Véletlen permutációk statisztikai vizsgálata

A doktori értekezés tézisei

KÉSZÍTETTE: Csiszár Villő

MATEMATIKA DOKTORI ISKOLA
ALKALMAZOTT MATEMATIKA DOKTORI PROGRAM

ISKOLAVEZETŐ: Laczkovich Miklós
PROGRAMVEZETŐ: Michaletzky György
TÉMAVEZETŐ: Tusnády Gábor,
tudományos tanácsadó, az MTA rendes tagja

Eötvös Loránd Tudományegyetem
Természettudományi Kar
2008

1. Előzmények

Értekezésemben a véletlen permutációk statisztikai elemzésének egyes kérdéseivel foglalkozom. A gyakorlatban a véletlen permutációk legtöbbször mint egy halmaz elemeinek sorbarendezései fordulnak elő. Számos szociológiai kutatásban szerepel olyan feladat, melyben a válaszadóknak adott számú lehetőséget kell rangsorolniuk fontosság vagy preferencia szerint. Máskor szavazók, bírálók vagy felvételi bizottságok jelölteket, pályázókat, jelentkezőket rangsorolnak, majd ezek alapján hoznak döntéseket. Ezen kívül egy permutáció két azonos elemszámú halmaz elemeinek összepárosítását is megadhatja: összepárosíthatók például feladatok és elvégzőik, hallgatók és témavezetők. Végül tetszőleges valós adatokat elemezhetünk csak a rangszámaik alapján.

Célom a permutáció-értékű adatokra illeszthető paraméteres modellek áttekintése, a paraméterbecslés és az illeszkedésvizsgálat tárgyalása, valamint új modellek kifejlesztése volt. Vizsgálataimat részben az a tapasztalat motiválta, hogy az Amerikai Pszichológiai Társaság 1980-as elnökválasztásának adataira – témavezetőmmel együtt – nem találtunk elfogadható illeszkedést mutató egyszerű modellt. Ez az adatsor az egyik legtöbbet vizsgált az irodalomban, lásd például [1, 3, 10, 11, 13]. Vizsgálataink során találtunk egy olyan egyszerű modellt, mely ugyan nem adott kielégítő illeszkedést, de az addig vizsgált modelleknél lényegesen jobb eredményt nyújtott. Ez a modell vezetett oda, hogy a véletlen permutációk feltételes függetlenségekkel jellemezhető, illetve faktorizálódó modelljeit vizsgáljam, melyeket, a kontingenciatáblák elemzésében meglévő szóhasználatot átvéve, az értekezésben hierarchikus modelleknek neveztem el.

A létező modellekről jó áttekintést nyújt Marden [10] könyve, melynek létrejöttéhez nagyban hozzájárult az 1990-ben Amherstben rendezett „Probability Models and Statistical Analyses for Ranking Data” című konferencia, illetve az ebből Fligner és Verducci által szerkesztett [5] kötet. Megemlítem még Critchlow, Fligner és Verducci [2] cikkét, mely ugyan jóval rövidebb, de fontos eredményeket tartalmaz az egyes modellek tulajdonságairól (megfordíthatóság, címke-invariancia, L-felbonthatóság, unimodalitás, és teljes konszenzus).

2. Módszerek

Az egyik elmélet, amit felhasználtam, az algebrai statisztika. Amint az elnevezés is mutatja, ez a tudományterület azzal foglalkozik, hogyan lehet statisztikai modellek vizsgálatához algebrai eszközöket felhasználni. Ezek az eszközök különösen alkalmasak arra, hogy exponenciális családok lezártját vizsgáljuk, mely mind elméleti, mind gyakorlati szempontból fontos. Az értekezésben felhasznált eredmények megtalálhatók Diaconis és Sturmfels [4], Geiger, Meek és Sturmfels [6], valamint Rapallo [12] cikkeiben.

