

VIZSGÁLATOK AZ EMPIRIKUS ELOSZLÁSFÜGGVÉNYRŐL*

Írta: CSÁKI ENDRE

Bevezetés

A matematikai statisztika alapvető feladata a mintából az alapsokaságra, elméleti eloszlásra való megbízható következtetések vizsgálata. Fontos szerepet játszik itt, különösen a nemparaméteres módszereknel az empirikus eloszlásfüggvény, amely az elméleti eloszlásfüggvénynek sok szempontból jó (torzítatlan, konzisztens stb.) becslése.

Az empirikus eloszlásfüggvénynek egy rögzített x pontban való viselkedése a valószínűségszámítás klasszikus eredményeivel, a *Bernoulli* (binomiális) — szkéma segítségével tárgyalható. A statisztika szempontjából nagyobb jelentőségűek azok az eredmények, melyek az empirikus eloszlásfüggvénynek x -ben egyenletes tulajdonságaira vonatkoznak. E téren alapvető KOLMOGOROV [64] és GLIVENKO [54] munkája 1933-ból. Azóta e kérdéskörben számos dolgozat látott napvilágot, melyet még csak felsorolni is lehetetlen vállalkozás lenne. A rendkívül kiterjedt irodalom jelzi a téma fontosságát mind elméleti, mind gyakorlati szempontból. Irodalomjegyzékünk az értekezés végén igyekszik felsorolni azokat a fontosabb irodalmi forrásokat, melyek az értekezés témájával és tárgyalásmódjával kapcsolatban állnak.

Az értekezésben egymintás, egyoldali eltérésekkel kapcsolatos kérdések egységes tárgyalását igyekszünk megadni. A tárgyalás egy igen elemi kombinatorikus módszeren, az ún. ballot-lemmán alapszik, amely lényegében TAKÁCS-tól származik.

Az 1. §-ban pontos formulákat adunk annak valószínűségére, hogy egy $(0, 1)$ -ben egyenletes eloszlásból vett minta empirikus eloszlásfüggvénye egy adott törtvonal felett halad. Ezek speciális esetként tartalmazzák az általunk ismert, egyoldali eltérésre és relatív eltérésre vonatkozó formulákat, a pontos irodalmi hivatkozásokkal együtt.

A 2. §-ban a pontos eloszlásokból származtatunk határeloszlásokat, ill. határérték tételeket.

A 3. §-ban erőfüggvényeket vizsgálunk. Pontos és aszimptotikus formulákat egyaránt adunk bizonyos alternatívák esetére, és tárgyaljuk az efficiencia különböző értelmezéseit.

A 4. §-ban empirikus eloszlásfüggvényre vonatkozó iterált logaritmus tételekkel foglalkozunk.

Az 5. §-ban néhány kiegészítő jellegű tárgyalás kerül sorra. Itt vizsgáljuk pl. a maximális eltéréssel kapcsolatos egyéb statisztikák, mint pl. max hely eloszlását.

* Kandidátusi értekezés, Budapest, 1974. május

Tárgyaljuk a véletlen mintaelemszám esetét, valamint bizonyos kétdimenziós problémákat.

Szeretnék köszönetet mondani RÉVÉSZ PÁLnak, TUSNÁDY GÁBORNak és VINCZE ISTVÁNNak probléma-felvetéseikért és értékes tanácsaikért.

1. §. VÉGES ELOSZLÁSTÉTELEK

1.1. A Ballot-lemma

Mint a Bevezetés-ben is említettük, az empirikus eloszlásfüggvény viselkedése általában nem csupán egy rögzített x pontban érdekes, hanem egy halmazon, rendszerint intervallumon vagy ezek egyesítésén. Érdekes, hogy erre vonatkozó első eredmények határértéktételek (KOLMOGOROV [64], GLIVENKO [54]) véges eloszlástételeket csak jóval később adtak meg. Az első ilyen eredmény SZMIRNOV nevéhez fűződik, aki megadta D_n^+ és D_n^- pontos eloszlását [92]. Később ő maga is, majd mások végeztek pontos eloszlásra vonatkozó vizsgálatokat. Ennek irodalma ma már olyan nagy, hogy a munkákat teljességében felsorolni lehetetlen, a fontosabb és érdekesebb cikkek, ill. könyvek közül megemlítiük a következőket: BARTON—MALLOWS [8], DARLING [37], GUPTA—PANCHAPAKESAN [55], HÁJEK—SIDÁK [56], KARLIN [61], SAHLER [82], SUZUKI [89], TAKÁCS [97], VINCZE [104], [105].

Disszertációnkban egy elemi, kombinatorikus módszert kívánunk bemutatni, amely az ún. ballot (szavazatszámálási) lemmán, ill. annak egy módosításán alapul. Ez a módszer lényegében TAKÁCS-tól származik [94].

A ballot lemma eredeti formája a következő (BERTRAND [10]): *Egy választás során, ahol 2 jelölt verseng egymással, a győztes végeredményben a , a vesztes b szavazatot kapott ($a > b$). Annak valószínűsége, hogy a győztes a szavazatok összeszámálása során végig előnyben volt, $\frac{a-b}{a+b}$.*

Könnyen látható, hogy TAKÁCS következő tétele [97] a ballot lemma általánosítása: *legyenek $k_1, k_2, \dots, k_n$ nem-negatív egész számok, melyekre $k_1 + k_2 + \dots + k_n = k \leq n$. Ekkor a $k_1, k_2, \dots, k_n$ számok $(x_1, x_2, \dots, x_n)$ ciklikus permutációi közül pontosan $n-k$ olyan van, melyre teljesülnek az alábbi egyenlőtlenségek:*

$$x_1 < 1, x_1 + x_2 < 2, \dots, x_1 + x_2 + \dots + x_n < n.$$

TAKÁCS nyomán ballot tételeknek fogjuk nevezni azokat a tételeket, melyek a fenti egyenlőtlenségek egyidejű teljesülésének valószínűségét adják meg, bár a feltételeket kissé módosítani fogjuk.

Az alábbi ballot tétel, melyre a továbbiakban szükségünk lesz, ekvivalens DANIELS egy tételével [36], melyet később mások is felfedeztek, ill. bizonyítottak. (Lásd pl. ROBBINS [80], CHAPMAN [23], RÉNYI [79], VINCZE [104], SARKADI [83]). Tekintettel arra, hogy ez a tétel képezi az e paragrafusban leírt vizsgálatok alapját, mi is adunk egy egyszerű bizonyítást (lásd CSÁKI—TUSNÁDY [32]).

