Diszkrét valószínűségeloszlások távolsága

Diplomamunka

Írta: Antal László

Matematikus szak

Témavezető:

Móri Tamás, egyetemi docens
Valószínűségelméleti és Statisztika Tanszék

Eötvös Loránd Tudományegyetem
Természettudományi Kar
2009

Előszó

Ez a dolgozat diszkrét valószínűségeloszlások közötti eltérésekkel foglalkozik. Az ilyen eltérések vizsgálatához többnyire valamilyen normát definiálnak; két diszkrét eloszlás távolsága lehet a különbségük normája. A kérdés főleg az, hogy mennyire lehet kicsi egy ilyen távolság, azaz mennyire vannak közel egymáshoz az eloszlások.

A normát például választhatjuk a variációs távolságnak, vagy másként fogalmazva, (konstans szorzótól eltekintve) a totális variáció normának.

Egy lehetséges megközelítés a következő is: a diszkrét eloszlások generátorfüggvényei közötti különbséget vizsgáljuk a $[0,1]$ intervallumon valamilyen normában, használhatjuk pl. a szuprémum-normát.

A dolgozatban főleg ez a kétféle megközelítés játszik hangsúlyos szerepet.

A következőkben három fő területet fogunk érinteni.

Az első fejezet a szita formuláról tartalmaz néhány gondolatot. Ez a rész kapcsolódik a második fejezetben tárgyalt Poisson approximáció témaköréhez, hiszen a szita formula használható a Poisson eloszlás közelítésére is.

A második fejezet az előbb említett Poisson approximáció. Ennek a témakörnek igen kiterjedt elmélete van, a dolgozat csak egy bevezető jellegű leírást tartalmaz. Ebben a Chen-Stein módszert vizsgáljuk.

A harmadik fejezet egy n-edfokú valós polinom k-adfokú tagja együtthatójának lehetséges legnagyobb értékét vizsgálja a polinom (megfelelő intervallumon vett) szuprémum-normájához viszonyítva. A témánkhoz kapcsolódó megfelelő kérdés az, hogy mekkora lehet két diszkrét eloszlás generátorfüggvényeinek különbségében a k-adfokú tag együtthatójának legnagyobb értéke.

Szeretnék köszönetet mondani témavezetőmnek, Móri Tamásnak, aki a témát a figyelmembe ajánlotta, és akihez bármikor fordulhattam segítségért.

1. fejezet

A szita formuláról

A szita formula segítségével véges sok esemény közül a bekövetkező események számának valószínűségét adhatjuk meg. Legyenek A_i, $i = 1, 2, \ldots, n$ a kérdéses események, I_i pedig az A_i esemény indikátora. A $W = \sum_{i=1}^{n} I_i$ jelölés mellett tehát $P(W = 0)$, illetve $P(W = k)$ a meghatározandó mennyiség. A szita formula azon a meggondoláson alapszik, hogy néha az események valamilyen részhalmazát kiválasztva, ezen események metszetének valószínűsége könnyebben számolható, és ezen metszet-valószínűségek segítségével kifejezhető $P(W = k)$. Legyen tetszőleges $M \subseteq \{1, 2, \ldots, n\}$ esetén

$$S_l = \sum_{|M|=l} P(\bigcap_{i \in M} A_i),$$

illetve $S_0 = 1$. S_l egy másik megadása lehet a következő. Legyen $E\left(\binom{W}{l}\right)$ az $A_1, A_2, \ldots, A_n$ események azon l-eseinek számának várható értéke, ahol mind az l esemény bekövetkezik. Annak a valószínűsége, hogy az $A_{i_1}, A_{i_2}, \ldots, A_{i_l}$ események mindegyike bekövetkezik, $P(\bigcap_{j=1}^{l} A_{i_j})$. Vagyis

$$E\left(\binom{W}{l}\right) = \sum_{1 \le i_1 < i_2 < \ldots < i_l \le n} P(\bigcap_{j=1}^{l} A_{i_j}) = S_l$$

Ezáltal W generátorfüggvényére, g_W-re is adhatunk formulát.

1.0.1. Lemma. $g_W(x) = \sum_{i=0}^{n} S_i (x-1)^i$.

Bizonyítás. $g_W(x) = E(x^W) = E\left((x-1+1)^W\right) = E\left(\sum_{i=0}^{W}\binom{W}{i}(x-1)^i\right) = \sum_{i=0}^{n} S_i(x-1)^i.$ $\square$

Ebből az $x=0$ helyen azt kapjuk, hogy

$$P(W=0) = \sum_{i=0}^{n}(-1)^i S_i.$$

Másképpen fogalmazva

$$P\left(\bigcup_{i=0}^{n} A_i\right) = 1 - P(W=0) = \sum_{i=1}^{n}(-1)^i S_i.$$

Másfelől g_W deriváltjainak segítségével $P(W=k)$ is kifejezhető:
$g_W^{(k)}(0) = k!P(W=k) = \sum_{i=0}^{n-k}(-1)^i(i+k)(i+k-1)\cdot\ldots\cdot(i+1)S_{i+k}$, vagyis

$$P(W=k) = \sum_{i=0}^{n-k}(-1)^i\binom{i+k}{k}S_{i+k}.$$

Becslés szempontjából a

$$P(W=k) \leq \sum_i c_i S_i$$

vagy

$$P(W=k) \geq \sum_i c_i S_i$$

érdekes lehet. Az ilyen egyenlőtlenségek az ún. Bonferroni típusú egyenlőtlenségek.

Vegyük észre, hogy a jobb oldali összegekben, ha $|M| = l$ rögzített, akkor minden $P\left(\bigcap_{i\in M} A_i\right)$ együtthatója azonos. Egy természetes általánosítás az, amikor ezek között lehetnek különbözőek, azaz

$$P(W=k) \leq \sum_{M\subseteq\{1,2,\ldots,n\}} c(M)P\left(\bigcap_{i\in M} A_i\right).$$

Ezek az ún. általánosított Bonferroni egyenlőtlenségek. A $P(W=0)$-ra vonatkozó becslés átvihető a g_W becslésévé is, amint az [6]-ben megtalálható.

1.0.2. Lemma. *Ha*

$$P(W=0) \le \sum_{M \subseteq \{1,2,\dots,n\}} c(M) P\left(\bigcap_{i \in M} A_i\right),$$

akkor

$$g_W(x) \le \sum_{M \subseteq \{1,2,\dots,n\}} c(M) P\left(\bigcap_{i \in M} A_i\right) (1-x)^{|M|},$$

ahol $x \in [0,1]$.

Bizonyítás. Vezessük be a B_i, $i \in \{1,2,\dots,n\}$ eseményeket úgy, hogy azok egymástól és az A_i, $i \in \{1,2,\dots,n\}$ eseményektől is függetlenek legyenek, továbbá mindegyikük $1-x$ valószínűséggel következzen be. Legyen $A_i' = A_i \cap B_i$, $i = 1,2,\dots,n$ és $W' = \sum_{i=1}^n I(A_i')$. A feltételből azt kapjuk, hogy

$$P(W'=0) \le \sum_{M \subseteq \{1,2,\dots,n\}} c(M) P\left(\bigcap_{i \in M} A_i'\right).$$

Itt viszont $P\left(\bigcap_{i \in M} A_i'\right) = P\left(\bigcap_{i \in M} A_i\right)(1-x)^{|M|}$ és $P(W'=0) = E(x^W) = g_W(x)$. Ezzel beláttuk az állítást. □

Ezt a lemmát lehet kissé általánosítani.

1.0.3. Lemma. *Ha*

$$P(W=k) \le \sum_{M \subseteq \{1,2,\dots,n\}} c(M) P\left(\bigcap_{i \in M} A_i\right),$$

akkor

$$g_W^{(k)}(x) \le \frac{1}{k!} \sum_{M \subseteq \{1,2,\dots,n\}} c(M) P\left(\bigcap_{i \in M} A_i\right) (1-x)^{|M|-k},$$

ahol $x \in [0,1)$.

Bizonyítás. Az előző bizonyításhoz képest a különbség annyi, hogy $P(W'=0) = \sum_{i=0}^n P(W=i)x^i = g_W(x)$ helyett a következő formulát kell használnunk:

$$P(W'=k) = (1-x)^k \sum_{i=k}^n \binom{i}{k} P(W=i) x^{i-k} = (1-x)^k k! \cdot g_W^{(k)}(x). \quad \square$$

A következő célunk az ún. Rényi szita belátása. Ehhez szükségünk van az alábbi tételre, amely megtalálható például [7]-ben.

Legyenek $A_1, A_2, \ldots, A_n$ tetszőleges események. Az A_i eseményeknek B polinomja, ha B előáll az A_i eseményekből úgy, hogy a komplementer, unió, metszet halmazműveleteket véges sokszor alkalmazzuk.

1.0.4. Tétel. *Legyenek $B_i = f_i(A_1, A_2, \ldots, A_n)$, $i = 1, 2, \ldots, k$ az A_j-k polinomjai és $b_1, b_2, \ldots, b_k$ valós számok. Ekkor*

$$\sum_{i=1}^{k} b_i P(B_i) \geq 0$$

pontosan akkor teljesül, ha teljesül minden olyan olyan esetben, amikor $A_j \in \{\emptyset, \Omega\}$ minden j-re. $\square$

Tegyük fel, hogy G egy egyszerű gráf a $V(G) = \{1, 2, \ldots, n\}$ csúcshalmazon. Legyen $V_1^r = V_1^r(G)$ azon r elemű csúcshalmazok összessége, amelyekben egyáltalán nincs él, és jelölje $V_2^r = V_2^r(G)$ azon r elemű csúcshalmazokat, amelyek legfeljebb egy élt feszítenek ki. Legyen továbbá
$T_r^1 = \sum_{M \in V_1^r} P\left(\bigcap_{i \in M} A_i\right)$ és $T_r^2 = \sum_{M \in V_2^r} P\left(\bigcap_{i \in M} A_i\right)$. Itt $T_0^i = 1$.

1.0.5. Tétel (Rényi szita).

$$P(W = 0) \leq \sum_{i \geq 0} T_{2i}^2 - \sum_{i \geq 0} T_{2i+1}^1.$$

Bizonyítás. Az előzőek szerint elég abban az esetben bizonyítani, amikor $A_i = \emptyset$ vagy Ω minden i-re. Legyen $H = \{i | 1 \leq i \leq n, A_i = \Omega\}$. Ha $H = \emptyset$, akkor mindkét oldal értéke 1. Emiatt feltehetjük azt, hogy $H \neq \emptyset$, sőt azt is, hogy $H = \{1, 2, \ldots, n\}$, mert egyébként vehetjük a H által meghatározott részgráfot. Ezáltal a bizonyítandó a következő lesz:

$$\sum_{i \geq 0} |V_1^{2i+1}| \leq \sum_{i \geq 0} |V_2^{2i}|.$$

Ezt lehet egyszerre bizonyítani a

$$\sum_{i\geq 0} |V_1^{2i}| \leq \sum_{i\geq 0} |V_2^{2i+1}|$$

egyenlőtlenséggel. Teljes indukcióval bizonyítunk n szerint. Ha $n = 1$, akkor az állítás az $1 \leq 1$ alakot ölti. Legyen $n > 1$. Ha G teljes gráf, akkor $\sum_{i\geq 0} |V_1^{2i+1}| = n$, $\sum_{i\geq 0} |V_2^{2i}| = \binom{n}{2}+1$, $\sum_{i\geq 0} |V_1^{2i}| = 1$, $\sum_{i\geq 0} |V_2^{2i+1}| = n$. Vagyis ekkor igaz az állítás.