Legyen $\mathcal{X} = \{x_1, \dots, x_s\}$ véges halmaz, $M = (m_{ij})$ pedig egy $t \times s$ méretű, nemnegatív egész elemű mátrix. Azt mondjuk, hogy a $p = (p(x_1), \dots, p(x_s))$ valószínűség-eloszlás az $\mathbf{F}(M)$ úgynevezett *torikus modellhez* tartozik, ha léteznek olyan nemnegatív $\lambda_1, \dots, \lambda_t$ paraméterek, melyekre

$$p(x_i) = c(\lambda) \prod_{j=1}^{t} \lambda_j^{m_{ji}}, \quad 1 \le i \le s.$$

A T halmaz *M-megvalósítható*, ha $\mathrm{Supp}(m_i) \not\subseteq \cup_{j \in T} \mathrm{Supp}(m_j)$ minden $i \notin T$-re, ahol m_i az M mátrix i. oszlopvektora, $\mathrm{Supp}(\cdot)$ pedig az argumentumba írt vektor tartóját jelöli. Az M-hez tartozó nemnegatív *torikus varietás*

$$X_M = \{x \in \mathbb{R}^s_{\ge 0} : \ x^u - x^v = 0 \quad \forall u, v \in \mathbb{N}^s \text{ melyre } Mu = Mv\},$$

ahol $x^u = \prod_i x_i^{u_i}$. A fenti képletben szereplő $x^u - x^v$ polinomok által generált torikus ideált jelölje I_M.

2.1. Tétel. (Geiger et al. [6]) *$cl(\mathbf{F}(M)) = X_M$, ahol $cl(\cdot)$ a lezárást jelöli. Továbbá $p \in X_M$-re $p \in \mathbf{F}(M)$ akkor és csak akkor, ha p tartója M-megvalósítható.*

2.2. Tétel. (Rapallo [12]) *Minden M-hez van olyan $M_{\max}$ maximális reprezentáció, melyre $cl(\mathbf{F}(M)) = \mathbf{F}(M_{\max})$.*

Azt mondjuk, hogy az $f_1, \dots, f_L : \mathcal{X} \to \mathbb{Z}$ függvények az $\mathbf{F}(M)$ modell *Markov bázisát* alkotják, ha minden u-ra az $Mg = u$ egyenletnek eleget tevő $g : \mathcal{X} \to \mathbb{N}$ gyakoriságvektorokon az f_i lépések erősen összefüggő gráfot generálnak (ahol az f_i lépés g-t $g + f_i$-be viszi). A Markov bázisok egyik gyakorlati alkalmazása, hogy kis mintaelemszám esetén az illeszkedést vizsgáló χ^2-próba kiváltható Markov lánc Monte Carlo próbával.

2.3. Tétel. (Diaconis és Sturmfels [4]) *Az $f_1, \ldots, f_L$ függvények akkor és csak akkor alkotnak Markov bázist, ha az $x^{f_i^+} - x^{f_i^-}$ polinomok generálják az I_M ideált, ahol f_i^+ (f_i^-) az f_i pozitív (negatív) része.*

A másik elmélet, amit felhasználtam, a hierarchikus és loglineáris modellek elmélete, illetve általánosabban a diszkrét valószínűség-eloszlások exponenciális családjainak elmélete. A felhasznált eredmények megtalálhatók például Lauritzen [8] kiváló könyvében. Az előbbi felállásban legyenek $A \in \in \mathcal{A}$ az $\mathcal{X}$ partíciói, az $M_\mathcal{A}$ mátrix sorai pedig legyenek az A partíciók osztályainak indikátorvektorai (azaz minden partícióhoz annyi sor tartozik, ahány osztálya van). Ekkor a $\mathrm{cl}(\mathbf{F}(M_\mathcal{A}))$ modellben egyértelműen létezik maximum likelihood becslés, melyet például *iteratív arányos illesztéssel* (IPS) kaphatunk meg. Minden $x \in \mathcal{X}$-re jelölje $x(A)$ az A partíció x-et tartalmazó osztályát. Ha pedig p az $\mathcal{X}$-en adott eloszlás, akkor legyen $p(x(A)) = \sum_{y \in \mathcal{X}: y(A) = x(A)} p(y)$ az $x(A)$ osztály p szerinti valószínűsége. Tegyük fel, hogy az r tapasztalati eloszlású mintához szeretnénk megtalálni azt a $\mathrm{cl}(\mathbf{F}(M_\mathcal{A}))$-beli eloszlást, mely szerint a minta likelihoodja maximális. Legyen $p^{(0)}$ az $\mathbf{F}(M_\mathcal{A})$ modell tetszőleges szigorúan pozitív eleme (pl. az egyenletes eloszlás). Az IPS algoritmus $(t+1)$. lépésben legyen

