 0{,}187$. Markov byl se svými výsledky spokojen a soudil, že jeho pozorování dostatečně dobře potvrdilo shodu skutečného pozorování hlásek s hypotézou existence prosté řetězové závislosti.

To, že se o problematiku využití matematických metod v lingvistice zajímal i později, svědčí například práce *Ob odnom primeneniju statističeskogo metoda* z roku 1916. Markov zde reaguje na stať N. A. Morozova *Lingvističeskije spektry*, která vyšla v 20. díle *Izvestij russkogo jazyka i slovesnosti.* Práce pojednává o využití „statistické metody" ke studiu jazyka různých spisovatelů. Morozov zkoumal útržky textu po tisíci slovech z děl L. N. Tolstého, A. S. Puškina, N. V. Gogola a I. S. Turgeněva. V těchto útržcích počítal výskyt různých slov a na základě jejich počtu se snažil určit charakteristiky jazyka jednotlivých spisovatelů. Markov tuto myšlenku považuje za velmi zajímavou a konstatuje, že podobné práce mohou mít velký význam. Důležité je uvědomit si, že jako důkaz „stability" zjištěných výsledků a jejich obecné platnosti je odkaz na obecný zákon velkých čísel či shoda s jinými výsledky nedostatečná. Zároveň Markov konstatuje řadu základních chyb. Není tu žádný náznak toho, že výsledky jsou typické pro ruské spisovatele, a nevztahují se pouze na těchto několik málo útržků. Součty, které provádí Markov sám, jsou často v rozporu se závěry Morozova (např. Gogolův jazyk se podle Morozova vyznačuje značnou převahou výskytu předložky „na" oproti předložce „v", Markov ale nachází v první tisícovce slov knihy *Mrtvé duše* předložku „na" 12krát a „v" dokonce 37krát, což podle tabulky odpovídá jazyku Puškina). Markov konstatuje, že zjištěné výskyty slov neopravňují autora k závěrům o individuálním stylu autora, neboť záměna jedné tisícovky slov jinou může tyto závěry zcela převrátit. Jistý stupeň opodstatnění těchto závěrů by byl možný, sčítal-li by autor ne pět tisícovek slov (někdy i méně), ale sto tisíc slov (pokud by se ovšem neukázala druhá poměrně pravděpodobná situace, že výsledky všech autorů se budou blížit jednomu a témuž číslu, podléhajíce obecným zákonům jazyka). A konečně Markov kritizuje využití takových pomocných prostředků, jakými je převedení jedné tabulky na druhou pomocí zvláštních dělitelů ($\frac{1}{26}$,

$\frac{1}{20}$ apod.) či náčrtky, které podstatu díla nezmění.

Již v roce 1913 položil A. A. Markov svou prací *Primer statističeskogo issledovanija nad tekstom „Evgenija Onegina", illjustrirujuščij svjaz ispytanij v cep* a některými pracemi pozdějšími základy jazykovědné disciplíny, kterou dnes nazýváme matematická lingvistika a jejíž počátky klademe do padesátých let 20. století. Nejspíše jako první důsledně aplikoval matematické poznatky v lingvistice (konkrétně teorii později označenou jako teorie markovských řetězců). Za přínosné můžeme považovat i to, že podnítil zájem matematiků o otázky jazykovědné (a rovněž jazykovědců o problematiku matematickou). Jazyk je totiž třeba chápat jako systém, ve kterém se stránka kvalitativní a kvantitativní vzájemně ovlivňují.

Literatura

[1] Černý, J., *Dějiny lingvistiky*, Votobia, Olomouc, 1996.

[2] `http://www-history.mcs.st-andrews.ac.uk/Biographies/Markov.html`; article by J. J. OConnor, E. F. Robertson, srpen 2006.

[3] Grodzenskij, Sergej Jakovlevič, *Andrej Andrejevič Markov (1856–1922)*, Izdatelstvo Nauka, Moskva, 1987.

[4] Hostinský, Bohuslav, *Počet pravděpodobnosti, první část*, Jednota československých matematiků a fyziků, Praha, 1950.

[5] Hostinský, Bohuslav, *Počet pravděpodobnosti, druhá část*, Přírodovědné nakladatelství, Praha, 1950.

[6] Majstrov, Leonid Jefimovič, *Teorija verojatnostej, istoričeskij očerk*, Nauka, Moskva, 1967.

[7] Markov, Andrej Andrejevič, *Izbranyje trudy (Teorija čisel, teorija verojatnostej)*, Izdatelstvo Akademii nauk SSSR, Leningrad, 1951.

[8] Markov, Andrej Andrejevič, Ob odnom primeneniji statističeskogo metoda, In: Izvestija Imperatorskoj Akademii nauk, série 6, díl 10, 1916, č. 4, s. 239–242.

[9] Markov, Andrej Andrejevič, *Primer statističeskogo issledovanija nad tekstom „Evgenija Onegina", illjustrirujuščij svjaz ispytanij v cep,*. In: Izvestija Imperatorskoj Akademii nauk série 6, díl 7, 1913, č. 3, s. 153–162.

[10] Sgall, Petr a kol, *Cesty moderní jazykovědy*, Orbis, Praha, 1964.

Mgr. Blanka Sedlačíková
Ústav matematiky a statistiky PřF MU Brno
Kotlářská 2, 611 37 Brno
e-mail: BSedlacikova@seznam.cz

★ ★ ★

JAK PŘIPRAVIT UČITELE MATEMATIKY

Ve dnech 23. až 25. září 2010 se v Profesním domě na Malostranském náměstí v Praze, v jedné z historických budov Matematicko-fyzikální fakulty Univerzity Karlovy, konala celostátní konference *Jak připravit učitele matematiky.* Organizovala ji Katedra didaktiky matematiky MFF UK.[8]

---

[8] Programový výbor: Jindřich Bečvář, Martina Bečvářová, Leo Boček, Adolf Karger, Oldřich Odvárko, Jarmila Robová (všichni KDM), Dag Hrubý (Gymnázium Jevíčko), František Kopecký (Gymnázium Jana Nerudy, Praha), Aleš Trojánek (Gymnázium Velké Meziříčí).
Organizační výbor: Antonín Slavík, Alena Blažková, Zdeněk Halas, Jana Hromadová, Pavla Pavlíková, Alena Šarounová (všichni KDM), Vlasta Chmelíková, Luboš Moravec, Karel Pazourek, Petra Surynková, Dana Trkovská (doktorandi KDM).