Tegyük fel, hogy G nem teljes, legyen például $n \in V(G)$ egy olyan csúcs, aminek foka $d < n$. Feltehető, hogy az n-nel összekötött csúcsok $1, 2, \ldots, d$. Tekintsük az $1, 2, \ldots, n-1$ illetve $d+1, d+2, \ldots, n-1$ csúcsok által feszített részgráfokat, legyenek ezek G' illetve G''. Ha $M \in V_1^r(G)$, akkor vagy $M \in V_1^r(G')$ vagy $M \in V_1^r(G'' \cup n)$. Ezért

$$\sum_{i\geq 0} |V_1^{2i}(G)| = \sum_{i\geq 0} |V_1^{2i}(G')| + \sum_{i\geq 1} |V_1^{2i-1}(G'')|$$

és

$$\sum_{i\geq 0} |V_1^{2i+1}(G)| = \sum_{i\geq 0} |V_1^{2i+1}(G')| + \sum_{i\geq 0} |V_1^{2i}(G'')|.$$

Másfelől, ha $M \in V_2^r(G')$ illetve $M \in V_2^r(G'')$, akkor nyilván $M \in V_2^r(G)$, illetve $M \cup \{n\} \in V_2^{r+1}(G)$. Emiatt most

$$\sum_{i\geq 0} |V_2^{2i}(G)| \geq \sum_{i\geq 0} |V_2^{2i}(G')| + \sum_{i\geq 1} |V_2^{2i-1}(G'')|$$

és

$$\sum_{i\geq 0} |V_2^{2i+1}(G)| \geq \sum_{i\geq 0} |V_2^{2i+1}(G')| + \sum_{i\geq 0} |V_2^{2i}(G'')|.$$

Itt nem írhatunk egyenlőséget, mert előfordulhatnak olyan halmazok is, amelyek például csak az $\{1, n\}$ élet tartalmazzák. Az indukciós feltevés szerint

$$\sum_{i\geq 0} |V_1^{2i+1}(G')| \leq \sum_{i\geq 0} |V_2^{2i}(G')|$$

és

$$\sum_{i\geq 0} |V_1^{2i}(G')| \leq \sum_{i\geq 0} |V_2^{2i+1}(G')|,$$

illetve

$$\sum_{i\geq 0}|V_1^{2i+1}(G'')| \leq \sum_{i\geq 0}|V_2^{2i}(G'')|$$

és

$$\sum_{i\geq 0}|V_1^{2i}(G'')| \leq \sum_{i\geq 0}|V_2^{2i+1}(G'')|.$$

Ebből már következik az állítás. $\square$

1.0.6. Megjegyzés. *Teljesen hasonló módon bizonyítható, hogy*

$$P(W=0) \geq \sum_{i\geq 0} T_{2i}^1 - \sum_{i\geq 0} T_{2i+1}^2.$$

2. fejezet

Egy kevés Poisson approximáció

2.1. A Chen-Stein módszer

Ez a fejezet nagy mértékben alapszik az [1] könyvre. Ebben a részben Poisson-approximációval foglalkozunk. Ehhez szükségünk lesz két, a nemnegatív egész számokon értelmezett valószínűségi mérték variációs távolságára. A nemnegatív egész számokra a $\mathbb{N}$ jelölést használjuk.

2.1.1. Definíció. *A μ és ν valószínűségi mértékek variációs távolsága:*

$$d(\mu, \nu) = \sup \left\{ \sum_{i \geq 0} |\mu(A_i) - \nu(A_i)| : \mathbb{N} = \bigcup_{i \geq 0} A_i \text{ partíció} \right\}.$$

Megmutatjuk, hogy

$$d(\mu, \nu) = 2 \sup\{|\mu(A) - \nu(A)| : A \subseteq \mathbb{N}\}.$$

Ehhez vegyük észre, hogy

$$\sum_{i \geq 0} |\mu(A_i) - \nu(A_i)| = \sum_{i \geq 0} \left| \sum_{j \in A_i} \mu(j) - \nu(j) \right| \leq$$

$$\sum_{i\geq 0}\sum_{j\in A_i}|\mu(j)-\nu(j)| = \sum_{j\in\mathbb{N}}|\mu(j)-\nu(j)|$$

igaz $\mathbb{N}$ tetszőleges partíciójára, ezért

$$d(\mu,\nu) \leq \sum_{j\in\mathbb{N}}|\mu(j)-\nu(j)|.$$

Legyen $A_1 = \{j \in \mathbb{N} : \mu(j) \geq \nu(j)\}$ és $A_2 = \{j \in \mathbb{N} : \mu(j) < \nu(j)\} = \mathbb{N}\backslash A_1$. Könnyen látható, hogy az előbbi egyenlőtlenséget az $\mathbb{N} = A_1 \cup A_2$ partíció egyenlőséggel teljesíti. Tehát

$$d(\mu,\nu) = \sum_{j\in\mathbb{N}}|\mu(j)-\nu(j)|.$$

Az is látszik, hogy

$$|\mu(A_1)-\nu(A_1)| = |\mu(A_2)-\nu(A_2)|.$$

Ez pedig pontosan azt jelenti, hogy

$$d(\mu,\nu) = 2\sup\{|\mu(A)-\nu(A)| : A \subseteq \mathbb{N}\}.$$

A továbbiakban $I_i, i \in \mathbf{I}$ fogja jelölni indikátor-változók egy (véges vagy megszámlálható) halmazát, amelyekre $E(I_i) = p_i$ és $q_i = 1 - p_i$. Legyen $W = \sum_{i\in\mathbf{I}} I_i$ és $\lambda = \sum_{i\in\mathbf{I}} p_i$. Ezen kívül, ha egy X valószínűségi változó eloszlására akarunk hivatkozni, akkor az $\mathcal{L}(X)$ jelölést használjuk.

Az egyik legrégebben vizsgált kérdés a Poisson-approximáció témaköréből a binomiális eloszlás Poisson eloszlástól való eltérése. Ha $|\mathbf{I}| = n$, az indikátorok függetlenek és $p_i = p, \forall i$-re, akkor nyilván W binomiális eloszlású, $W \sim Bin(n,p)$. Megmutatható, hogy

$$P(W=k) = \binom{n}{k}p^k(1-p)^{n-k} = \frac{(np)^k}{k!}e^{-np}(1+O(np^2, k^2n^{-1})).$$

Ebből tetszőleges $A \subset \mathbb{N}$ esetén, megfelelő c konstanssal:

$$|P(W\in A) - Po(np)\{A\}| \leq c\sum_{k\in A}\frac{(np)^k}{k!}e^{-np}(np^2+k^2n^{-1}) \leq c(2np^2+p),$$

ahol $Po(\tau)$ egy τ paraméterű Poisson eloszlást jelöl. Ez tulajdonképpen azt jelenti, hogy

$$d(Bin(n,p), Po(np)) \leq O(\max(np^2, p)).$$

Most hagyjuk el a $p_i = p, \forall i$-re feltételt. Ekkor azt kaphatjuk, hogy

$$P(W=k) = \prod_{i=1}^{n} q_i \sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \prod_{j=1}^{k} \frac{p_{i_j}}{q_{i_j}}.$$

Másfelől:

$$\lambda^k = \left(\sum_{i=1}^{n} p_i \right)^k = \sum_{\{i_1, \ldots i_k\} \in \{1, \ldots n\}^k} \prod_{j=1}^{k} p_{i_j}.$$

Ezért

$$0 \leq \lambda^k - k! \sum_{\{i_1, \ldots i_k\} \in \{1, \ldots n\}^k} \prod_{j=1}^{k} p_{i_j} \leq$$

$$\binom{k}{2} \sum_{i=1}^{n} p_i^2 \sum_{\{i_1, \ldots i_{k-2}\} \in \{1, \ldots n\}^{k-2}} \prod_{j=1}^{k-2} p_{i_j} \leq (\max p_i) \binom{k}{2} \lambda^{k-1}.$$

Ez azt jelenti, hogy

$$\sum_{1 \leq i_1 < i_2 < \ldots < i_k \leq n} \prod_{j=1}^{k} p_{i_j} = \frac{\lambda^k}{k!} (1 + O(k^2 \lambda^{-1} \max p_i)).$$

Ezt $P(W=k)$ alakjával összevetve kapjuk, hogy

$$P(W=k) = Po(\lambda)\{k\} \exp\{O(\lambda \cdot \max p_i, k^2 \lambda^{-1} \max p_i)\}.$$

Tulajdonképpen ez az egyenlőség is becslést ad az $\mathcal{L}(W)$ és $Po(\lambda)$ közötti variációs távolságra.

Másképp is adható becslés a kérdéses távolságra, mint azt Le Cam következő eredményei mutatják:

$$d(\mathcal{L}(W), Po(\lambda)) \leq 9 \max p_i, \qquad d(\mathcal{L}(W), Po(\lambda)) \leq \frac{16}{\lambda} \sum_{i=1}^{n} p_i^2.$$

Ez utóbbi becslésben az $\frac{1}{\lambda}$ szorzótényezőre kell leginkább felfigyelnünk, ha azt elhagyjuk, akkor az állítás könnyen bizonyítható a következő lemma segítségével. A μ és ν valószínűségi mértékek konvolúcióját $\mu * \nu$ fogja jelölni. Legyenek $\mu_1, \mu_2, \nu_1, \nu_2$ valószínűségi mértékek $\mathbb{N}$-en.

2.1.2. Lemma. $d(\mu_1 * \mu_2, \nu_1 * \nu_2) \leq d(\mu_1, \nu_1) + d(\mu_2, \nu_2)$.

Bizonyítás. Legyen $\mu_1 = (\alpha_0, \alpha_1, \ldots), \mu_2 = (\beta_0, \beta_1, \ldots)$ és $\nu_1 = (\gamma_0, \gamma_1 \ldots)$.

$d(\mu_1 * \mu_2, \nu_1 * \nu_2) = \sum_{j=0}^{\infty} |\mu_1 * \mu_2(j) - \nu_1 * \nu_2(j)| =$

$\sum_{j=0}^{\infty} |\mu_1 * \mu_2(j) - \nu_1 * \mu_2(j) + \nu_1 * \mu_2(j) - \nu_1 * \nu_2(j)| \leq$

$\sum_{j=0}^{\infty} |\mu_1 * \mu_2(j) - \nu_1 * \mu_2(j)| + \sum_{j=0}^{\infty} |\nu_1 * \mu_2(j) - \nu_1 * \nu_2(j)|$.

A két tag hasonlósága miatt elég csak az elsőt vizsgálnunk.

$\sum_{j=0}^{\infty} |\mu_1 * \mu_2(j) - \nu_1 * \mu_2(j)| = \sum_{j=0}^{\infty} |(\mu_1 - \nu_1) * \mu_2(j)| =$

$\sum_{j=0}^{\infty} | \sum_{i=0}^{j} (\alpha_i - \beta_i)\gamma_{j-i}| \leq \sum_{j=0}^{\infty} \sum_{i=0}^{j} |(\alpha_i - \beta_i)\gamma_{j-i}| =$

$\sum_{j=0}^{\infty} |\alpha_j - \beta_j| \sum_{j=0}^{\infty} |\gamma_j| = \sum_{j=0}^{\infty} |\mu_1(j) - \nu_1(j)| = d(\mu_1, \nu_1)$.

Ugyanígy látható, hogy $\sum_{j=0}^{\infty} |\nu_1 * \mu_2(j) - \nu_1 * \nu_2(j)| \leq d(\mu_2, \nu_2)$.

Ezeket összeadva éppen a bizonyítandót nyerjük. $\square$

Könnyen kiszámítható, hogy $d(\mathcal{L}(I_i), Po(p_i)) = 2p_i(1 - e^{-p_i}) \approx 2p_i^2$. Tehát a Lemma szerint

$$d(\mathcal{L}\left(\sum_{i=1}^{n} I_i\right), Po(\lambda)) \leq \sum_{i=1}^{n} d(\mathcal{L}\left(I_i\right), Po(p_i)) \leq 2 \sum_{i=1}^{n} p_i^2.$$

Eddig mindig feltételeztük az indikátorok függetlenségét. Sokkal nehezebb becslést adni akkor, amikor ez nem áll fenn. E két lehetséges esetet egyszerre kezelő, nagyon hatékony eljárás az ún. **Chen-Stein módszer**. A módszer alapját a következő lemma képezi. Egy $g : \mathbb{N} \to \mathbb{R}$ függvény esetén legyen $\|g\| = \sup\{|g(i)| : i \in \mathbb{N}\}$ és $\Delta g(j) = g(j+1) - g(j)$.