$$p^{(t+1)}(x) = \frac{r(x(A))}{p^{(t)}(x(A))} p^{(t)}(x), \quad x \in \mathcal{X},$$

ahol A ciklikusan befutja $\mathcal{A}$ elemeit.

3. Eredmények

3.1. McCullagh inverziós modellje

Peter McCullagh [11] vezette be következő modellt. A k elemű $C \subseteq \subseteq [n]$ halmaz egy σ_C permutációja $(k-1)$-*ed rendű inverzió*, ha benne egy elem sincs a nagyság szerint növekvő sorrend szerinti helyén (ahol $[n] = = \{1, \ldots, n\}$). Azt mondjuk, hogy egy $\pi \in S_n$ permutáció tartalmazza a

σ_C inverziót (jelölésben $\sigma_C \subseteq \pi$), ha C elemei π-ben éppen σ_C sorrendben állnak. Ekkor a $\sigma^1_{C_1}, \dots, \sigma^s_{C_s}$ inverziókra épülő modellben

$$\log p_\theta(\pi) = \sum_{i:\sigma^i_{C_i} \subseteq \pi} \theta_i, \quad \theta = (\theta_1, \dots, \theta_s) \in \mathbb{R}^s. \tag{1}$$

A disszertációban bebizonyítom McCullagh következő sejtését.

3.1. Tétel. *Az* (1) *modellben, ha* $\theta \neq \tau$, *akkor* $p_\theta \neq p_\tau$.

A tétel ekvivalens a következő érdekes kombinatorikai átfogalmazással. Definiáljuk az S_n halmazon azt a G_H gráfot, melyben a π permutációból azokba a permutációkba vezet irányított él, melyeket π-ből egy elem helyre rakásával kapunk (a többi elemet, ha kell, odébb csúsztatjuk). Például $n =$ $= 5$-re a (24351) permutációból az (12435), (42351), (23541), (24315) permutációkba vezet él.

3.2. Tétel. *A* G_H *gráfban nincs irányított kör.*

3.2. EM algoritmusok Plackett-Luce-féle modellekre

Van n sportoló, az i-edik képességét a λ_i paraméter fejezi ki. A Plackett-Luce modell szerint ha az $I \subseteq [n]$ sportolók versenyeznek egymással, akkor a π sorrend valószínűsége

$$p(\pi) = \prod_{k=1}^{|I|} \frac{\lambda_{\pi(k)}}{\sum_{j=k}^{|I|} \lambda_{\pi(j)}}, \tag{2}$$

ahol $\pi(1)$ az első helyezett, $\pi(|I|)$ az utolsó helyezett. Ezt a modellt Luce [9] a sorbarendezési posztulátumból és a kiválasztási axiómából vezette le. Könnyen látszik, hogy ha a Z_i $(i = 1, \dots, n)$ valószínűségi változók függetlenek és λ_i paraméterű exponenciális eloszlásúak, akkor (2) jobb oldala éppen a $P(Z_{\pi(1)} < \ldots < Z_{\pi(|I|)})$ valószínűség. Ez alapján a λ_i paraméterek maximum likelihood becslésére a következő EM algoritmus számolható ki. Tegyük fel, hogy m megfigyelésünk van, az r-edikben az I_r halmazba eső játékosok π_r sorrendjét figyeljük meg. Jelölje minden $i \in I_r$-re $\alpha_r(i)$, hogy i hányadik helyen áll a π_r sorrendben. Jelölje továbbá m_i azon megfigyelések számát, melyben az i játékos szerepel. Ekkor egy EM-lépés:

$$\lambda_i^{(t+1)} = m_i \left[\sum_{r: i \in I_r} \sum_{k=1}^{\alpha_r(i)} \frac{1}{\sum_{j=k}^{|I_r|} \lambda_{\pi_r(j)}^{(t)}} \right]^{-1} \qquad 1 \leq i \leq n.$$

Hunter [7] erre a modellre, illetve általánosításaira MM algoritmusokat vezetett le, és feltételeket adott az algoritmusok konvergenciájára. Megmutatom, hogy ugyanezekre a feladatokra EM algoritmusok is természetes módon szóba jönnek, bár ezek – tapasztalataim szerint – lassabban konvergálnak.

3.3. L-felbonthatóság

A $\pi = (\pi(1), \ldots, \pi(n))$ permutációra legyen $\pi\{i..j\} = \{\pi(i), \ldots, \pi(j)\}$ és $\pi(i..j) = (\pi(i), \ldots, \pi(j))$. A Π véletlen permutáció (illetve eloszlása) *L-felbontható*, ha a $\Pi\{1..k\}$, $k = 1, \ldots, n$ halmazok Markov láncot alkotnak. Az „L” Luce-ra utal, mivel ezek pontosan azok az eloszlások, melyek eleget tesznek Luce sorbarendezési posztulátumának. Jelölje (π, ρ) két rész-permutáció egymás után írását, valamint egy $C \subseteq [n]$ halmazra álljon S_C a C elemeinek permutációiból.

3.3. Tétel. *Az L-felbontható eloszlások zárt torikus modellt alkotnak. A hozzá tartozó ideált az $x_{(\pi_1,\rho_1)}x_{(\pi_2,\rho_2)} - x_{(\pi_1,\rho_2)}x_{(\pi_2,\rho_1)}$ polinomok generálják, ahol $\pi_1, \pi_2 \in S_C$ és $\rho_1, \rho_2 \in S_{[n]\setminus C}$ valamely $C \subseteq [n]$-re.*

3.4. Tétel. *Az $n = 4$ és $n = 5$ esetben az L-felbontható modell minimális Markov bázisa egyértelmű, és megegyezik az előző tételben leírt bázissal. Az $n \geq 6$ esetben a leírt bázis nem minimális, és a minimális Markov bázis nem egyértelmű.*

A következő tétel a maximum likelihood (ML) becslés tulajdonságairól szól.

3.5. Tétel. *Az L-felbontható modellben a ML becslés egyértelműen létezik, explicit képlettel megadható, pontos eloszlása kiszámolható, valamint teljesül rá a következő hiper-Markov tulajdonság: minden k-ra a $\{\hat{P}(\Pi(1..k) =$*

$= u)\}_u$ *és* $\{\hat{P}(\Pi(k+1..n) = v)\}_v$ *véletlen eloszlások feltételesen függetlenek a* $\{\hat{P}(\Pi\{1..k\} = C)\}_C$ *véletlen eloszlásra nézve, ahol* $\hat{P}$ *a ML becslést jelöli.*

3.4. Duplán L-felbonthatóság

A következő tulajdonságot az irodalomban eddig nem vizsgálták. Nevezzük a Π véletlen permutációt (illetve eloszlását) *duplán L-felbonthatónak*, ha Π és Π^{-1} is L-felbontható. Ezek tanulmányozásához vezettem be a véletlen permutációk hierarchikus modelljeit. Legyen $\mathcal{D}$ (illetve $\mathcal{R}$) az $[n]$ halmaz egy d (illetve r) osztályú partíciója. A π permutáció *durvítása* a $\mathcal{P} = \mathcal{D} \times \mathcal{R}$ szorzatpartícióra a következő $d \times r$-es mátrix:

$$|\pi(\mathcal{P})| = (t_{ij}), \quad t_{ij} = |\{1 \le s \le n : s \in D_i, \pi(s) \in R_j\}|.$$