1.1. TÉTEL. *Legyen $A_0, A_1, \dots, A_n$ teljes eseményrendszer, melyre $P(A_0) = p \geq 0$, $P(A_i) = q \geq 0$ ($i=1, 2, \dots, n$). Természetesen $p + nq = 1$. Végezzünk n független kísérletet, melynek során az A_j esemény v_j -szer következett be, ($j=0, 1, 2, \dots, n$).*

Ekkor

$$(1.1) \quad P(v_1 < 1, v_1 + v_2 < 2, \dots, v_1 + v_2 + \dots + v_n < n) = p.$$

Bizonyítás. A tételt teljes indukcióval bizonyítjuk. $n=1$ -re nyilván $P(v_1 < 1) = P(v_0 < 1) = p$. Tegyük fel, hogy az állítást $n < N$ -re, tetszőleges p, q -ra már belátuk. Ekkor

$$\begin{aligned} & P(v_1 < 1, \dots, v_1 + v_2 + \dots + v_N < N | v_0) = \\ & = P(v_1 < 1, \dots, v_1 + v_2 + \dots + v_{N-v_0} < N - v_0 | v_0) = \\ & = P(v_1 < 1, \dots, v_1 + v_2 + \dots + v_{N-v_0} < N - v_0 | v_1 + \dots + v_N = N - v_0) = \frac{v_0}{N}. \end{aligned}$$

Itt az utolsó lépés $v_0=0$ esetén nyilvánvaló, $v_0>0$ esetén pedig az indukciós feltevés miatt igaz. Így a teljes valószínűség tétele szerint

$$P(v_1 < 1, v_1 + v_2 < 2, \dots, v_1 + v_2 + \dots + v_N < N) = E\left(\frac{v_0}{N}\right) = p,$$

ami a tétel állítása.

1.2. Pontos eloszlástételek

A továbbiakban vizsgáljuk az $X_1, X_2, \dots, X_n$ valószínűségi változók empirikus eloszlásfüggvényét, azaz

$$(1.2) \quad F_n(x) = \frac{1}{n} \sum_{X_i \leq x} 1.$$

Feltételezzük, hogy az $X_1, X_2, \dots, X_n$ változók függetlenek és egyenletes eloszlásúak a $(0, 1)$ intervallumban. A folytonos eloszlások alkalmas transzformációval nyilván erre az esetre visszavezethetők, de a nem folytonos alapeloszlásokat is visszavezetjük a fenti esetre (lásd 1.4 pont, ill. NOETHER [72], SCHMID [85], VINCZE [104]).

Az alábbi tétel TAKÁCS [97], ill. EICKER [50] tételének speciális esete, de a formula az általános esethez képest lényegesen egyszerűsödik. $a=1$ -re lásd DEMPSTER [40] és DWASS [48].

1.2. TÉTEL. Legyen $0 < a \leq 1, 0 < b, 0 < c, 0 < ca - b < 1$. Ekkor

$$\begin{aligned} (1.3) \quad & P(F_n(x) > cx - b, 0 \leq x \leq a) = \\ & = 1 - \frac{b}{c} \sum_{j=0}^{[n(ca-b)]} \binom{n}{j} \left(\frac{b}{c} + \frac{j}{nc}\right)^{j-1} \left(1 - \frac{b}{c} - \frac{j}{nc}\right)^{n-j}. \end{aligned}$$

Bizonyítás. Az (1.3) formula annak az eseménynek a valószínűségét adja, hogy $F_n(x)$ az $y=cx-b$ egyenes felett halad, azaz az egyenest nem metszi a $(0, a]$ intervallum egyetlen pontjában sem. Határozzuk meg e helyett annak valószínűségét, hogy $F_n(x)$ metszi az egyenest valamilyen x pontban, ahol $0 < x \leq a$. Ez egy $\frac{j}{n}$

magasságon történhet, ahol $0 \leq j \leq [n(ca-b)]$. Jelölje tehát B_j azt az eseményt, hogy $F_n(x)$ átmetszi az $y=cx-b$ egyenest a $\frac{j}{n}$ magasságon, tehát létezik olyan x_j ($0 < x_j \leq a$), hogy $\frac{j}{n} = F_n(x_j) = cx_j - b$. Ez azt jelenti, hogy a $(0, x_j)$ intervallumban, ahol $x_j = \frac{b}{c} + \frac{j}{nc}$, pontosan j mintaelem van. Ennek valószínűsége nyilván

$$(1.4) \quad P(B_j) = \binom{n}{j} x_j^j (1-x_j)^{n-j}.$$

Mi a $B = \sum_{j=0}^{[n(ca-b)]} B_j$ esemény valószínűségét keressük. A B_j események természetesen nem diszjunktak, ezért

$$\begin{aligned} P(B) &= P\left(\sum_{j=0}^{[n(ca-b)]} B_j\right) = P\left(\sum_{j=0}^{[n(ca-b)]} \bar{B}_0 \bar{B}_1 \dots \bar{B}_{j-1} B_j\right) = \\ &= \sum_{j=0}^{[n(ca-b)]} P(\bar{B}_0 \bar{B}_1 \dots \bar{B}_{j-1} | B_j) P(B_j). \end{aligned}$$

A $P(B_j)$ valószínűséget az (1.4) formula adja, míg a $P(\bar{B}_0 \bar{B}_1 \dots \bar{B}_{j-1} | B_j)$ feltételes valószínűséget az 1.1. Tételből határozhatjuk meg. Jelölje ui. v_0 a $(0, x_0)$ intervallumba eső mintaelemek számát, v_i pedig az (x_{j-i}, x_{j-i+1}) intervallumba eső mintaelemek számát ($i=1, 2, \dots, j$). Ekkor a B_j esemény ekvivalens a $\{v_0 + v_1 + \dots + v_j = j\}$ eseménnyel, a $\bar{B}_0 \bar{B}_1 \dots \bar{B}_{j-1}$ esemény pedig a

$$v_1 < 1, v_1 + v_2 < 2, \dots, v_1 + v_2 + \dots + v_j < j$$

egyenlőtlenségek egyidejű teljesülésével. Így alkalmazható az 1.1. Tétel $n=j$, $p = \frac{x_0}{x_j}$ vel, tehát

$$(1.5) \quad P(\bar{B}_0 \bar{B}_1 \dots \bar{B}_{j-1} | B_j) = \frac{x_0}{x_j},$$

azaz

$$P(B) = x_0 \sum_{j=0}^{[n(ca-b)]} \binom{n}{j} x_j^{j-1} (1-x_j)^{n-j}.$$

Ezzel bizonyítottuk az 1.2. Tételt, hiszen az ott szereplő valószínűség a B esemény kiegészítő eseményének valószínűsége.