2.1.3. Lemma. $Z \sim Po(\lambda) \Leftrightarrow E\left(\lambda g(Z+1) - Zg(Z)\right) = 0$ *minden* g*-re, amire* $\|g\| < \infty$.

Bizonyítás. Legyen $Z \sim Po(\lambda)$. Ekkor

$$E\left(\lambda g(Z+1) - Zg(Z)\right) = \sum_{i=0}^{\infty}\left(\lambda g(i+1)\frac{\lambda^i}{i!}e^{-\lambda} - i\cdot g(i)\frac{\lambda^i}{i!}e^{-\lambda}\right) =$$

$$\lambda e^{-\lambda}\left(\sum_{i=0}^{\infty} g(i+1)\frac{\lambda^i}{i!} - \sum_{i=1}^{\infty} g(i)\frac{\lambda^{i-1}}{(i-1)!}\right) = 0.$$

A megfordításhoz vegyük a $g(n) = I(n = i)$ függvényt. Ezzel
$E\left(\lambda g(Z+1) - Zg(Z)\right) = \lambda P(Z = i-1) - iP(Z = i) = 0$, vagyis
$P(Z = i) = \frac{\lambda}{i}P(Z = i-1)$, azaz $P(Z = i) = \frac{\lambda^i}{i!}P(Z = 0)$. Ezáltal
$1 = \sum_{i=0}^{\infty} P(Z = i) = \sum_{i=0}^{\infty} \frac{\lambda^i}{i!}P(Z = 0)$. $\Rightarrow P(Z = 0) = e^{-\lambda}$. $\square$

Ezek után a fő gondolat az, hogy amennyiben egy X változóra
$E\left(\lambda g(X+1) - Xg(X)\right)$,,közel" van 0-hoz minden megfelelő g-re, akkor X eloszlása ,,közel" van a Poisson eloszláshoz.

Legyen g olyan függvény, amire $g(0) = 0$, és $A \subseteq \mathbb{N}$-re g teljesíti a

$$\lambda g(j+1) - jg(j) = I(j \in A) - Po(\lambda)\{A\} \tag{1}$$

egyenletet minden $j \geq 0$-ra.
Ezzel például

$$P(W \in A) - Po(\lambda)\{A\} = E(\lambda g(W+1) - Wg(W)). \tag{2}$$

Ha az I_i indikátorok függetlenek, akkor

$$W_i = \sum_{j \neq i} I_j$$

független I_i-től, ezért

$$E(Wg(W)) = \sum_{i=1}^{n} E(I_i g(W)) = \sum_{i=1}^{n} E(I_i g(W_i+1)) = \sum_{i=1}^{n} p_i E(g(W_i+1))$$

Tehát

$$E(\lambda g(W+1) - Wg(W)) = \sum_{i=1}^{n} p_i E(g(W+1) - g(W_i+1)).$$

W és W_i csak akkor nem egyenlőek, ha $I_i = 1$. Az előzőeket összerakva tehát a kövtkező becslést kapjuk:

$$|P(W \in A) - Po(\lambda)\{A\}| \leq \sup_{j \in \mathbb{N}} |g(j+1) - g(j)| \sum_{i=1}^{n} p_i^2 = \|\Delta g(j)\| \sum_{i=1}^{n} p_i^2$$

$$\leq 2\|g(j)\| \sum_{i=1}^{n} p_i^2.$$

A következő célunk lehet $\|g\|$-re illetve $\|\Delta g\|$-re becslést adni. Ezt g speciális, (1) alakja teszi lehetővé. Legyen $U_i = \{0, 1, \dots, i\}$. Ekkor g-re teljesülni fog, hogy

$$g(j+1) = \lambda^{-j-1} j! e^{\lambda} (Po(\lambda)\{A \cap U_j\} - Po(\lambda)\{A\} Po(\lambda)\{U_j\}) =$$

$$\lambda^{-j-1} j! e^{\lambda} (Po(\lambda)\{A \cap U_j\} Po(\lambda)\{U_j^c\} - Po(\lambda)\{A \cap U_j^c\} Po(\lambda)\{U_j\}). \quad (3)$$

Ezért

$$|g(j+1)| \leq \lambda^{-j-1} j! e^{\lambda} Po(\lambda)\{U_j\} Po(\lambda)\{U_j^c\}. \quad (4)$$

2.1.4. Állítás. *$\|g\| \leq \min(\frac{5}{4}, 2\lambda^{-1/2})$ és $\|\Delta g\| \leq \lambda^{-1}(1 - e^{-\lambda})$.*

Bizonyítás. Legyen először $j < \lambda$. (4)-ből:

$$|g(j+1)| \leq \lambda^{-j-1} j! e^{\lambda} Po(\lambda)\{U_j\} \leq$$

$$\lambda^{-1} \sum_{i=0}^{j} \lambda^{-i} \frac{j!}{(j-i)!} \leq \lambda^{-1} \sum_{i=0}^{j} \left(\frac{j}{\lambda}\right)^i \leq (\lambda - j)^{-1}.$$

Másfelől (4)-ből következik az is, hogy

$$|g(j+1)| \leq \lambda^{-j-1} j! e^{\lambda} Po(\lambda)\{U_j^c\} \leq$$

$$\frac{j+2}{(j+1)(j+2-\lambda)} = \frac{1}{j+2-\lambda} + \frac{1}{(j+1)(j+2-\lambda)},$$

ha $j > \lambda - 2$.

Ezek miatt $|g(j+1)| \leq \frac{5}{4}$ tetszőleges λ-ra és $j \geq 1$-re, hiszen

$$\text{ha } j \le \lambda - \frac{4}{5} < \lambda, \text{ akkor } |g(j+1)| \le \frac{1}{\lambda - j} \le \frac{5}{4},$$

$$\text{ha pedig } j > \lambda - \frac{4}{5} > \lambda - 2, \text{ akkor } |g(j+1)| \le \frac{1}{j+2-\lambda} +$$
$$\frac{1}{(j+1)(j+2-\lambda)} \le \frac{1}{\frac{6}{5}} + \frac{1}{\frac{12}{5}} = \frac{5}{4}.$$

$|g(1)|$-re pedig (3)-ból: $|g(1)| \le \frac{1}{\lambda}(1 - e^{-\lambda}) \le 1$. Vagyis $\|g\| \le \frac{5}{4}$.

A $\|g\| \le 2\lambda^{-1/2}$ egyenlőtlenséghez is a fenti két, $|g(j+1)|$-re vonatkozó egyenlőlenséget fogjuk használni. Ha $|j - \lambda| \le \lambda^{1/2}$, akkor az előző becslések megfelelőek ehhez. Ha $|j - \lambda| > \lambda^{1/2}$, akkor a becsléseket meghatározó összegekben az első néhány tagot egyszerűen 1-gyel becsüljük, csak azokra a tagokra alkalmazzuk a geometriai sor összegképletét, amelyek már kisebbek, mint $1 - \lambda^{-1/2}$. Például: ha $\lambda - \lambda^{1/2} \le j < \lambda$, akkor

$$\lambda^{-1} \sum_{i=0}^{j} \lambda^{-i} \frac{j!}{(j-i)!} \le \lambda^{-1} \left(1 + \lambda^{1/2} + \frac{1 - \lambda^{-1/2}}{1 - (1 - \lambda^{-1/2})} \right) = 2\lambda^{-1/2}.$$

Nézzük most $\|\Delta g\|$ becslését. g (3) alakjából $A = \{j\}$ mellett az következik, hogy $g(i+1)$ negatív és monoton csökkenő, amikor $i < j$, és pozitív és csökkenő, amikor $i \ge j$. Ezért $g(i+1) - g(i)$ csak $i = j$-re lesz pozitív. Ekkor

$$g(j+1) - g(j) = e^{-\lambda} \lambda^{-1} \left(\sum_{i=j+1}^{\infty} \left(\frac{\lambda^i}{i!} \right) + \sum_{i=1}^{j} \left(\frac{\lambda^i}{i!} \right) \frac{i}{j} \right) \le \min\{\frac{1}{j}, \frac{1}{\lambda}(1 - e^{-\lambda})\}.$$

$j = 0$-ra ez a különbség negatív lesz. Tetszőleges A esetén is igaz lesz ez az egyenlőtlenség. Ezek alapján valóban, $\|\Delta g\| \le \frac{1}{\lambda}(1 - e^{-\lambda})$. $\square$

Ezzel a becsléssel tehát azt kaptuk, hogy

$$d(\mathcal{L}(W), Po(\lambda)) \le \frac{2}{\lambda}(1 - e^{-\lambda}) \sum_{i=1}^{n} p_i^2.$$

Ezen eljárás ugyan még mindig csak független indikátorokra vonatkozik, azonban csekély módosítással használható összefüggő esetben is. A függetlenséget csak az

$$E(I_i g(W)) = E(I_i g(W_i + 1)) = p_i E(g(W_i + 1))$$

egyenlőségben használtuk, ezt érdemes valahogyan módosítani összefüggőség esetén.

Egy lehetséges módosítás a következő:

$$E(I_i g(W)) = p_i E(g(W)|I_i = 1).$$

Ebből (2) a következőbe megy át:

$$P(W \in A) - Po(\lambda)\{A\} = E(\lambda g(W+1) - W g(W)) =$$

$$\sum_{i=1}^{n} p_i \left(E(g(W+1)) - E(g(W)|I_i = 1) \right).$$

Vegyük $\forall i$-re az U_i és V_i változókat, amelyekre az igaz, hogy U_i eloszlása illetve $V_i + 1$ eloszlása megegyezik W eloszlásával illetve W-nek az $I_i = 1$-re vonatkozó feltételes eloszlásával. Így

$$|P(W \in A) - Po(\lambda)\{A\}| = \left| \sum_{i=1}^{n} p_i \left(E(g(U_i + 1)) - E(g(V_i + 1)) \right) \right| \leq$$

$$\|\Delta g\| \sum_{i=1}^{n} p_i E|U_i - V_i| \leq \frac{1}{\lambda} \sum_{i=1}^{n} p_i E|U_i - V_i|.$$

Ez azt jelenti, hogy

$$d(\mathcal{L}(W), Po(\lambda)) \leq 2 \sum_{i=1}^{n} p_i E \left(\min \left\{ \frac{1}{\lambda}(1 - e^{-\lambda})|U_i - V_i|, 2 \min \left\{ \frac{5}{4}, \frac{2}{\lambda} \right\} \right\} \right).$$

Tehát ilyen esetben akkor kaphatunk a variációs távolságra ,,kis” értéket, ha $E|U_i - V_i|$,,kicsi”.

Ha U_i és V_i megválasztható úgy, hogy $U_i \geq V_i \quad m.m.$, akkor

$$\sum_{i=1}^{n} p_i E|U_i - V_i| = \sum_{i=1}^{n} p_i E((U_i+1)-(V_i+1)) = \lambda E(W+1) - \sum_{i=1}^{n} E(I_i W) =$$

$$\lambda + E(W)^2 - E\left(\sum_{i=1}^{n} I_i W\right) = E(W) - D^2(W).$$

Tehát ilyenkor elég csak a várható értéket és a szórásnégyzetet ismerni.