3.6. Definíció. *Legyenek* $\mathcal{P}_1, \ldots, \mathcal{P}_s$ *az* $[n] \times [n]$ *halmaz szorzatpartíciói. Az* S_n*-en adott szigorúan pozitív* p *eloszlás akkor tartozik a* $\mathcal{P}_1, \ldots, \mathcal{P}_s$ *generátorok definiálta hierarchikus modellhez, jelölésben* $p \in \mathcal{L}(\mathcal{P}_1, \ldots, \mathcal{P}_s)$*, ha valamilyen* θ_i *függvényekre*

$$\log p(\pi) = \sum_{i=1}^{s} \theta_i(|\pi(\mathcal{P}_i)|) \quad \forall \pi \in S_n.$$

Jelölje $\mathcal{D}' \succeq \mathcal{D}$ azt, hogy a $\mathcal{D}'$ partíció finomabb $\mathcal{D}$-nél.

3.7. Tétel. *Legyen* $\mathcal{L}(\mathcal{D}_i \times \mathcal{R} : i = 1, \ldots, s)$ *és* $\mathcal{L}(\mathcal{D} \times \mathcal{R}_j : j = 1, \ldots, t)$ *két hierarchikus modell, ahol* $\mathcal{D} \succeq \mathcal{D}_i$ *és* $\mathcal{R} \succeq \mathcal{R}_j$ *minden* $1 \le i \le s$, $1 \le j \le t$ *esetén. Ekkor a két modell metszete az* $\mathcal{L}(\mathcal{D}_i \times \mathcal{R}_j : i = 1, \ldots, s, j = 1, \ldots, t)$ *hierarchikus modell.*

Ez a tétel alkalmazható az L-felbontható modell és az annak invertálásával kapott modell metszetére, mely éppen a duplán L-felbontható modell (illetve annak szigorúan pozitív része). Ebből, illetve további számolásokból kaptam a következő eredményt.

3.8. Tétel. *A szigorúan pozitív duplán L-felbontható eloszlások családjának* $\sum_{i=1}^{n-1} i^2$ *szabad paramétere van.*

Az értekezésben megadom a család két paraméterezését, azaz a duplán L-felbontható eloszlások logaritmusai által $\mathbb{R}^{n!}$-ben generált altér két bázisát. Az egyik bázis egymásra merőleges vektorból, a másik csupa indikátorvektorból áll.

Minden hierarchikus modellhez tartozik egy olyan $M_{\mathcal{A}}$ mátrix, melyet a Módszereknél írtam le. Az L-felbontható hierarchikus modellhez tartozó mátrixot jelölje M_L, a duplán L-felbontható hierarchikus modellhez tartozót pedig M_B. Az $\mathbf{F}(M_B)$ torikus modell Markov bázisát általános esetben nem tudtam meghatározni, csak $n = 4$-re.

3.9. Tétel. *Az $\mathbf{F}(M_B)$ modell minimális Markov bázisa $n = 4$-re 10 darab másodfokú polinomot (melyek az $\mathbf{F}(M_L)$-hez illetve annak invertáltjához tartozó Markov bázisból származnak), és 8 darab negyedfokú polinomot tartalmaz.*

Ennek segítségével sikerült az alábbi tételt – most már tetszőleges n-re – igazolni.