MEGJEGYZÉS. Az 1.2. Tételnél feltételeztük, hogy $ca-b < 1$. Ha ui. $ca-b \geq 1$, akkor $F_n(x)$ a $(0, a]$ intervallumon biztosan átmetszi a $cx-b$ egyenest, ha máshol nem, az 1 magasságon, ill. az x_n pontban, ahol $cx_n - b = 1$ és most $x_n \leq a$. Tekint-

sük annak valószínűségét, hogy csak itt történjék metszés. Ez kétféleképpen is meghatározható. Alkalmazzuk először az 1.2. Tétel módszerét. Ekkor az $1 - P\left(\sum_{j=0}^{n-1} B_j\right)$ valószínűséget kell meghatároznunk, így kapjuk

$$(1.6) \quad P\left(F_n(x) > cx - b, 0 \leq x < \frac{1+b}{c}\right) = \\ = 1 - \frac{b}{c} \sum_{j=0}^{n-1} \binom{n}{j} \left(\frac{b}{c} + \frac{j}{nc}\right)^{j-1} \left(1 - \frac{b}{c} - \frac{j}{nc}\right)^{n-j}.$$

Másrészt pedig itt a $P(\bar{B}_0 \bar{B}_1 \dots \bar{B}_{n-1} B_n) = P(\bar{B}_0 \bar{B}_1 \dots \bar{B}_{n-1} | B_n) P(B_n)$ valószínűségről van szó, tehát

$$(1.7) \quad P\left(F_n(x) > cx - b, 0 \leq x < \frac{1+b}{c}\right) = \frac{b}{c} \left(\frac{1+b}{c}\right)^{n-1} = \frac{b(1+b)^{n-1}}{c^n}.$$

A két valószínűség természetesen azonos, ami pl. az *Abel*-azonosságból is belátható.

Az Abel-féle

$$(1.8) \quad \sum_{j=0}^n \binom{n}{j} A(A+jB)^{j-1} (1-A-jB)^{n-j} = 1$$

azonosságot felhasználva bizonyos esetekben az 1.2. Tétel (1.3) formulája is egyszerűsíthető. Igaz ui., hogy

$$1 - \frac{b}{c} \sum_{j=0}^{[n(ca-b)]} \binom{n}{j} \left(\frac{b}{c} + \frac{j}{nc}\right)^{j-1} \left(1 - \frac{b}{c} - \frac{j}{nc}\right)^{n-j} = \\ = \frac{b}{c} \sum_{j=[n(ca-b)]+1}^n \binom{n}{j} \left(\frac{b}{c} + \frac{j}{nc}\right)^{j-1} \left(1 - \frac{b}{c} - \frac{j}{nc}\right)^{n-j}.$$

Ebből pl. $a=1$, $b=\frac{p}{1-p} \left(1 - \frac{1}{n}\right)$, $c=\frac{1-\frac{1}{n}}{1-p}$ -re megkaphatjuk SZMIRNOV következő formuláját (lásd [93], ill. RÉNYI [79])

$$(1.9) \quad P\left(F_n(x) > \frac{x-p}{1-p} \left(1 - \frac{1}{n}\right), 0 \leq x \leq 1\right) = p \left(1 + \frac{1-p}{n-1}\right)^{n-1}.$$

Ez a formula egyébként ekvivalens a következő ballot-lemmával:

Legyen $n \geq 2$ és $A_0, A_1, \dots, A_{n-1}$ egy teljes eseményrendszer, melyre

$$P(A_0) = p, \quad P(A_i) = q \quad (i = 1, 2, \dots, n-1).$$

Természetesen $p+(n-1)q=1$. Legyen v_i ($i=0, 1, 2, \dots, n-1$) az A_i esemény gyakorisága n független kísérlet során. Ekkor

$$(1.10) \quad P(v_1 \leq 1, v_1 + v_2 \leq 2, \dots, v_1 + v_2 + \dots + v_{n-1} \leq n-1) = p(1+q)^{n-1}.$$

A továbbiakban vizsgáljunk egy $G(x)$, a $[0, 1]$ intervallumban értelmezett nem-csökkenő függvényt, mely szakaszonként lineáris és $G(0) \leq 0, G(1) \leq 1$. Az a feltevés, hogy $G(x)$ szakaszonként lineáris, tulajdonképpen nem jelenti az általánosság megszorítását, tetszőleges más függvény is visszavezethető erre az esetre. Általában azonban a szakaszok száma függhet a mintaelemszámtól. Az érdekes eset nyilván az, amikor a szakaszok száma nem nagy, és nem függ a mintaelemszámtól. Ezeket a későbbiekben részletesen tárgyaljuk.

Tekintsük a $P(F_n(x) > G(x), 0 \leq x \leq 1)$ valószínűséget. A fenti valószínűség egyéb előállításaira (pl. n -változós integrál, rekurziós formula, determináns) a következő dolgozatokra utalunk: WALD—WOLFOWITZ [106], DANIELS [36], WHITTLE [108], SUZUKI [89], DURBIN [44], NOÉ—VANDEWIELE [71], STECK [87], SARKADI [84]. Az 1.3. Tételben erre egy annyi változós összeget adunk meg, ahány szakaszból $G(x)$ áll. Bizonyos esetekben 2, vagy 3 szakasz össze is vonható (mint pl. az 1.2. Tételnél, és így az összegben szereplő változók száma csökken).

Legyen tehát $G(x)$ adva a következő formulával:

$$(1.11) \quad G(x) = c_i x + b_i, \quad a_{i-1} < x \leq a_i, \quad i = 1, 2, \dots, k,$$

ahol

$$a_0 = 0 < a_1 < a_2 < \dots < a_{k-1} < a_k = 1,$$

$$c_i \geq 0, \quad c_i a_i + b_i \leq c_{i+1} a_i + b_{i+1}, \quad c_k + b_k \leq 1,$$

valamint vagy $b_1 < 0$, vagy $b_1 = 0, c_1 = 0$.