Tegyük fel, hogy tetszőleges j esetén léteznek olyan $X_{i,j}$, $i \in \mathbf{I}$ valószínűségi változók ugyanazon a valószínűségi mezőn, amelyeknek együttes eloszlása megegyezik az $\mathbf{I}$-be tartozó indikátorok $I_j = 1$-re vonatkozó feltételes eloszlásával:

$$\mathcal{L}(X_{i,j},\ i \in \mathbf{I}) = \mathcal{L}(I_i,\ i \in \mathbf{I} | I_j = 1).$$

Legyen továbbá $X^j = \sum_{i \neq j} X_{i,j}$. Ezzel a korábbiak miatt

$$d(\mathcal{L}(W), Po(\lambda) \leq 2\frac{1-e^{-\lambda}}{\lambda} \sum_{j \in \mathbf{I}} p_j E\left(|W - X^j|\right) =$$

$$2\frac{1-e^{-\lambda}}{\lambda} \sum_{j \in \mathbf{I}} p_j E\left(|I_j + \sum_{i \neq j}(I_i - X_{i,j})|\right).$$

Legyen $\mathbf{I}_j = \mathbf{I} \backslash \{j\}$. Vegyük a következő indexhalmazokat:
$\mathbf{I}_j^+ = \{i \in \mathbf{I}_j | X_{i,j} \geq I_i\}$, $\mathbf{I}_j^- = \{i \in \mathbf{I}_i | X_{i,j} \leq I_i\}$ és $\mathbf{I}_j^0 = \mathbf{I}_i \backslash (\mathbf{I}_i^+ \cup \mathbf{I}_i^-)$.

Ha $\mathbf{I}_j = \mathbf{I}_j^+$ vagy $\mathbf{I}_j = \mathbf{I}_j^-$, akkor azt mondjuk, hogy az X_i-k az I_i indikátorokkal monoton módon párosítottak.

2.1.5. Definíció. *Ha minden j-ra meg lehet úgy adni az $X_{i,j}$, $i \in \boldsymbol{I}$ változókat, hogy $\boldsymbol{I} = \boldsymbol{I}_j^+$, (illetve $\boldsymbol{I} = \boldsymbol{I}_j^-$) akkor azt mondjuk, hogy az I_j, $j \in \boldsymbol{I}$ indikátorok pozitív (illetve negatív) kapcsolatban állnak.*

Ezekkel a jelölésekkel $p_j E(X_{i,j}) = P(I_j = 1) E(I_i | I_j = 1) = E(I_i I_j)$. Ezáltal I_i és I_j kovarianciája:

$$Cov(I_i, I_j) = E(I_i I_j) - p_i p_j = p_j E(X_{i,j} - I_i).$$

Vagyis $Cov(I_i, I_j) \leq 0$, ha $i \in \mathbf{I}_i^-$ és $Cov(I_i, I_j) \geq 0$, ha $i \in \mathbf{I}_i^+$.

Emiatt

$$p_j E\left(|W - X^j|\right) = p_j E\left(|I_j + \sum_{i \neq j}(I_i - X_{i,j})|\right) \leq$$

$$p_j E(I_j) + p_j \sum_{i \in \mathbf{I}_j^-}(I_i - X_{i,j}) + p_j \sum_{i \in \mathbf{I}_j^+}(X_{i,j} - I_i) + p_j \sum_{i \in \mathbf{I}_j^0}(I_i + X_{i,j}) =$$

$$p_j^2 - \sum_{i \in \mathbf{I}_j^-} Cov(I_i, I_j) + \sum_{i \in \mathbf{I}_j^+} Cov(I_i, I_j) + \sum_{i \in \mathbf{I}_j^0}(p_i p_j + E(I_i I_j)).$$

Tehát

$$d(\mathcal{L}(W), Po(\lambda)) \leq 2\frac{1 - e^{-\lambda}}{\lambda}\Big(\sum_{j \in \mathbf{I}} p_j^2 + \sum_{j \in \mathbf{I}} \sum_{i \in \mathbf{I}_j^-} |Cov(I_i, I_j)| +$$

$$\sum_{j \in \mathbf{I}} \sum_{i \in \mathbf{I}_j^+} Cov(I_i, I_j) + \sum_{j \in \mathbf{I}} \sum_{i \in \mathbf{I}_j^0}(p_i p_j + E(I_i I_j))\Big).$$

Ezen formula bizonyos speciális esetekben egyszerűsödik.

2.1.6. Következmény. *Ha* $\boldsymbol{I}_j^+ = \emptyset$, *akkor*

$$d(\mathcal{L}(W), Po(\lambda)) \leq 2\frac{1 - e^{-\lambda}}{\lambda}\left(\lambda - D^2(W) + 2\sum_{j \in \boldsymbol{I}} \sum_{i \in \boldsymbol{I}_j^0} E(I_i I_j)\right).$$

2.1.7. Következmény. *Ha* $\boldsymbol{I} = \boldsymbol{I}_j^-$, *akkor*

$$d(\mathcal{L}(W), Po(\lambda)) \leq 2(1 - e^{-\lambda})\left(1 - \frac{D^2(W)}{\lambda}\right).$$

2.1.8. Következmény. *Ha* $\boldsymbol{I}_j^- = \emptyset$, *akkor*

$$d(\mathcal{L}(W), Po(\lambda)) \leq 2\frac{1 - e^{-\lambda}}{\lambda}\left(D^2(W) - \lambda + 2\sum_{j \in \boldsymbol{I}} p_j^2\right).$$

2.1.9. Következmény. *Ha* $\boldsymbol{I} = \boldsymbol{I}_j^+$, *akkor*

$$d(\mathcal{L}(W), Po(\lambda)) \leq 2\frac{1-e^{-\lambda}}{\lambda}\left(\sum_{j\in \boldsymbol{I}} p_j^2 + \sum_{j\in \boldsymbol{I}}\sum_{i\in \boldsymbol{I}_j^0}(E(I_iI_j + p_ip_j))\right).$$

A $E(I_ig(W)) = E(I_ig(W_i+1)) = p_iE(g(W_i+1))$ egyenlőséget máshogyan is átírhatjuk. Legyen $\mathbf{I} = \mathbf{J}_i^1 \cup \mathbf{J}_i^2 \cup \{i\}$ az $\mathbf{I}$ partíciója. $\mathbf{J}_i^1$-be, illetve $\mathbf{J}_i^2$-be azok az i-től különböző j-k tartozzanak, amelyekre I_j ,,erősen" illetve ,,gyengén" függ I_i-től. Legyen $Z_i = \sum_{j\in\mathbf{J}_i^1} I_j$ és $Y_i = \sum_{j\in\mathbf{J}_i^2} I_j = W - Z_i - I_i$. Ezzel

$$E(I_ig(W)) = E(I_ig(Y_i+1)) + E(I_i(g(Y_i+Z_i+1) - g(Y_i+1))).$$

A második összeadandót becsülhetjük:

$$|E(I_i(g(Y_i+Z_i+1) - g(Y_i+1)))| \leq E(I_iZ_i)\|\Delta g\|.$$

A másik összeadandóra pl. akkor tudunk használható becslést adni, amikor $|E(I_ig(Y_i+1)) - p_iE(g(Y_i+1))|$,,kicsi", hiszen ekkor

$$|p_iE(g(Y_i+1)) - p_iE(g(W+1))| \leq p_i(EI_i + E(Z_i))\|\Delta g\| = p_i(p_i + E(Z_i))\|\Delta g\|.$$

A $|E(I_ig(Y_i+1)) - p_iE(g(Y_i+1))|$,,kicsi" feltevés általában nem túl erős, hiszen Y_i azon változók összege, amelyek ,,gyengén" függnek I_i-től.

Ezek alapján a következő tétel mondható ki:

2.1.10. Tétel. *A fenti jelölések mellett:*

$$d(\mathcal{L}(W), Po(\lambda)) \leq 2\sum_{i=1}^n \left(p_i^2 + p_iE(Z_i) + E(I_iZ_i)\right)\min\{1, 1/\lambda\} +$$

$$2\sum_{i=1}^n \eta_i \min\{1, 1/\lambda^{1/2}\},$$

ahol η_i *olyan, hogy* $|E(I_ig(Y_i+1)) - p_iE(g(Y_i+1))| \leq \eta_i\|g\|$ *teljesül, például* $\eta_i = E\left(\left|E(I_i|(I_j, j\in \boldsymbol{J}_i^2)) - p_i\right|\right)$.

2.2. Két tétel

A következő két állítás alapján is képet kaphatunk a Chen-Stein módszer bizonyos előnyeiről. Az első tételt a második a Chen-Stein módszer segítségével általánosítja, és a bizonyítás sem túlságosan bonyolódik el.

Legyenek most minden n-re $I_{1,n}, I_{2,n}, \ldots, I_{n,n}$ független indikátorok, $P(I_{i,n} = 1) = p_{i,n}$. Legyen $W^{(n)} = \sum_{i=1}^{n} I_{i,n}$ és $\sum_{i=1}^{n} p_{i,n} = \lambda$.

Mint már korábban említettük, ha $\forall i$-re $p_{i,n} = p = \frac{\lambda}{n}$, akkor

$$\lim_{n\to\infty} P(W^{(n)} = k) = Po(\lambda)\{k\},$$

sőt

$$\lim_{n\to\infty} d\left(Bin\left(n, \frac{\lambda}{n}\right), Po(\lambda)\right) = 0,$$

vagyis

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} \left|P(W^{(n)} = k) - Po(\lambda)\{k\}\right| = 0.$$

2.2.1. Tétel. *Tegyük fel, hogy $p_{i,n} = p = \frac{\lambda}{n}$ minden i-re. Legyen $h(k) \geq 0$ $\forall k \geq 0$-ra. Ekkor*

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} h(k) \left|P(W^{(n)} = k) - Po(\lambda)\{k\}\right| = 0$$

pontosan akkor teljesül, ha $\sum_{k=0}^{\infty} h(k) Po(\lambda)\{k\} < \infty$.

Bizonyítás. Legyen

$$S = \sup\left\{\frac{P(W^{(n)} = k)}{Po(\lambda)\{k\}} : k \geq 0, n \geq 1 + \lambda\right\}.$$

Megmutatjuk, hogy $S < \infty$.
Legyen $a_{k,n} = \dfrac{P(W^{(n)} = k)}{Po(\lambda)\{k\}}$. Ekkor

$$\frac{a_{k,n}}{a_{k-1,n}} = \frac{P(W^{(n)} = k)}{Po(\lambda)\{k\}} \cdot \frac{Po(\lambda)\{k-1\}}{P(W^{(n)} = k-1)} = \frac{k}{\lambda} \cdot \frac{\binom{n}{k} p^k (1-p)^{n-k}}{\binom{n}{k-1} p^{k-1} (1-p)^{n-k+1}} =$$

$$\frac{k}{\lambda} \cdot \frac{(n-k+1)p}{k(1-p)} = \frac{n-k+1}{n-\lambda},$$

ha $1 \leq k \leq n$ és $n \geq 1+\lambda$.

Legyen $s = 1 + \lfloor \lambda \rfloor$. Ezáltal $a_{k,n} \leq a_{s,n}$, ha $k \geq 0, n \geq 1+\lambda$. Mivel $\lim_{n\to\infty} P(W^{(n)} = k) = Po(\lambda)\{k\}$, ezért $\lim_{n\to\infty} a_{s,n} = 1$. Vagyis tényleg, $S < \infty$.

Tegyük fel most, hogy $\sum_{k=0}^{\infty} h(k)Po(\lambda)\{k\} < \infty$. A $\sum_{k=0}^{\infty} h(k) \left| P(W^{(n)} = k) - Po(\lambda)\{k\} \right|$ összegben a k-adik összeadandó korlátos:

$$0 \leq h(k) \left| P(W^{(n)} = k) - Po(\lambda)\{k\} \right| \leq (S+1)h(k)Po(\lambda)\{k\}.$$

Ezért a dominált konvergencia tételt használva megkapjuk a bizonyítandót.