3.10. Tétel. *Az $\mathbf{F}(M_B)$ modell nem zárt, de még lezártja sem adja ki az összes duplán L-felbontható eloszlást.*

3.5. S-felbonthatóság

A duplán L-felbontható eloszlások tanulmányozásakor fontos szerepet játszott egy erősebb tulajdonság, az S-felbonthatóság. A Π véletlen permutáció és annak p eloszlása *S-felbontható*, ha léteznek $\Lambda(C) \geq 0$ ($C \subseteq [n]$) paraméterek, melyekkel $p(\pi) = \prod_{k=1}^{n} \Lambda(\pi\{1..k\})$ alakba írható. Π *duplán S-felbontható*, ha Π és Π^{-1} is S-felbontható.

3.11. Tétel. *A p szigorúan pozitív eloszlás akkor és csak akkor S-felbontható, ha L-felbontható és léteznek olyan $\Lambda'(C) > 0$ paraméterek, melyekre*

$$P(\Pi(k+1) = x | \, \Pi\{1..k\} = C) = \frac{\Lambda'(C \cup x)}{\sum_{y \notin C} \Lambda'(C \cup y)}.$$

A szigorúan pozitív S-felbontható eloszlások is hierarchikus modellt alkotnak egy M_S mátrixszal. A hozzá tartozó $\mathbf{F}(M_S)$ torikus modell nem zárt, megadható viszont egy Markov bázisa.

3.12. Tétel. *Legyenek $C_1, D_1, C_2, D_2, \ldots, C_j, D_j \subseteq [n]$ olyan halmazok, melyekre $|C_i| = k$, $|D_i| = k+1$, és $C_i, C_{i+1} \subset D_i$ (ahol $C_{j+1} := C_1$). Legyen még $\pi_i \in S_{C_i}$ és $\rho_i \in S_{[n]\setminus D_i}$. Minden ilyen választásra készítsük el a*

$$\prod_{i=1}^{j} x_{(\pi_i, D_i \setminus C_i, \rho_i)} - \prod_{i=1}^{j} x_{(\pi_i, D_{i-1} \setminus C_i, \rho_{i-1})}$$

polinomot, ahol $D_0 = D_j$ és $\rho_0 = \rho_j$. Ezek a polinomok az $\mathbf{F}(M_L)$ modell korábban leírt Markov bázisával együtt $\mathbf{F}(M_S)$ Markov bázisát adják.

Ezek az eredmények azt mutatják, hogy az S-, illetve duplán S-felbontható eloszláscsaládok „algebrailag bonyolultabbak” az L-, illetve duplán L-felbontható családoknál.

3.6. Címke-invariancia

Tegyük fel, hogy egy halmaz elemeit megcímkéztük az $1, \ldots, n$ számokkal, és ezen elemek egy sorbarendezése π. Ha most a halmaz elemeit átcímkézzük, azaz az eddigi i címkét $\sigma(i)$-re cseréljük, akkor ugyanezt a sorbarendezést a két permutáció $\sigma\pi$ szorzata fejezi ki. Hasonlóan, ha π^{-1} az eredeti címkézés szerinti helyezésvektor, akkor az új címkézés szerinti helyezésvektor $\pi^{-1}\sigma^{-1}$. Ez motiválja azt a kérdést, hogy a véletlen permutációkra adott modellek invariánsak-e a balról-, illetve jobbról szorzásokra.

3.13. Tétel. *Legyen $n \geq 4$. Az L-felbontható eloszlások családja invariáns a balról szorzásokra. A σ-val való jobbról szorzásra pontosan akkor invariáns, ha σ eleme az $(n\, n{-}1 \ldots 21)$ és a $(2134 \ldots n)$ permutációk által generált nyolc elemű csoportnak.*

Adódik a kérdés, hogy az L-felbontható eloszlásoknak melyik az a részhalmaza, amely az összes jobbról szorzásra invariáns. Pontosabban azokat az eloszlásokat keressük, melyek minden jobbról szorzás után is L-felbonthatók.