1.3. TÉTEL. *A fenti feltétel mellett*

$$(1.12) \quad P(F_n(x) \geq G(x), 0 \leq x \leq 1) = \\ = n! \sum_{(\alpha_1, \dots, \alpha_k)} \dots \sum_{i=1}^k \frac{(a_i - a_{i-1})^{\alpha_i}}{\alpha_i!} p_n \left(c_i (a_i - a_{i-1}), \frac{\alpha_1 + \dots + \alpha_{i-1}}{n} - b_i - c_i a_{i-1}; \alpha_i \right),$$

ahol

$$p_n(u, z; \alpha) = \begin{cases} 1 - \frac{z}{u^\alpha} \sum_{j=0}^{[\alpha(u-z)]} \binom{\alpha}{j} \left(z + \frac{j}{n} \right)^{j-1} \left(u - z - \frac{j}{n} \right)^{\alpha-j}, & u \geq z \\ 1, & u < z \end{cases}$$

és az összegzés azon $\alpha_1, \alpha_2, \dots, \alpha_k$ nem-negatív egész számokra terjed ki, melyekre

$$\alpha_1 + \alpha_2 + \dots + \alpha_k = n,$$

$$\alpha_1 + \alpha_2 + \dots + \alpha_i \geq n(c_{i+1} a_i + b_{i+1}), \quad i = 1, 2, \dots, k-1.$$

Bizonyítás. Legyen v_i az (a_{i-1}, a_i) intervallumba eső mintaelemek száma. A $\{v_1 = \alpha_1, \dots, v_k = \alpha_k\}$ esemény valószínűsége a polinomiális eloszlásból adódik, ezen feltétel mellett az $\{F_n(x) \geq G(x), a_{i-1} \leq x \leq a_i\}$ események különböző i -re nyilván függetlenek. Tekintsük tehát az $\{F_n(x) \geq G(x), a_{i-1} \leq x \leq a_i\}$, azaz az

$\{F_n(x) \geq c_i x + b_i, a_{i-1} \leq x \leq a_i\}$ eseményt a $v_1 = \alpha_1, v_2 = \alpha_2, \dots, v_k = \alpha_k$ feltétel mellett. Ekkor $F_n(a_{i-1}) = \frac{\alpha_1 + \dots + \alpha_{i-1}}{n}$ és $F_n(a_i) = \frac{\alpha_1 + \dots + \alpha_i}{n}$. Alkalmazzuk az 1.2. Tétel bizonyításánál kifejtett gondolatmenetet. Ekkor annak valószínűsége, hogy $F_n(x)$ az $y = c_i x + b_i$ egyenest az $\frac{\alpha_1 + \dots + \alpha_{i-1} + j}{n}$ magasságon messe, $\binom{\alpha_i}{j} x_j^j (1-x_j)^{\alpha_i-j}$, ahol most $x_j = \frac{\alpha_1 + \dots + \alpha_{i-1} + j}{nc_i(a_i - a_{i-1})} - \frac{b_i + a_{i-1}c_i}{c_i(a_i - a_{i-1})}$. Annak valószínűsége pedig, hogy ez legyen az első metszés, $\frac{x_0}{x_j}$. Így annak valószínűsége, hogy $F_n(x)$ az egyenest ne messe, a tételben szereplő p_n . Az α -kra vonatkozó egyenlőtlenség biztosítja, hogy $F_n(x)$ az $a_1, a_2, \dots, a_k$ pontokban a $G(x)$, szakaszonként lineáris függvény felett legyen. Ezzel az 1.3. Tételt bizonyítottuk.

Az 1.3. Tétel több fontos statisztika eloszlását speciális esetként tartalmazza. Ezeket kívánjuk részletezni a továbbiakban. Mint jeleztük, bizonyos esetekben 2 vagy 3 szakasz össze is vonható, ami által a formula egyszerűsödik. Ezeket külön tételként mondjuk ki.

Az alábbi, 1.4. Tétel a $P(F_n(x) > cx + b, a_1 \leq x \leq a_2)$ valószínűséget adja meg. Ezzel ekvivalens formulákat ad meg EICKER [50], de véleményünk szerint az alábbi tétel mind módszerében, mind formulájában egyszerűbb, a speciális esetre vonatkozó legegyszerűbb formulák azonnal kiadódnak.

1.4. TÉTEL. Legyen $0 \leq a_1 < a_2 \leq 1, c > 0, ca_2 + b \leq 1$. Ekkor

$$(1.13) \quad P(F_n(x) \geq cx + b, a_1 \leq x \leq a_2) =$$

$$= \begin{cases} \sum_{\alpha > n(ca_1 + b)} \binom{n}{\alpha} a_1^\alpha (1-a_1)^{n-\alpha} - \\ - \sum_{\alpha > n(ca_1 + b)} (x_0 - a_1)^{[n(ca_2 + b)] - \alpha} \sum_{j=0}^{[n(ca_2 + b)] - \alpha} \frac{n!}{\alpha! j! (n - \alpha - j)!} a_1^\alpha (x_j - a_1)^{j-1} (1-x_j)^{n-\alpha-j}, & \text{ha} \\ 1 - z_0 \sum_{j=0}^{[n(ca_2 + b)]} \binom{n}{j} z_j^{j-1} (1-z_j)^{n-j}, & \text{ha } ca_1 + b > 0, \end{cases}$$

$$x_j = \frac{\alpha + j}{nc} - \frac{b}{c}, \quad z_j = \frac{j}{nc} - \frac{b}{c}.$$

Bizonyítás. Ha $ca_1 + b > 0$, legyen v a $(0, a_1)$ intervallumban levő mintaelemek száma. A $\{v = \alpha\}$ esemény valószínűsége nyilván $\binom{n}{\alpha} a_1^\alpha (1-a_1)^{n-\alpha}$, ezen feltétel mellett az $\{F_n(x) > cx + b, a_1 \leq x \leq a_2\}$ esemény valószínűségét az 1.2. Tétel módszerével, az 1.3. Tételnél eszközölt módosítást alkalmazva határozhatjuk meg. Eszerint

$$P(F_n(x) \geq cx + b, a_1 \leq x \leq a_2 | v = \alpha) =$$

$$= 1 - \sum_{(j)} \binom{n-\alpha}{j} \left(\frac{x_j - a_1}{1 - a_1} \right)^{j-1} \left(\frac{1 - x_j}{1 - a_1} \right)^{n-\alpha-j} \frac{x_0 - a_1}{1 - a_1},$$

feltéve, hogy $\alpha > n(ca_1 + b)$. így

$$\begin{aligned} P(F_n(x) \cong cx + b, a_1 \leq x \leq a_2) &= \\ &= \sum_{\alpha > n(ca_1 + b)} P(F_n(x) \cong cx + b, a_1 \leq x \leq a_2 | v = \alpha) P(v = \alpha) = \\ &= \sum_{(\alpha)} \binom{n}{\alpha} a_1^\alpha (1 - a_1)^{n-\alpha} \left(1 - \sum_{(j)} \binom{n-\alpha}{j} \left(\frac{x_j - a_1}{1 - a_1} \right)^{j-1} \left(\frac{1 - x_j}{1 - a_1} \right)^{n-\alpha-j} \frac{x_0 - a_1}{1 - a_1} \right), \end{aligned}$$

ami adja az állítást, ha $ca_1 + b > 0$. (Megjegyezzük, hogy a formula $a_1 = 0, b > 0$ esetén 0-t ad, aminek így is kell lenni.)