Fordítva, ha $\lim_{n\to\infty} \sum_{k=0}^{\infty} h(k) \left| P(W^{(n)} = k) - Po(\lambda)\{k\} \right| = 0$, akkor

$$0 \leq \sum_{k=n+1}^{\infty} h(k)Po(\lambda)\{k\} \leq \sum_{k=0}^{\infty} h(k) \left| P(W^{(n)} = k) - Po(\lambda)\{k\} \right| \to 0,$$

ha $n \to \infty$. Tehát $\sum_{k=0}^{\infty} h(k)Po(\lambda)\{k\} < \infty$. $\square$

A Chen-Stein módszert használva nem kell feltennünk a $p_{i,n}$-ek egyenlőségét:

2.2.2. Tétel. *Legyen* $\lambda = \sum_{i=1}^{n} p_{i,n}$ *rögzített. Tegyük fel, hogy* $\max_{1\leq i\leq n} p_{i,n} \to 0$, *ha* $n \to \infty$. *Ekkor*

$$\lim_{n\to\infty} \sum_{k=0}^{\infty} h(k) \left| P(W^{(n)} = k) - Po(\lambda)\{k\} \right| = 0$$

pontosan akkor teljesül, ha $\sum_{k=0}^{\infty} h(k)Po(\lambda)\{k\} < \infty$.

Bizonyítás. Látszik, hogy az előző Tétel bizonyítása csak az $S < \infty$ egyenlőtlenségen múlott, ezért most csak ennek a belátásával foglalkozunk. Legyen $W_i^{(n)} = \sum_{j\neq i} I_{j,n}$. A korábban megállapítottak szerint minden korlátos g függvényre

$$E\left(W^{(n)} g(W^{(n)})\right) = \sum_{i=1}^{n} p_{i,n} E\left(g(W_i^{(n)} + 1)\right). \tag{1}$$

Legyen $g(l) = 1$, ha $l = k+1$ és $g(l) = 0$, ha $l \neq k+1$, ahol $0 \leq k \leq n$. Ebből azt kapjuk, hogy

$$(k+1)P(W^{(n)} = k+1) = \sum_{i=1}^{n} p_{i,n} P(W_i^{(n)} = k).$$

Mivel

$$1 \geq P(W_i^{(n)} = k | W^{(n)} = k) = \frac{P(W_i^{(n)} = k, W^{(n)} = k)}{P(W^{(n)} = k)} =$$

$$(1 - p_{i,n}) \frac{P(W_i^{(n)} = k)}{P(W^{(n)} = k)} \geq (1 - \max_{1 \leq i \leq n} p_{i,n}) \frac{P(W_i^{(n)} = k)}{P(W^{(n)} = k)},$$

ezért ebből következik, hogy

$$\frac{P(W^{(n)} = k+1)}{P(W^{(n)} = k)} \leq \frac{\lambda}{(k+1)(1 - \max_{1 \leq i \leq n} p_{i,n})}. \tag{2}$$

(1)-et tovább írva:

$$E\left(W^{(n)} g(W^{(n)})\right) = \sum_{i=1}^{n} p_{i,n} E\left(g(W_i^{(n)} + 1)\right) =$$

$$\lambda E\left(g(W^{(n)} + 1\right) + \sum_{i=1}^{n} p_{i,n} E\left(g(W_i^{(n)} + 1) - g(W^{(n)} + 1)\right) =$$

$$\lambda E\left(g(W^{(n)} + 1\right) + \sum_{i=1}^{n} p_{i,n} E\left[I_{i,n}\left(g(W_i^{(n)} + 1) - g(W_i^{(n)} + 2)\right)\right] =$$

$$\lambda E\left(g(W^{(n)} + 1\right) + \sum_{i=1}^{n} p_{i,n}^2 E\left(g(W_i^{(n)} + 1) - g(W_i^{(n)} + 2)\right).$$

Ezek alapján

$$\frac{\dfrac{P(W^{(n)} = k+1)}{Po(\lambda)\{k+1\}}}{\dfrac{P(W^{(n)} = k)}{Po(\lambda)\{k\}}} = \frac{(k+1)P(W^{(n)} = k+1)}{\lambda P(W^{(n)} = k)} =$$

$$1 + \frac{1}{\lambda P(W^{(n)} = k)} \sum_{i=1}^{n} p_{i,n}^2 \left(P(W_i^{(n)} = k) - P(W_i^{(n)} = k-1)\right).$$

(2)-höz teljesen hasonló módon kaphatjuk, hogy

$$\frac{P(W_i^{(n)} = k)}{P(W_i^{(n)} = k-1)} \leq \frac{\lambda - p_{i,n}}{k(1 - \max_{j \neq i} p_{j,n})} \leq \frac{\lambda}{k(1 - \max_i p_{i,n})}.$$

Az utóbbi kifejezés ≤ 1, ha $n \geq k \geq 1 + \lfloor \lambda \rfloor$ és $n \geq n_0$. Vagyis $P(W_i^{(n)} = k) - P(W_i^{(n)} = k-1) \leq 0$, ha $i = 1, 2, \ldots, n$. Ezért

$$\frac{\dfrac{P(W^{(n)} = k+1)}{Po(\lambda)\{k+1\}}}{\dfrac{P(W^{(n)} = k)}{Po(\lambda)\{k\}}} \leq 1.$$

Tehát ha $k \geq 1 + \lfloor \lambda \rfloor$ és $n \geq n_0$, akkor

$$\frac{P(W^{(n)} = k)}{Po(\lambda)\{k\}} \leq \frac{P(W^{(n)} = 1 + \lfloor \lambda \rfloor)}{Po(\lambda)\{1 + \lfloor \lambda \rfloor\}} \leq S < \infty,$$

Ebből az állítás már következik a dominált konvergencia tételt használva ugyanúgy, mint az előző tételben. $\square$

2.2.3. Megjegyzés. *Ha* $Z \sim Po(\lambda)$, *akkor*

$$\sum_{k=0}^{\infty} h(k) Po(\lambda)\{k\} = E\left(h(Z)\right).$$

3. fejezet

Legnagyobb együtthatók

3.1. Polinomok

A valós n-edfokú polinomok vizsgálata során érdemes azokra figyelmet fordítani, amelyek valamilyen szempontból extremális tulajdonságúak. Természetes módon adódhat például az a kérdés, hogy létezik-e ,,0-hoz legközelebbi" polinom, pontosabban, az 1 főegyütthatós n-edfokú polinomok között van-e legkisebb szuprémum-normájú a $[-1, 1]$ intervallumon. A válasz igen. Ráadásul minden n-re (-1-gyel való szorzástól eltekintve) egyértelműen létezik ilyen polinom. Ezek az ún. Csebisev polinomok.

Az előbbi állítást úgy is megfogalmazhatjuk, hogy a $[-1, 1]$-en adott szuprémum-normájú n-edfokú polinomok között a Csebisev polinom főegyütthatója a legnagyobb. Valóban, ilyen esetben az 1 főegyütthatós polinomok közül a Csebisev polinomot kell a legnagyobb abszolút értékű számmal megszoroznunk ahhoz, hogy a kívánt normát elérjük.

A felmerülő kérdés, amire ebben a részben választ keresünk, hogy lehet-e hasonló állítást kimondani az n-edfokú polinomok k-adik együtthatóiról. A Csebisev polinomok segítségével ez is megválaszolható.

Jelölje tehát T_n az n-edfokú (elsőfajú) Csebisev polinomot, amit $|x| \leq 1$

esetén a

$$T_n(x) = \cos(n \arccos x),$$

vagyis (az $x = \cos y$ helyettesítéssel) a $T_n(\cos y) = \cos(ny)$ képlettel definiálunk. Ekkor $T_n(\cos y)$ másképpen is felírható, $\cos y$ hatványait használva. Ehhez vegyük észre, hogy

$$\cos(ny) + i\sin(ny) = e^{iny} = \left(e^{iy}\right)^n = (\cos y + i \sin y)^n.$$

Az utóbbit a binomiális tétellel kifejtve és a $\sin^2 y = 1 - \cos^2 y$ azonosságot felhasználva, a valós részekre a következő egyenlőség adódik:

$$\cos(ny) = \sum_{l \geq 0} \binom{n}{4l} (\cos y)^{n-4l} (1-\cos^2 y)^{2l} -$$

$$\sum_{l \geq 0} \binom{n}{4l+2} (\cos y)^{n-4l-2} (1-\cos^2 y)^{2l+1} \tag{1}$$

Tehát $\cos(ny)$ felírható $\cos y$ polinomjaként:

$$T_n(\cos y) = \cos(ny) = \sum_{i=0}^{n} c_i^{(n)} (\cos y)^i.$$

A Csebisev polinomok ortogonálisak a következő értelemben:

3.1.1. Állítás.

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} dx = = \begin{cases} \pi & ha\ n = m = 0 \\ \frac{1}{2}\pi & ha\ n = m \neq 0 \\ 0 & ha\ n \neq m. \end{cases}$$

Bizonyítás. Az integrálban az $x = \cos y$ helyettesítést alkalmazva kapjuk, hogy

$$\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}} dx = \int_0^{\pi} \cos(ny)\cos(my) dy =$$

$$\frac{1}{2} \int_0^{\pi} \left\{ \cos\left((n-m)y\right) + \cos\left((n+m)y\right) \right\} dy.$$

Itt az utolsó lépésnél a $\cos(\alpha+\beta)+\cos(\alpha-\beta) = 2\cos\alpha\cos\beta$ trigonometrikus azonosságot alkalmaztuk. Ha az előbbi kifejezésben $n = m = 0$, akkor ez a kifejezés nyilván π. Ha $n = m \neq 0$, akkor

$$\frac{1}{2}\int_0^\pi \{\cos((n-m)y) + \cos((n+m)y)\}\, dy = \frac{\pi}{2} + \frac{1}{2}\int_0^\pi \cos(2ny)\, dy =$$

$$\frac{\pi}{2} + \frac{1}{2}\left[\frac{\sin(2ny)}{2n}\right]_0^\pi = \frac{\pi}{2}.$$

Ha $n \neq m$, akkor

$$\frac{1}{2}\int_0^\pi \{\cos((n-m)y) + \cos((n+m)y)\}\, dy =$$

$$\frac{1}{2}\left\{\left[\frac{\sin((n-m)y)}{n-m}\right]_0^\pi + \left[\frac{\sin((n+m)y)}{n+m}\right]_0^\pi\right\} = 0.$$

Ez pedig az állítást jelenti. $\square$

A $c_i^{(n)}$ együtthatókra adhatunk képletet, mint azt a következő állítás mutatja.

3.1.2. Állítás. *Ha $n \equiv i \ (mod\, 2)$, akkor*

$$c_i^{(n)} = (-1)^{\frac{n-i}{2}} \cdot 2^i \frac{n}{n+i}\binom{\frac{n+i}{2}}{i},$$

és ha $n \not\equiv i \ (mod\, 2)$, akkor

$$c_i^{(n)} = 0.$$

Ahhoz, hogy ezt beláthassuk, szükségünk lesz a Csebisev polinomok között fennálló rekurzióra.