3.14. Tétel. *Legyen $n \geq 4$. Az S_n-en adott szigorúan pozitív p eloszlás akkor és csak akkor lesz minden jobbról szorzás után L-felbontható, ha kvázi-független, azaz léteznek $c_i(x)$, $1 \leq i, x \leq n$ paraméterek, melyekkel $p(\pi) = \prod_{i=1}^{n} c_i(\pi(i))$.*

4. Következtetések

Az értekezésben leírt vizsgálatokból azt a következtetést vonhatjuk le, hogy bár a véletlen permutációk leírására sok jól kezelhető, a gyakorlatban alkalmazható, és bizonyos mechanizmusokat hűen tükröző modell létezik, van még létjogosultsága új modellek bevezetésének is. Itt elsősorban a hierarchikus modellekre gondolok, melyek egy része jól interpretálható, mint bizonyos feltételes függetlenségi relációkat kielégítő eloszlások családja. Foglalkoztat az a kérdés, hogyan lehetne az „egyszerű struktúrájú" hierarchikus modelleket jellemezni, melyek a klasszikus eset felbontható grafikus modelljeinek felelnének meg. Érdekes lenne több hierarchikus modell Markov bázisát kiszámolni, ezek fokára tételeket bizonyítani. Egy másik nyitott kérdés, hogy hierarchikus modellek metszete hogyan írható le általában.

Köszönetnyilvánítás. Szeretném megköszönni témavezetőmnek, Tusnády Gábornak a közel nyolc évnyi közös munkát, a tőle kapott ismereteket, de még inkább azt a matematikai szemléletmódot és emberi hozzáállást, amit folyamatosan sugároz. Köszönöm családom és kollégáim segítségét, támogatását, és nem utolsó sorban azt a sok nógatást, ami nélkül a disszertáció nem jött volna létre.

Az értekezés témájához kapcsolódó dolgozataim

– Conditional independence relations and log-linear models for random matchings. *Acta Math. Hungar.*, Online First (2008).

– (társszerzők: Rejtő L. és Tusnády G.) Statistical Inference on Random Structures. In: *Horizons of Combinatorics*, Bolyai Society Mathematical Studies **17**, Springer (2008), 37-66.

– Markov bases of conditional independence models for permutations. *Kybernetika*, közlésre elfogadva.

– On L-decomposability of random permutations. *J. Math. Psych.*, átdolgozás alatt.

– An acyclic operation on the symmetric group. Benyújtva.

Hivatkozások

[1] Chung, L. and Marden, J. I.: Extensions of Mallows' ϕ model. In: [5], 108-139.

[2] Critchlow, D. E., Fligner, M. A. and Verducci, J. S.: Probability models on rankings. *J. Math. Psych.* **35** (1991), 294-318.

[3] Diaconis, P.: A generalization of spectral analysis with application to ranked data. *Ann. Statist.* **17** (1989), 949-979.

[4] Diaconis, P. and Sturmfels, B.: Algebraic algorithms for sampling from conditional distributions. *Ann. Statist.* **26** (1998), 363-397.

[5] Fligner, M. A. and Verducci, J. S. (eds.): *Probability Models and Statistical Analyses for Ranking Data.* Springer-Verlag, New York (1993).

[6] Geiger, D., Meek, C. and Sturmfels, B.: On the toric algebra of graphical models. *Ann. Statist.* **34** (2006), 1463-1492.

[7] Hunter, D. R.: MM algorithms for generalized Bradley-Terry models. *Ann. Statist.* **32** (2004), 384-406.

[8] Lauritzen, S.: *Graphical Models.* Clarendon Press, Oxford (1996).

[9] Luce, R. D.: *Individual choice behavior.* Wiley, New York (1959).

[10] Marden, J. I.: *Analyzing and Modelling Rank Data.* Chapman&Hall, London (1995).

[11] McCullagh, P.: Permutations and regression models. In: [5], 196-215.

[12] Rapallo, F.: Toric statistical models: parametric and binomial representations. *Ann. Inst. Statist. Math.*, Online First (2006).

[13] Stern, H.: Probability models on rankings and the electoral process. In: [5], 173-195.