Ha $ca_1 + b \leq 0$, akkor az állítás közvetlenül az 1.2. Tételből adódik. Ekkor ui.

$$P(F_n(x) \cong cx + b, a_1 \leq x \leq a_2) = P(F_n(x) \cong cx + b, 0 \leq x \leq a_2).$$

Az 1.4. Tétel következményeként megadjuk néhány (egyoldali) statisztika pontos eloszlását. Ezek egyike-másika természetesen nem új, az irodalomban már megtalálható, ismert formula. Ezeknél megjelöljük azt a forrást, ahol az illető formula tudomásunk szerint először meg van adva (helyesen). Megjegyezzük, hogy BIRNBAUM és LIENTZ [12] egy formulagyűjteményt bizonyítással vélnek megadni, de csaknem minden eredményük — egy-két kivétellel — hibás. Az alábbiakban elvégezzük e formulagyűjtemény korrekcióját is.

Előbb kimondunk még egy lemmát, amely lehetővé teszi, hogy $\{F_n(x) \leq \dots\}$ események valószínűségét visszavezessük $\{F_n(x) \geq \dots\}$ események valószínűségére.

1.1. LEMMA. Legyen $G(x)$ a $[0, 1]$ intervallumban értelmezett függvény. Ekkor

$$P(F_n(x) \geq G(x), 0 \leq x \leq 1) = P(F_n(x) \leq 1 - G(1 - x), 0 \leq x \leq 1).$$

Bizonyítás. Végezzük el az $X_i^{(1)} = 1 - X_i$ transzformációt. Ekkor $(X_1^{(1)}, X_2^{(2)}, \dots, X_n^{(1)})$ ugyancsak a $(0, 1)$ -ben egyenletes eloszlásból vett véletlen minta és az $\{F_n(x) \geq G(x)\}$ esemény minden x -re ekvivalens az $\{F_n^{(1)}(x) \leq 1 - G(1 - x)\}$ eseménnyel, ahol $F_n^{(1)}(x)$ az $X_1^{(1)}, \dots, X_n^{(1)}$ mintaelemek empirikus eloszlásfüggvénye.

1.4.1. KOROLLÁRIUM. Tekintsük a $\sup_{a_1 \leq x \leq a_2} (x - F_n(x))$, ill. a $\sup_{1 - a_2 \leq x \leq 1 - a_1} (F_n(x) - x)$ statisztikákat, melyeknek az 1.1. Lemma értelmében azonos az eloszlásuk. Ekkor

$$(1.14) \quad P\left(\sup_{a_1 \leq x \leq a_2} (x - F_n(x)) < \varepsilon\right) = P\left(\sup_{1 - a_2 \leq x \leq 1 - a_1} (F_n(x) - x) < \varepsilon\right) =$$

$$= \begin{cases} \sum_{\alpha > n(a_1 - \varepsilon)} \binom{n}{\alpha} a_1^\alpha (1 - a_1)^{n-\alpha} - \\ - \sum_{\alpha > n(a_1 - \varepsilon)} \binom{\alpha + \varepsilon - a_1}{n} \sum_{j=0}^{[n(a_2 - \varepsilon)] - \alpha} \frac{n!}{\alpha! j! (n - \alpha - j)!} a_1^\alpha \left(\frac{\alpha + j + \varepsilon - a_1}{n} \right)^{j-1} \left(1 - \frac{\alpha + j}{n} - \varepsilon \right)^{n-\alpha-j}, \\ 1 - \varepsilon - \sum_{j=0}^{[n(a_2 - \varepsilon)]} \binom{n}{j} \left(\frac{j}{n} + \varepsilon \right)^{j-1} \left(1 - \frac{j}{n} - \varepsilon \right)^{n-j}, \quad a_1 \leq \varepsilon < 1. \end{cases} \quad 0 < \varepsilon < a_1.$$

Bizonyítás.

$$\begin{aligned} P\left(\sup_{a_1 \leq x \leq a_2} (x - F_n(x)) < \varepsilon\right) &= P(x - F_n(x) < \varepsilon, a_1 \leq x \leq a_2) = \\ &= P(F_n(x) > x - \varepsilon, a_1 \leq x \leq a_2), \end{aligned}$$

így alkalmazható az 1.4. Tétel $b = -\varepsilon, c = 1$ -gyel.

Az $a_1 = 0, a_2 = 1$ eset felel meg az egyoldali Kolmogorov—Szmirnov statisztika eloszlásának. Tudomásunk szerint ez volt az első pontos eloszlás, melyet meghatároztak, mégpedig SZMIRNOV [92]. Később BIRNBAUM és TINGEY [15] is megadták ugyanezt a formulát, néhol — tévesen — a formula meghatározását nekik tulajdonítják. Az $a_1 = 0, a_2$ tetszőleges, ill. a_1 tetszőleges, $a_2 = 1$ esetet G. ISHII [58], majd később SZMIRNOV [93] vizsgálta. Lásd még CSÖRGŐ [33]. Az általános esetet TAKÁCS [94] adja meg, nem a fenti formában. BIRNBAUM és LIENTZ formulái [12] hibásak.

1.4.2. KOROLLÁRIUM. Tekintsük a

$$\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{1 - x} \right), \quad \text{ill. } a \sup_{1 - a_2 \leq x \leq 1 - a_1} \left(\frac{F_n(x) - x}{x} \right)$$

statisztikákat, melyeknek az 1.1. Lemma értelmében azonos az eloszlásuk. Ekkor

$$\begin{aligned} (1.15) \quad P\left(\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{1 - x} \right) < \varepsilon\right) &= P\left(\sup_{1 - a_2 \leq x \leq 1 - a_1} \left(\frac{F_n(x) - x}{x} \right) < \varepsilon\right) = \\ &= \begin{cases} \sum_{\alpha > n(1+\varepsilon)a_1 - \varepsilon} \binom{n}{\alpha} a_1^\alpha (1 - a_1)^{n-\alpha} - \sum_{\alpha > n(1+\varepsilon)a_1 - \varepsilon} \left(\frac{\alpha}{n(1+\varepsilon)} + \frac{\varepsilon}{1+\varepsilon} - a_1 \right) \times \\ \times \sum_{j=0}^{[n((1+\varepsilon)a_2 - \varepsilon)] - \alpha} \frac{n! a_1^\alpha}{\alpha! j! (n - \alpha - j)!} \left(\frac{\alpha + j}{n(1+\varepsilon)} + \frac{\varepsilon}{1+\varepsilon} - a_1 \right)^{j-1} \left(1 - \frac{\alpha + j}{n(1+\varepsilon)} - \frac{\varepsilon}{1+\varepsilon} \right)^{n-\alpha-j}, \\ 0 < \varepsilon < \frac{a_1}{1 - a_1}, \\ 1 - \frac{\varepsilon}{1+\varepsilon} \sum_{j=0}^{[n((1+\varepsilon)a_2 - \varepsilon)]} \binom{n}{j} \left(\frac{j}{n(1+\varepsilon)} + \frac{\varepsilon}{1+\varepsilon} \right)^{j-1} \left(1 - \frac{j}{n(1+\varepsilon)} - \frac{\varepsilon}{1+\varepsilon} \right)^{n-j}, & \frac{a_1}{1 - a_1} \leq \varepsilon. \end{cases} \end{aligned}$$