3.1.3. Lemma. *$T_0(x) = 1$, $T_1(x) = x$, és ha $n \geq 1$, akkor*

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

Bizonyítás. Alkalmazva ismét a $\cos(\alpha+\beta) + \cos(\alpha-\beta) = 2\cos\alpha\cos\beta$ egyenlőséget, kapjuk, hogy

$$T_{n+1}(x) + T_{n-1}(x) = \cos((n+1)y) + \cos((n-1)y) =$$

$$2\cos y\cos(ny) = 2xT_n(x). \qquad \square$$

Bizonyítás. (A 3.1.2. állításé.) Az állítás második része (1)-ből nyilvánvaló. Az első részt teljes indukcióval bizonyítjuk. Ha $i = n$, akkor a 3.1.3. rekurzióból könnyen látszik, hogy $c_n^{(n)} = 2^{n-1}$, és ezt adja az állítás formulája is. Legyen most $i < n$. Mivel $n \equiv k \ (mod\, 2)$, ezért írjuk át $c_i^{(n)}$-t $c_{n-2j-2}^{(n)}$ alakba. Az indukciós feltevés: tegyük fel, hogy a formula igaz tetszőleges i esetén $c_{k-2i}^{(k)}$-re, ahol $k < n$ és igaz $c_{n-2i}^{(n)}$-re, ha $0 \le i \le j$. Belátjuk, hogy $i = j+1$-re is igaz. Mivel minden n esetén tudjuk, hogy $c_n^{(n)} = 2^{n-1}$, ez elég. A rekurzióból kapjuk, hogy

$$c_{n-2j-2}^{(n)} = 2c_{n-2j-3}^{(n-1)} - c_{n-2j-2}^{(n-2)}.$$

Az indukciós feltevés szerint

$$c_{n-2j-3}^{(n-1)} = (-1)^{j+1}2^{n-2j-3}\frac{n-1}{2n-2j-4}\binom{n-j-2}{n-2j-3} =$$

$$(-1)^{j+1}2^{n-2j-3}\frac{n-1}{2n-2j-4}\binom{n-j-2}{j+1} =$$

$$\frac{(-1)^{j+1}}{(j+1)!}2^{n-2j-4}(n-1)(n-j-3)\cdot\ldots\cdot(n-2j-2)$$

és

$$c_{n-2j-2}^{(n-2)} = (-1)^{j}2^{n-2j-2}\frac{n-2}{2n-2j-4}\binom{n-j-2}{n-2j-2} =$$

$$\frac{(-1)^{j}}{(j)!}2^{n-2j-3}(n-2)(n-j-3)\cdot\ldots\cdot(n-2j-1).$$

Vagyis

$$c_{n-2j-2}^{(n)} = 2c_{n-2j-3}^{(n-1)} - c_{n-2j-2}^{(n-2)} =$$

$$\frac{(-1)^{j+1}}{j!}2^{n-2j-3}(n-j-3)\cdot\ldots\cdot(n-2j-1)\left[\frac{(n-1)(n-2j-2)}{j+1} + (n-2)\right]$$

Vizsgáljuk a szögletes zárójelben lévő kifejezést:

$$\frac{1}{j+1}(n-1)(n-2j-2) + (n-2) = \frac{(n-1)n}{j+1} - \frac{(n-1)(2j+2)}{j+1} + n - 2 =$$

$$\frac{(n-1)n}{j+1} - n = \frac{1}{j+1}n(n-j-2).$$

Tehát

$$2c^{(n-1)}_{n-2j-3} - c^{(n-2)}_{n-2j-2} = \frac{(-1)^{j+1}}{(j+1)!} 2^{n-2j-3} n(n-j-2)(n-j-3)\cdot \ldots \cdot (n-2j-1).$$

Ez pedig pontosan az állítást adja:

$$c^{(n)}_{n-2j-2} = (-1)^{j+1} 2^{n-2j-2} \frac{n}{2n-2j-2} \binom{n-j-1}{n-2j-2}. \qquad \square$$

A Csebisev polinomok számos tulajdonsága közül még az explicit alakjukat említjük meg.

3.1.4. Tétel.

$$T_n(x) = \frac{1}{2}\left[\left(x+\sqrt{x^2-1}\right)^n + \left(x-\sqrt{x^2-1}\right)^n\right].$$

Bizonyítás. A 3.1.3. lemmából ez is könnyen adódik indukcióval. Ha $n=0$ vagy 1, akkor az állítás nyilvánvaló. Tegyük fel, hogy igaz az állítás igaz T_k-ra, ha $k \leq n$. Ekkor a rekurzió alapján:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x) = 2x\frac{1}{2}\left[\left(x+\sqrt{x^2-1}\right)^n + \left(x-\sqrt{x^2-1}\right)^n\right] -$$

$$\frac{1}{2}\left[\left(x+\sqrt{x^2-1}\right)^{n-1} + \left(x-\sqrt{x^2-1}\right)^{n-1}\right] =$$

$$\frac{1}{2}\left(x+\sqrt{x^2-1}\right)^{n-1}\left(2x(x+\sqrt{x^2-1})-1\right) +$$

$$\frac{1}{2}\left(x-\sqrt{x^2-1}\right)^{n-1}\left(2x(x-\sqrt{x^2-1})-1\right) =$$

$$\frac{1}{2}\left(x+\sqrt{x^2-1}\right)^{n-1}(x+\sqrt{x^2-1})^2 + \frac{1}{2}\left(x-\sqrt{x^2-1}\right)^{n-1}(x-\sqrt{x^2-1})^2 =$$

$$\frac{1}{2}\left[\left(x+\sqrt{x^2-1}\right)^{n+1} + \left(x-\sqrt{x^2-1}\right)^{n+1}\right]. \qquad \square$$

[3]-ben megtalálható, hogy a következő állítás V. A. Markov eredménye 1892-ből.

Ha $p(x) = \sum_{i=0}^n a_i x^i,\ \ a_n \neq 0$ valós együtthatós polinom, akkor legyen

$$p^+(x) = a_n x^n + a_{n-2}x^{n-2} + \ldots + a_{n-2\lfloor\frac{n}{2}\rfloor} x^{n-2\lfloor\frac{n}{2}\rfloor}$$

és

$$p^-(x) = a_{n-1}x^{n-1} + a_{n-3}x^{n-3} + \ldots + a_{n-2\lceil\frac{n}{2}\rceil+1} x^{n-2\lceil\frac{n}{2}\rceil+1}$$

Legyen

$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_{k+1}x^{k+1} + x^k + a_{k-1}x^{k-1} + \ldots + a_1 x + a_0$

valós együtthatós polinom, $a_n \neq 0$.

3.1.5. Állítás. *Tegyük fel, hogy* $0 < k < n$.

Ha $n \equiv k \pmod 2$, *akkor* $\max_{x\in[-1,1]} |p(x)| \geq \frac{1}{|c_k^{(n)}|}$.

Ha $n \not\equiv k \pmod 2$, *akkor* $\max_{x\in[-1,1]} |p(x)| \geq \frac{1}{|c_k^{(n-1)}|}$.

Továbbá, ha az első egyenlőtlenségben egyenlőség teljesül, akkor $p(x) = \frac{1}{c_k^{(n)}} T_n(x)$.

A második egyenlőtlenségben, ha $n > 2$, *akkor nem teljesülhet egyenlőség.*

Bizonyítás. Legyen $n \equiv k \pmod 2$. Indirekt tegyük fel, hogy $\max_{x\in[-1,1]} |p(x)| < \frac{1}{|c_k^{(n)}|}$. Először belátjuk, hogy

$$\max_{x\in[-1,1]} |p(x)| \geq \max_{x\in[-1,1]} |p^+(x)|.$$

(Megjegyezzük, hogy a következőhöz teljesen hasonló bizonyítással belátható, hogy $\max_{x\in[-1,1]} |p(x)| \geq \max_{x\in[-1,1]} |p^-(x)|$.)

Mivel $\exists x_0 \in [-1,1]$, amelyre $|p^+(x_0)| = \max_{x\in[-1,1]} |p^+(x)|$, ezért $p^+(x_0) = (-1)^n p^+(-x_0)$ és $p^-(x_0) = (-1)^{n-1} p^-(-x_0)$ miatt: $\max\{|p(x_0)|, |p(-x_0)|\} = \max\{|p^+(x_0) + p^-(x_0)|, |p^+(-x_0) + p^-(-x_0)|\} = \max\{|p^+(x_0) + p^-(x_0)|, |p^+(x_0) - p^-(x_0)|\} \geq |p^+(x_0)|$, vagyis $\max_{x\in[-1,1]} |p(x)| \geq \max_{x\in[-1,1]} |p^+(x)|$.

Emiatt $\max_{x\in[-1,1]} |p^+(x)| < \frac{1}{|c_k^{(n)}|}$, továbbá $p^+(x)$-ben x^k együtthatója 1.

A $T_n(x)$ polinomra igaz, hogy $\max_{x\in[-1,1]} |T_n(x)| = 1$, és $T_n(\cos\frac{i\pi}{n}) = (-1)^i$, $i = 0, 1, \ldots, n$. Ezek alapján az $\frac{1}{c_k^{(n)}} T_n(x) - p^+(x)$ polinomnak van gyöke a $(\cos\frac{i\pi}{n}, \cos\frac{(i+1)\pi}{n})$, $\quad i = 0, 1, \dots, n-1$ intervallumon. Legyenek ezek a gyökök $y_1, -y_1, y_2, -y_2, \dots, y_{\lfloor\frac{n}{2}\rfloor}, -y_{\lfloor\frac{n}{2}\rfloor}$, és ha n páratlan, akkor $y_0 = 0$. Itt $y_i \neq 0$, ha $i \neq 0$. Mivel van n gyök, ezért $\frac{c_n^{(n)}}{c_k^{(n)}} - a_n \neq 0$.

Ha n páratlan, akkor

$$\frac{1}{c_k^{(n)}} T_n(x) - p^+(x) = \left(\frac{c_n^{(n)}}{c_k^{(n)}} - a_n\right) x \prod_{i=1}^{\lfloor\frac{n}{2}\rfloor} (x - y_i)(x + y_i) =$$

$$\left(\frac{c_n^{(n)}}{c_k^{(n)}} - a_n\right) x \prod_{i=1}^{\lfloor\frac{n}{2}\rfloor} (x^2 - y_i^2).$$

Ha n páros, akkor

$$\frac{1}{c_k^{(n)}} T_n(x) - p^+(x) = \left(\frac{c_n^{(n)}}{c_k^{(n)}} - a_n\right) \prod_{i=1}^{\frac{n}{2}} (x - y_i)(x + y_i) =$$

$$\left(\frac{c_n^{(n)}}{c_k^{(n)}} - a_n\right) \prod_{i=1}^{\frac{n}{2}} (x^2 - y_i^2).$$

Mivel $\frac{1}{c_k^{(n)}} T_n(x) - p^+(x)$-ben x^k együtthatója 0, ezért azt kapjuk, hogy

$$0 = \left(\frac{c_n^{(n)}}{c_k^{(n)}} - a_n\right) (-1)^{\frac{n-k}{2}} \sum_{1 \le i_1 < i_2 < \dots < i_{\frac{n-k}{2}} \le \lfloor\frac{n}{2}\rfloor} y_{i_1}^2 y_{i_2}^2 \dots y_{i_{\frac{n-k}{2}}}^2,$$

ami ellentmondás.

Legyen most $n \not\equiv k \quad (mod\, 2)$. Az előzőekből már következik, hogy

$$\max_{x \in [-1,1]} |p(x)| \ge \max_{x \in [-1,1]} |p^-(x)| \ge \frac{1}{|c_k^{(n-1)}|}.$$

Az egyenlőség kérdése van még hátra.

Legyen $n \equiv k \quad (mod\, 2)$ Tegyük fel, hogy $p^+(x) \neq \frac{1}{c_k^{(n)}} T_n(x)$. Azt is feltehetjük, hogy az $\frac{1}{c_k^{(n)}} T_n(x) - p^+(x)$ polinomnak nincs a $[\cos\pi, \cos\frac{(n-1)\pi}{n})$, $(\cos\frac{(n-1)\pi}{n}, \cos\frac{(n-2)\pi}{n}), \dots, (\cos\frac{\pi}{n}, 1]$ intervallumok mindegyikén gyöke, hiszen azt az esetet már tulajdonképpen láttuk. Ha egy ilyen intervallumon nincs gyök, akkor valamelyik végpontja gyök kell, hogy legyen. Azaz $\exists i_0$: $\frac{1}{c_k^{(n)}} T_n(\cos \frac{i_0\pi}{n}) = p^+(\cos \frac{i_0\pi}{n})$. Állítjuk, hogy ebben az esetben $\frac{1}{c_k^{(n)}} T_n(x) - p^+(x)$-nek $\cos(\frac{i_0\pi}{n})$ legalább kétszeres gyöke. Ugyanis

$$\left(\frac{1}{c_k^{(n)}} T_n(x) - p^+(x) \right)' \Bigg|_{x=\cos(\frac{i_0\pi}{n})} = \frac{1}{c_k^{(n)}} T_n'(\cos \frac{i_0\pi}{n}) - (p^+)'(\cos \frac{i_0\pi}{n}) = 0,$$

mert $\cos(\frac{i_0\pi}{n})$ szélsőértékhelye a $T_n(x)$ és a $p^+(x)$ polinomnak is. Vagyis minden $[\cos(\frac{(i+1)\pi}{n}), \cos(\frac{i\pi}{n})], i = 1, \ldots, n-1$ intervallumon van gyök, és ha ez az egyik végpontban van, akkor ez legalább kétszeres. Tehát az $\frac{1}{c_k^{(n)}} T_n(x) - p^+(x)$ polinomnak van n gyöke a $[-1, 1]$ intervallumon (multiplicitással számolva). Nyilván n-nél több nem lehet, és ebből az is látszik, hogy ekkor minden gyök (így esetleg 0 is) legfeljebb kétszeres. Ugyanúgy, mint korábban, kapjuk, hogy

$$0 = \left(\frac{c_n^{(n)}}{c_k^{(n)}} - a_n \right) (-1)^{\frac{n-k}{2}} \sum_{1 \le i_1 < \ldots < i_{\frac{n-k}{2}} \le \lfloor \frac{n}{2} \rfloor} y_{i_1}^2 \ldots y_{i_{\frac{n-k}{2}}}^2 .$$

Mivel 0 legfeljebb kettő multiplicitású zérushely, és a feltételek miatt $n > 2$, ezért a fenti szummában van 0-tól különböző tag. Ez $\frac{c_n^{(n)}}{c_k^{(n)}} \neq a_n$ miatt ellentmondás. Tehát $p^+(x) = \frac{1}{c_k^{(n)}} T_n(x)$. A $\max_{x\in[-1,1]} |p(x)| \ge \max_{x\in[-1,1]} |p^+(x)|$ egyenlőtlenség bizonyításából látszik, hogy most $p^-(\cos \frac{i\pi}{n}) = 0$, $i = 0, 1, \ldots, \lfloor \frac{n-1}{2} \rfloor, \lceil \frac{n+1}{2} \rceil, \ldots, n$. Ez legalább n gyököt jelent, amiből következik, hogy $p^- \equiv 0$.

Vizsgáljuk most a második egyenlőtlenséget. Látszik, hogy $\max_{x\in[-1,1]} |p(x)| = \frac{1}{|c_k^{(n-1)}|}$ csak akkor teljesülhet, ha $p^-(x) = \frac{1}{c_k^{(n-1)}} T_n(x)$. Hasonlóan, mint az előbb, kapjuk, hogy $p^+(\cos \frac{i\pi}{n-1}) = 0, i = 0, 1, \ldots, n-1$. Ezzel p^+ összes gyökét meghatároztuk. ($p^+ \neq 0$, mert $a_n \neq 0$.) Ha $n > 2$, akkor $\cos \frac{\pi}{n-1} \in (-1, 1)$, és így $p'(\cos \frac{\pi}{n-1}) = (p^-)'(\cos \frac{\pi}{n-1}) + (p^+)'(\cos \frac{\pi}{n-1}) \neq 0$, mert p^+ $\cos \frac{\pi}{n-1}$-ben előjelet vált. Ezért $\cos \frac{\pi}{n-1}$ nem szélsőértékhelye p-nek.$\square$

3.1.6. Megjegyzés. *$n = 2$ és $k = 1$ esetén $p(x) = ax^2 + x - a, a \in [-\frac{1}{2}, \frac{1}{2}] \setminus \{0\}$ olyan polinom, amire $\max_{x\in[-1,1]} |p(x)| = 1 = \frac{1}{|c_1^{(1)}|}$.*

3.1.7. Megjegyzés. *A második egyenlőtlenségben a konstans nem javítható, mint azt az* $\frac{1}{m}x^n + \frac{1}{c_k^{(n-1)}}T_{n-1}(x), m = 1, 2, \dots$ *polinomsorozat mutatja.*

Jelölje P_n a legfeljebb n-edfokú polinomok halmazát.

3.1.8. Következmény. *Ha tekintjük azokat P_n-beli polinomokat, amelyeknek a $[-1, 1]$ intervallumon 1 a szuprémum-normájuk, akkor ezek között a megfelelő Csebisev polinom k-adfokú tagja együtthatójának abszolút értéke a legnagyobb. Pontosabban:*

ha $n \equiv k \pmod 2$, *akkor* $\max\{|a_k| : p \in P_n, \max_{x\in[-1,1]} |p(x)| = 1\} = |c_k^{(n)}|$,

ha $n \not\equiv k \pmod 2$, *akkor* $\max\{|a_k| : p \in P_n, \max_{x\in[-1,1]} |p(x)| = 1\} = |c_k^{(n-1)}|$.

Bizonyítás. Ha $n \equiv k \pmod 2$, akkor egy tetszőleges p polinomot, (amire $a_k \neq 0$) szorozzunk be azzal a c valós számmal, amire $ca_k = c_k^{(n)}$. Ebből azt fogjuk kapni az előző állítás alapján, hogy $\max_{x\in[-1,1]} |cp(x)| \geq 1$, ezért $|c| \geq 1$, és egyenlőség pontosan akkor áll fenn, ha $p = T_n$. Ha $n \not\equiv k \pmod 2$, akkor azt a c-t vegyük, amelyre $ca_k = c_k^{(n-1)}$, és ezek után ugyanúgy folytathatjuk a bizonyítást, mint az előbb.□

Nekünk valójában a $[0, 1]$ interallumon lesz szükségünk egy, az előbbihez teljesen hasonló eredményre.

Legyen $T_n(2x - 1) = \sum_{i=0}^{n} d_i^{(n)} x^i$.

3.1.9. Állítás. *Tegyük fel, hogy* $0 < k < n$. *Legyen*
$p(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_{k+1}x^{k+1} + x^k + a_{k-1}x^{k-1} + \ldots + a_1 x + a_0$
valós együtthatós polinom, $a_n \neq 0$. *Ekkor*

$$\max_{x\in[0,1]} |p(x)| \geq \frac{1}{|d_k^{(n)}|}.$$

Egyenlőség pontosan akkor áll fenn, ha $p(x) = T_n(2x - 1)$.

Bizonyítás. A bizonyítás az előző állítás bizonyításának egyszerűsített változata, ezért csak vázoljuk azt.

$T_n(\cos\frac{i\pi}{n}) = (-1)^i$, azaz $T_n(2\cos^2\frac{i\pi}{2n} - 1) = (-1)^i$, $\quad i = 0, 1, \ldots, n$. Ha

$$max_{x\in[0,1]}|p(x)| < \max_{x\in[0,1]}\left|\frac{1}{d_k^{(n-1)}}T_n(2x-1)\right|,$$

akkor $\frac{1}{d_k^{(n-1)}}T_n(2x-1)-p(x)$-nek van n különböző gyöke a $[0,1]$ intervallumon. A gyökök: $y_1, y_2, \ldots, y_n \geq 0$. Ezzel

$$\frac{1}{d_k^{(n-1)}}T_n(2x-1) - p(x) = \left(\frac{d_n^{(n)}}{d_k^{(n)}} - a_n\right)\prod_{i=1}^{n}(x - y_i),$$

a k-adfokú tag együtthatójára pedig

$$0 = \left(\frac{d_n^{(n)}}{d_k^{(n)}} - a_n\right)(-1)^{n-k}\sum_{1\leq i_1<i_2<\ldots<i_{n-k}\leq n} y_{i_1}y_{i_2}\ldots y_{i_{n-k}}.$$

Ez pedig ellentmondás. Az, hogy egyenlőség pontosan $p(x) = T_n(2x-1)$ esetén áll fenn, ugyanúgy látható be, mint az előző állításban.□

3.1.10. Következmény. *Ha tekintjük azokat P_n-beli polinomokat, amelyeknek a $[0,1]$ intervallumon 1 a szuprémum-normájuk, akkor ezek között a $T_n(2x-1)$ polinom k-adfokú tagja együtthatójának abszolút értéke a legnagyobb.*

Bizonyítás. Ugyanúgy bizonyíthatunk, mint az előbb. □

A $T_n(2x-1)$ polinom együtthatóit is meghatározhatjuk a korábbiak alapján. Erre egy becslés miatt lesz szükségünk.

3.1.11. Állítás.

$$d_k^{(n)} = (-1)^{n-k}2^{2k}\frac{n}{k+n}\binom{k+n}{2k}.$$

Bizonyítás. Az állítást ismét a a 3.1.3. rekurzió segítségével bizonyíthatjuk teljes indukcióval, mint a a 3.1.2. állításban. A $T_n(2x-1)$ polinomra tehát a következőt kapjuk:

$$T_n(2x-1) = 2(2x-1)T_{n-1}(2x-1) - T_{n-2}(2x-1).$$

Ebből könnyen látható, hogy $d_n^{(n)} = 2^{2n-1}$ minden n-re. A fenti formula is ezt adja. Az együtthatót írjuk át ismét $d_{n-i}^{(n)}$ alakba. Ekkor az együtthatókra a következő összefüggés adódik:

$$d_{n-i}^{(n)} = 4d_{n-i-1}^{(n-1)} - 2d_{n-i}^{(n-1)} - d_{n-i}^{(n-2)}, \tag{2}$$

ahol persze $d_l^{(m)} = 0$, ha $l > m$. Az indukciós feltevés a következő: tegyük fel, hogy az állítás igaz $d_{n-i}^{(m)}$-re, ha $m < n$ és i tetszőleges, illetve ha $m = n$ és $0 \leq i \leq j$. Belátjuk, hogy ekkor igaz $m = n$ és $i = j+1$-re is. A (2) képletben szereplő együtthatók közül az indukciós feltevés szerint a jobb oldalon álló együtthatókra érvényes a formula.Ha $j + 1 = 1$, akkor az állítás a

$$d_{n-1}^{(n)} = 4d_{n-2}^{(n-1)} - 2d_{n-2}^{(n-2)}$$

alakot ölti. A jobb oldal az indukciós feltevés miatt a következő:

$$4d_{n-2}^{(n-1)} - 2d_{n-2}^{(n-2)} = 4 \cdot (-1)2^{2n-4}\frac{n-1}{2n-3}\binom{2n-3}{2n-4} - 2 \cdot 2^{2n-3} = -n2^{2n-2}.$$

Ezt adja az állítás is. Ha $j + 1 \geq 2$, akkor írjuk fel a (2) jobb oldalán álló tagokat egyenként:

$$4d_{n-j-2}^{(n-1)} = 4(-1)^{j+1}2^{2n-2j-4}\frac{n-1}{2n-j-3}\binom{2n-j-3}{2n-2j-4} =$$

$$\frac{(-1)^{j+1}}{(j+1)!}2^{2n-2j-2}(n-1)(2n-j-4) \cdot \ldots \cdot (2n-2j-3),$$

$$2d_{n-j-1}^{(n-1)} = 2(-1)^{j}2^{2n-2j-2}\frac{n-1}{2n-j-2}\binom{2n-j-2}{2n-2j-2} =$$

$$\frac{(-1)^{j}}{j!}2^{2n-2j-1}(n-1)(2n-j-3) \cdot \ldots \cdot (2n-2j-1),$$

$$d_{n-j-1}^{(n-2)} = (-1)^{j-1}2^{2n-2j-2}\frac{n-2}{2n-j-3}\binom{2n-j-3}{2n-2j-2} =$$

$$\frac{(-1)^{j-1}}{(j-1)!}2^{2n-2j-2}(n-2)(2n-j-4) \cdot \ldots \cdot (2n-2j-1).$$