$a_2 = 1$ esetén a formula tovább egyszerűsíthető:

$$\begin{aligned} (1.16) \quad P\left(\sup_{a_1 \leq x \leq 1} \left(\frac{x - F_n(x)}{1 - x} \right) < \varepsilon\right) &= \\ &= \begin{cases} \sum_{\alpha > n(1+\varepsilon)a_1 - \varepsilon} \binom{n}{\alpha} a_1^\alpha (1 - a_1)^{n-\alpha-1} \left(\frac{\alpha}{n(1+\varepsilon)} + \frac{\varepsilon}{1+\varepsilon} - a_1 \right), & 0 < \varepsilon < \frac{a_1}{1 - a_1} \\ \frac{\varepsilon}{1+\varepsilon}, & \frac{a_1}{1 - a_1} \leq \varepsilon. \end{cases} \end{aligned}$$

Bizonyítás.

$$P\left(\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{1-x}\right) < \varepsilon\right) = P\left(\frac{x - F_n(x)}{1-x} < \varepsilon, a_1 \leq x \leq a_2\right) = \\ = P(F_n(x) > (1+\varepsilon)x - \varepsilon, a_1 \leq x \leq a_2),$$

így alkalmazható az 1.4. Tétel $b = -\varepsilon, c = 1 + \varepsilon$ -nal. $a_2 = 1$ esetén a $\{v = \alpha\}$ feltétel mellett, ill. $\frac{a_1}{1-a_1} \leq \varepsilon$ esetén közvetlenül az 1.1. Tétel alkalmazható.

Az $a_1 = 0, a_2 = 1$ esetben az állítás ekvivalens az 1.1. Tétellel, tehát DANIELSTÖL származik [36]. Az $a_1 = 0, a_2$ tetszőleges esetet G. ISHII [59], a_1 tetszőleges, $a_2 = 1$ esetet CHANG [20] vizsgálta. Mind a két eset megtalálható SZMIRNOVNÁL [93] is. Lásd még CSÖRGŐ [34]. Az általános esetet TAKÁCS [94] adja, de nem a fenti formában. BIRNBAUM és LIENTZ formulái [12] hibásak.

1.4.3. KOROLLÁRIUM. *Tekintsük a*

$$\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{x}\right), \quad \text{ill. } a \quad \sup_{1-a_2 \leq x \leq 1-a_1} \left(\frac{F_n(x) - x}{1-x}\right)$$

statisztikákat, melyeknek az 1.1. Lemma értelmében azonos az eloszlásuk. Ekkor $a_1 > 0$ esetén

$$(1.17) \quad P\left(\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{x}\right) < \varepsilon\right) = P\left(\sup_{1-a_2 \leq x \leq 1-a_1} \left(\frac{F_n(x) - x}{x}\right) < \varepsilon\right) = \\ = \sum_{\alpha > n(1-\varepsilon)a_1} \binom{n}{\alpha} a_1^\alpha (1-a_1)^{n-\alpha} - \\ - \sum_{\alpha > n(1-\varepsilon)a_1} \left(\frac{\alpha}{n(1-\varepsilon)} - a_1\right)^{[n(1-\varepsilon)a_2] - \alpha} \frac{n! a_1^\alpha}{\alpha! j! (n-\alpha-j)!} \times \\ \times \left(\frac{\alpha+j}{n(1-\varepsilon)} - a_1\right)^{j-1} \left(1 - \frac{\alpha+j}{n(1-\varepsilon)}\right)^{n-\alpha-j}, \quad \text{ha } \frac{a_2-1}{a_2} < \varepsilon < 1.$$

Bizonyítás.

$$P\left(\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{x}\right) < \varepsilon\right) = P\left(\frac{x - F_n(x)}{x} < \varepsilon, a_1 \leq x \leq a_2\right) = \\ = P(F_n(x) > (1-\varepsilon)x, a_1 \leq x \leq a_2),$$

így alkalmazható az 1.4. Tétel $b=0, c=1-\varepsilon$ -nal. Itt az $a_1=0$ eset értelmetlen, hiszen 1 valószínűséggel fennáll $\sup_{0 \leq x \leq a_2} \left(\frac{x - F_n(x)}{x}\right) = 1$, azaz konstans.

BIRNBAUM és LIENTZ formulái [12] hibásak.

Tudomásunk szerint az (1.17) formula explicite még speciális esetben sem található meg az irodalomban.

1.4.4. KOROLLÁRIUM. Tekintsük a

$$\sup_{A_1 \leq F_n(x) \leq A_2} (x - F_n(x)), \quad \text{ill. a} \quad \sup_{1-A_2 \leq F_n(x) \leq 1-A_1} (F_n(x) - x)$$

statisztikákat, melyek eloszlása az 1.1. Lemma értelmében azonos. Ekkor

$$(1.18) \quad P\left(\sup_{A_1 \leq F_n(x) \leq A} (x - F_n(x)) < \varepsilon\right) = P\left(\sup_{1-A_2 \leq F_n(x) \leq 1-A_1} (F_n(x) - x) < \varepsilon\right) = 1, \\ \text{ha} \quad \varepsilon \geq 1 - A_1$$

$$\left\{ \sum_{\alpha > nA_1} \binom{n}{\alpha} (A_1 + \varepsilon)^\alpha (1 - A_1 - \varepsilon)^{n-\alpha} - \sum_{\alpha > nA_1} \left(\frac{\alpha}{n} - A_1\right) \sum_{j=0}^{[nA_2]-\alpha} \frac{n! (A_1 + \varepsilon)^\alpha}{\alpha! j! (n - \alpha - j)!} \left(\frac{\alpha + j}{n} - A_1\right)^{j-1} \left(1 - \frac{\alpha + j}{n} - \varepsilon\right)^{n-\alpha-j} \right\}, \\ \text{ha} \quad -A_1 \leq \varepsilon < 1 - A_1.$$

Bizonyítás. Könnyen látható, hogy

$$P\left(\sup_{A_1 \leq F_n(x) \leq A_2} (x - F_n(x)) < \varepsilon\right) = P\left(\sup_{a_1 \leq x \leq a_2} (x - F_n(x)) < \varepsilon\right),$$

ahol $a_1 = \min(A_1 + \varepsilon, 1)$ és $a_2 = \min(A_2 + \varepsilon, 1)$, így a fenti formulát az 1.4.1. Korolláriumból kaphatjuk. $A_1 = 0$ -ra a formula a következőképpen egyszerűsíthető:

$$P\left(\sup_{0 \leq F_n(x) \leq A_2} (x - F_n(x)) < \varepsilon\right) = 1 - \varepsilon \sum_{j=0}^{[nA_2]} \binom{n}{j} \left(\frac{j}{n} + \varepsilon\right)^{j-1} \left(1 - \frac{j}{n} - \varepsilon\right)^{n-j}.$$

$A_1 = 0$ -ra, ill. $A_2 = 1$ -re CSÖRGŐ [33] határozta meg az eloszlást. $A_1 = 0$ -ra BIRNBAUM és LIENTZ formulája [12] korrekt, míg az általános eset tudomásunk szerint az irodalomban nem található.

1.4.5. KOROLLÁRIUM. Tekintsük a

$$\sup_{A_1 \leq F_n(x) \leq A_2} \left(\frac{x - F_n(x)}{1 - F_n(x)}\right), \quad \text{ill. a} \quad \sup_{1-A_2 \leq F_n(x) \leq 1-A_1} \left(\frac{F_n(x) - x}{F_n(x)}\right)$$

statisztikákat, melyeknek eloszlása az 1.1. Lemma értelmében azonos. Ekkor

$$(1.19) \quad P\left(\sup_{A_1 \leq F_n(x) \leq A_2} \left(\frac{x - F_n(x)}{1 - F_n(x)}\right) < \varepsilon\right) = P\left(\sup_{1-A_2 \leq F_n(x) \leq 1-A_1} \left(\frac{F_n(x) - x}{F_n(x)}\right) < \varepsilon\right) = \\ = \sum_{\alpha > nA_1} \binom{n}{\alpha} (A_1(1 - \varepsilon) + \varepsilon)^\alpha (1 - \varepsilon)^{n-\alpha} (1 - A_1)^{n-\alpha} - \\ - \sum_{\alpha > nA_1} \left(\frac{\alpha}{n} - A_1\right) \sum_{j=0}^{[nA_2]-\alpha} \frac{n! (A_1(1 - \varepsilon) + \varepsilon)^\alpha}{\alpha! j! (n - \alpha - j)!} (1 - \varepsilon)^{n-\alpha} \left(\frac{\alpha + j}{n} - A_1\right)^{j-1} \left(1 - \frac{\alpha + j}{n}\right)^{n-\alpha-j}, \\ - \frac{A_1}{1 - A_1} < \varepsilon < 1.$$

Ha $A_1=0$, akkor

$$P\left(\sup_{0 \leq F_n(x) \leq A_2} \left(\frac{x - F_n(x)}{1 - F_n(x)}\right) < \varepsilon\right) = \\ = 1 - \varepsilon \sum_{j=0}^{[nA_2]} \binom{n}{j} \left(\frac{j}{n}(1-\varepsilon) + \varepsilon\right)^{j-1} (1-\varepsilon)^{n-j} \left(1 - \frac{j}{n}\right)^{n-j}, \quad 0 < \varepsilon < 1.$$

Ha $A_2=1$, akkor

$$P\left(\sup_{A_1 \leq F_n(x) \leq 1} \left(\frac{x - F_n(x)}{1 - F_n(x)}\right) < \varepsilon\right) = \\ = \sum_{\alpha > nA_1} \binom{n}{\alpha} (A_1(1-\varepsilon) + \varepsilon)^\alpha (1-\varepsilon)^{n-\alpha} (1-A_1)^{n-\alpha-1} \left(\frac{\alpha}{n} - A_1\right), \quad -\frac{A_1}{1-A_1} < \varepsilon < 1.$$

Ha $A_1=0$, $A_2=1$, akkor

$$P\left(\sup_{0 \leq F_n(x) \leq 1} \left(\frac{x - F_n(x)}{1 - F_n(x)}\right) < \varepsilon\right) = \varepsilon, \quad 0 < \varepsilon < 1,$$

azaz az eloszlás egyenletes.

Bizonyítás. Könnyen látható, hogy

$$P\left(\sup_{A_1 \leq F_n(x) \leq A_2} \left(\frac{x - F_n(x)}{1 - F_n(x)}\right) < \varepsilon\right) = P\left(\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{1 - x}\right) < \frac{\varepsilon}{1 - \varepsilon}\right),$$

ahol $a_1 = A_1(1-\varepsilon) + \varepsilon$ és $a_2 = A_2(1-\varepsilon) + \varepsilon$. Így a formulát az 1.4.2. Korolláriumból kaphatjuk.

Az $A_1=0$ eset Csörgőtől származik [34]. BIRNBAUM és LIENTZ formulája [12] hibás. Az általános eset tudomásunk szerint az irodalomban nem található.

1.4.6. KOROLLÁRIUM. Tekintsük a

$$\sup_{A_1 \leq F_n(x) \leq A_2} \left(\frac{x - F_n(x)}{F_n(x)}\right), \quad \text{ill. } a \quad \sup_{1-A_2 \leq F_n(x) \leq 1-A_1} \left(\frac{F_n(x) - x}{1 - F_n(x)}\right)$$

statisztikákat, melyeknek eloszlása az 1.1. Lemma értelmében azonos. Ekkor $A_1 > 0$ -ra

$$(1.20) \quad P\left(\sup_{A_1 \leq F_n(x) \leq A_2} \left(\frac{x - F_n(x)}{F_n(x)}\right) < \varepsilon\right) = P\left(\sup_{1-A_2 \leq F_n(x) \leq 1-A_1} \left(\frac{F_n(x) - x}{1 - F_n(x)}\right) < \varepsilon\right) = \\ = \sum_{\alpha > nA_1} \binom{n}{\alpha} ((1+\varepsilon)A_1)^\alpha (1 - (1+\varepsilon)A_1)^{n-\alpha} - \\ - \sum_{\alpha > nA_1} \left(\frac{\alpha}{n} - A_1\right) \sum_{j=0}^{[nA_2]-\alpha} \frac{n! A_1^\alpha (1+\varepsilon)^{\alpha+j}}{\alpha! j! (n-\alpha-j)!} \left(\frac{\alpha+j}{n} - A_1\right)^{j-1} \left(1 - \frac{\alpha+j}{n} (1+\varepsilon)\right)^{n-\alpha-j}, \\ -1 < \varepsilon < \frac{1-A_1}{A_1}.$$

Bizonyítás. Könnyen látható, hogy

$$P\left(\sup_{A_1 \leq F_n(x) \leq A_2} \left(\frac{x - F_n(x)}{F_n(x)}\right) < \varepsilon\right) = P\left(\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{x}\right) < \frac{\varepsilon}{1 + \varepsilon}\right),$$

ahol $a_1 = (1 + \varepsilon)A_1$ és $a_2 = (1 + \varepsilon)A_2$, így a formulát az 1.4.3. Korolláriumból kaphatjuk.