Ebből kapjuk, hogy

$$4d_{n-j-2}^{(n-1)} - 2d_{n-j-1}^{(n-1)} - d_{n-j-1}^{(n-2)} =$$

$$\frac{(-1)^{j+1}}{(j+1)!}2^{2n-2j-2}(2n-j-4)\cdot\ldots\cdot(2n-2j-1)\cdot$$
$$((n-1)(2n-2j-2)(2n-2j-3)+2(j+1)(n-1)(2n-j-3)-$$
$$j(j+1)(n-2)).$$

Számoljuk ki külön-külön a zárójelben álló három tényezős szorzatokat:

$$(n-1)(2n-2j-2)(2n-2j-3)=4n^3+4j^2n+18jn-6n^2+16n-4j^2-10j-6,$$

$$2(j+1)(n-1)(2n-j-3)=4jn^2-2j^2n+4n^2+2j^2-12jn+8j-10n+6,$$

$$-j(j+1)(n-2)=-j^2n+2j^2-jn+2j.$$

Ezek összege:

$$4n^3+j^2n+5jn-10n^2-4jn^2+6n=n(2n-j-2)(2n-j-3).$$

Ez azt jelenti, hogy

$$4d_{n-j-2}^{(n-1)}-2d_{n-j-1}^{(n-1)}-d_{n-j-1}^{(n-2)}=\frac{(-1)^{j+1}}{(j+1)!}2^{2n-2j-2}n(2n-j-2)\cdot\ldots\cdot(2n-2j-1).$$

Tehát

$$d_{n-j-1}^{(n)}=(-1)^{j+1}2^{2n-2j-2}\frac{n}{2n-j-1}\binom{2n-j-1}{2n-2j-2},$$

ami az állítás volt. $\square$

Az állítás segítségével becslést kaphatunk $|d_k^{(n)}|$-re:

$$|d_k^{(n)}|=2^{2k}\frac{n}{k+n}\binom{k+n}{2k}\leq\frac{2^{2k}n^{2k}}{(2k)!},$$

a számtani és mértani közepek közötti egyenlőtlenség miatt.

3.2. Generátorfüggvények együtthatói

Az előző eredmények tulajdonképpen a következőt jelentik a mi számunkra. Legyenek $\mathbf{p}=(p_0,p_1,\ldots,p_{n_1})$ és $\mathbf{q}=(q_1,q_2,\ldots,q_{n_2})$ diszkrét valószínűségeloszlások a $g_\mathbf{p}$ és $g_\mathbf{q}$ generátorfüggvényekkel. Tegyük fel, hogy ezekben csak véges sok 0-tól különböző valószínűség szerepel. Jelölje $g_{\mathbf{p}} - g_{\mathbf{q}}$ szuprémum-normáját a $[0,1]$ intervallumon Δ. A k-adfokú tag együtthatójára ekkor

$$|a_k| \le |d_k^{(n)}|\Delta,$$

ahol $n = \max\{n_1, n_2\}$. Ezt az eredményt próbáljuk meg a következőkben valamivel általánosabbá tenni.

Legyen

$$\mathcal{F} = \{f \mid f(x) = \sum_{i=0}^{\infty} a_i x^i, \quad \sum_{i=0}^{\infty} |a_i| \le 2, \quad \sum_{i=0}^{\infty} a_i = 0\}.$$

(Nyilván $g_{\mathbf{p}} - g_{\mathbf{q}} \in \mathcal{F}$.)

Látható, hogy $|f(x)| = |\sum_{i=0}^{\infty} a_i x^i| \le \sum_{i=0}^{\infty} |a_i| \le 2 < \infty, \forall x \in [0,1]$, ha $f \in \mathcal{F}$. Ebből a Weierstrass kritérium miatt f hatványsora egyenletesen konvergens a $[0,1]$ intervallumon.

Legyen $f \in \mathcal{F}$, amelyre $\max_{x\in[0,1]} |f(x)| = \Delta$. Legyen h olyan függvény, amelyre $h(\Delta) > \Delta \quad \forall \Delta > 0$-ra. Legyen továbbá ϱ tetszőleges függvény. Tudjuk, hogy $\exists N$, hogy $\forall m \ge N$ esetén $|\sum_{i=m+1}^{\infty} a_i x^i| \le h(\Delta) - \Delta$.

3.2.1. Állítás. *Tegyük fel, hogy $0 < k \le N$. Ha $N \le \varrho(\Delta)$, akkor*

$$|a_k| \le \frac{2^{2k}}{(2k)!} \cdot h(\Delta) \left(\varrho(\Delta)\right)^{2k}.$$

Bizonyítás.

$$\left| \sum_{i=0}^{N} a_i x^i \right| = \left| f(x) - \sum_{i=N+1}^{\infty} a_i x^i \right| \le \Delta + h(\Delta) - \Delta = h(\Delta).$$

Ebből adódóan $|a_k| \le h(\Delta) \cdot |d_k^{(N)}| \le h(\Delta) \cdot \frac{2^{2k} N^{2k}}{(2k)!} \le \frac{2^{2k}}{(2k)!} h(\Delta) \left(\varrho(\Delta)\right)^{2k}$. □

[6]-ben és [7]-ben felső illetve alsó becslések találhatók $|a_k|$-re.

[6]-ben pédául az olvasható, hogy tetszőleges $0 < \varepsilon \le 1$ esetén létezik $c_k(\varepsilon)$, $k = 0, 1, \ldots$, amire

$$|a_k| \le c_k(\varepsilon)\Delta^{1-\varepsilon}$$

tetszőleges $f \in \mathcal{F}$ esetén.

[7]-ben megtalálható, hogy tetszőleges $0 < C < \dfrac{1}{2(2k)!\left[\log(1+\sqrt{2})\right]^{2k}}$ esetén, ha $\Delta > 0$,,elég kicsi," akkor van olyan $f \in \mathcal{F}$, amire

$$|a_k| \geq C\Delta \left(\log \frac{1}{\Delta}\right)^{2k}.$$

A 3.2.1. állítást alkalmazzuk a $h(\Delta) = 2\Delta$ és $\varrho(\Delta) = \log \frac{1}{\Delta}$ választásokkal. Ezzel azt kapjuk, hogy ha $N \leq \log \frac{1}{\Delta}$, akkor

$$|a_k| \leq \frac{2^{2k+1}}{(2k)!} \cdot \Delta \left(\log \frac{1}{\Delta}\right)^{2k}.$$

A függvények ezen választása mellett a következőt állapíthatjuk meg. Adott $f \in \mathcal{F}$ esetén tetszőleges $C > 0$ konstansra $C \cdot f$-hez tartozhat ugyanaz az N, mint f-hez. Ez azt jelenti, hogy ehhez az f-hez létezik olyan $C = C(f)$, amivel $N < \log \frac{1}{C\Delta}$ lesz. Ebből pedig kaphatunk becslést $|a_k|$-re:

$$\frac{C|a_k|}{C\Delta\left(\log \frac{1}{C\Delta}\right)^{2k}} = \frac{|a_k|}{\Delta\left(\log \frac{1}{C\Delta}\right)^{2k}} \leq \frac{2^{2k+1}}{(2k)!}.$$

Valójában a becslés során csak azt használtuk, hogy $\varrho(\Delta) = \log \frac{1}{\Delta}$-ra $|\varrho(\Delta)| \to \infty$, ha $\Delta \searrow 0$. Ha C-re adott valamilyen alsó becslés, akkor ez egy felső becslést jelent $|a_k|$-re is.

Egy kissé más megközelítés a következő. Legyen

$$\varphi(x) = \frac{1}{x^{2k}} \sup_{n \geq x} n^{2k} \sum_{i>n} |a_i|.$$

Ez egy monoton csökkenő függvény. Ha igaz, hogy $\sum_{i=0}^{\infty} |a_i| i^{2k} < \infty$, akkor $\lim_{x\to\infty} \varphi(x) = 0$. Legyen $\varphi^{-1}(x) = \min\{l > 0 : \varphi(l) \leq x\}$. Vegyük észre, hogy φ^{-1} értéke mindig egész szám.

3.2.2. Állítás. *Ha $k \leq \varphi^{-1}(\Delta)$, akkor $|a_k| \leq \frac{2^{2k+1}}{(2k)!}\Delta(\varphi^{-1}(\Delta))^{2k}$.*

Bizonyítás. Válasszuk N-et úgy, hogy $N \geq \varphi^{-1}(\Delta)$ legyen. Ezzel

$$\left|\sum_{i=0}^{N} a_i x^i\right| \leq \Delta + \sum_{i>N} |a_i| \leq \Delta + \varphi(N) \leq 2\Delta.$$

Ez azt jelenti, hogy $|a_k| \leq \frac{2^{2k}}{(2k)!} N^{2k} \cdot 2\Delta$. Vagyis az $N = \varphi^{-1}(\Delta)$ választással a bizonyítás kész. □

Az előbbi állítás segítségével egy további felső becslést adhatunk $|a_k|$-re.

3.2.3. Állítás. *Tegyük fel, hogy van olyan* $0 < \varepsilon \leq 2$, *amire* $R(\varepsilon) = \sum_{i=0}^{\infty} |a_i| e^{i\varepsilon} < \infty$. *Ha* $\Delta \leq e^{-2k} R(\varepsilon)$, *akkor*

$$|a_k| \leq \frac{2^{2k+1}}{(2k)!} \Delta \left\lceil \frac{1}{\varepsilon} \log\left(\frac{R(\varepsilon)}{\Delta}\right)\right\rceil^{2k}.$$

Bizonyítás. Vegyük észre, hogy $\sum_{i>n} |a_i| < e^{-n\varepsilon} R(\varepsilon)$, azaz $n^{2k} \sum_{i>n} |a_i| < n^{2k} e^{-n\varepsilon} R(\varepsilon)$. A $t^{2k} e^{-t\varepsilon}$ függvény a maximumát $t = \frac{2k}{\varepsilon}$-ban veszi fel. Emiatt $\varphi(x) < e^{-x\varepsilon} R(\varepsilon)$, ha $x \geq \frac{2k}{\varepsilon}$.

Válasszuk most N-et a következőképpen: $N = \left\lceil \frac{1}{\varepsilon} \log\left(\frac{R(\varepsilon)}{\Delta}\right)\right\rceil$. Ezzel $N \geq \frac{1}{\varepsilon} \log\left(\frac{R(\varepsilon)}{\Delta}\right) \geq \frac{2k}{\varepsilon}$ lesz, vagyis $\varphi(N) \leq \varphi\left(\frac{1}{\varepsilon} \log\left(\frac{R(\varepsilon)}{\Delta}\right)\right) < \Delta$. Tehát $\varphi^{-1}(\Delta) \leq N$. Ez pedig adja az állítást. □

Tartalomjegyzék

Irodalomjegyzék

[1] A. D. Barbour, L. Holst, S. Janson: *Poisson Approximation,* Oxford Univ. Press, 1972

[2] L. H. Y. Chen: On the Convergence of Poisson Binomial to Poisson Distributions, *Ann. Probab.* **2** (1974), 178-180

[3] D. Dryanov, R. Fournier: Refinement of an Inequality of P. L. Chebyshev, *Acta Math. Hungar.* **122** (2009), 59-69

[4] Lovász László: *Kombinatorikai Problémák és Feladatok,* Typotex, Budapest, 2000

[5] Michaletzky György: *Kockázati Folyamatok,* ELTE Eötvös Kiadó, Budapest, 1995

[6] T. F. Móri: Bonferroni Inequalities and Deviations of Discrete Distributions, *J. Appl. Probab.* **33** (1996), 115-121

[7] T. F. Móri: Deviation of Discrete Distributions – Positive and Negative Results, *Statist. Probab. Lett.* **79** (2009), 1089-1096

[8] G. Simons, N. L. Johnson: On the Convergence of Binomial to Poisson Distributions, *Ann. Math. Statist.* **42** (1971) 1735-1736

[9] Stoyan Gisbert, Takó Galina: *Numerikus Módszerek 1.,* Typotex, Budapest 2002