További, hasonló statisztikák eloszlása, mint pl.

$$\sup_{a_1 \leq x \leq a_2} (F_n(x) - \gamma x), \quad \text{ill.} \quad \sup_{a_1 \leq x \leq a_2} \left(\frac{F_n(x) - \gamma x}{x}\right)$$

(TAKÁCS [94]) stb. is meghatározható a fenti módszerrel, ezek részletezéséről azonban eltekintünk.

A következőkben két lineáris szakaszból álló $G(x)$ esetével foglalkozunk.

1.5. TÉTEL. Legyen $0 < a_1 \leq a_2 < 1$, $c_1 > 0$, $c_2 > 0$, $b_1 < 0$, $c_1 a_1 + b_1 > 0$, $c_2 a_2 + b_2 \leq c_1 a_1 + b_1$, $c_2 + b_2 \leq 1$.
Ekkor

$$\begin{aligned} (1.21) \quad & P(F_n(x) > c_1 x + b_1, \quad 0 \leq x \leq a_1; \quad F_n(x) > c_2 x + b_2, \quad a_2 \leq x < 1) = \\ & = \sum_{\alpha > n(c_2 a_2 + b_2)} \binom{n}{\alpha} a_2^\alpha (1 - a_2)^{n - \alpha} \times \\ & \times \left(1 + \frac{b_1}{a_2 c_1} \sum_{j=0}^{[n(c_1 a_1 + b_1)]} \binom{\alpha}{j} \left(\frac{j - b_1}{a_2 c_1} \right)^{j-1} \left(1 - \frac{j - b_1}{a_2 c_1} \right)^{\alpha - j} \right) \times \\ & \times \left(1 - \frac{\frac{\alpha}{nc_2} - \frac{b_2}{c_2} - a_2}{1 - a_2} \sum_{i=0}^{[n(c_2 + b_2)] - \alpha} \binom{n - \alpha}{i} \left(\frac{\frac{\alpha + i}{nc_2} - \frac{b_2}{c_2} - a_2}{1 - a_2} \right)^{i-1} \left(1 - \frac{\frac{\alpha + i}{nc_2} - \frac{b_2}{c_2}}{1 - a_2} \right)^{n - \alpha - i} \right). \end{aligned}$$

Bizonyítás. Legyen v a $(0, a_2)$ intervallumba eső mintaelemek száma. A $\{v = \alpha\}$ esemény valószínűsége $\binom{n}{\alpha} a_2^\alpha (1 - a_2)^{n - \alpha}$, ezen feltétel mellett $\alpha > n(c_2 a_2 + b_2)$ esetén az $\{F_n(x) > c_1 x + b_1, 0 \leq x \leq a_1\}$ és $\{F_n(x) > c_2 x + b_2, a_2 \leq x < 1\}$ események függetlenek. A valószínűségeket éppen úgy kaphatjuk mint az 1.4. Tételben, amivel az 1.5. Tételt bizonyítottuk.

Az 1.5. Tétel folyományaként megkaphatjuk néhány újabb statisztika eloszlását.

1.5.1. KOROLLÁRIUM. Vizsgáljuk a

$$\sup_{\substack{x \leq a_1 \\ a_2 \leq x}} (x - F_n(x)), \quad \text{ill.} \quad a \quad \sup_{\substack{x \leq 1 - a_2 \\ 1 - a_1 \leq x}} (F_n(x) - x)$$

statisztikákat, melyeknek eloszlása az 1.1. Lemma értelmében azonos. Ekkor

$$\begin{aligned}
 (1.22) \quad P\left(\sup_{\substack{x \leq a_1 \\ a_2 \leq x}} (x - F_n(x)) < \varepsilon\right) &= P\left(\sup_{\substack{x \leq 1-a_2 \\ 1-a_1 \leq x}} (F_n(x) - x) < \varepsilon\right) = \\
 &= \sum_{\alpha > n(a_2 - \varepsilon)} \binom{n}{\alpha} a_2^\alpha (1-a_2)^{n-\alpha} \times \\
 &\quad \times \left(1 - \frac{\varepsilon}{a_2} \sum_{j=0}^{[n(a_1 - \varepsilon)]} \binom{\alpha}{j} \left(\left(\frac{j}{n} + \varepsilon\right) \frac{1}{a_2}\right)^{j-1} \left(1 - \left(\frac{j}{n} + \varepsilon\right) \frac{1}{a_2}\right)^{\alpha-j}\right) \times \\
 &\quad \times \left(1 - \frac{\frac{\alpha}{n} + \varepsilon - a_2}{1-a_2} \sum_{i=0}^{[n(1-\varepsilon)]-\alpha} \binom{n-\alpha}{i} \left(\frac{\frac{\alpha+i}{n} + \varepsilon - a_2}{1-a_2}\right)^{i-1} \left(1 - \frac{\frac{\alpha+i}{n} + \varepsilon - a_2}{1-a_2}\right)^{n-\alpha-i}\right).
 \end{aligned}$$

Bizonyítás. Alkalmazzuk az 1.5. Tételt $c_1 = c_2 = 1, b_1 = b_2 = -\varepsilon$ -nal.

1.5.2. KOROLLÁRIUM. Tekintsük a

$$\sup_{\substack{x \leq a_1 \\ a_2 \leq x}} \left(\frac{x - F_n(x)}{1-x}\right), \quad \text{ill. } a \quad \sup_{\substack{x \leq 1-a_2 \\ 1-a_1 \leq x}} \left(\frac{F_n(x) - x}{x}\right)$$

statisztikákat, melyeknek eloszlása az 1.1. Lemma értelmében azonos. Ekkor

$$\begin{aligned}
 (1.23) \quad P\left(\sup_{\substack{x \leq a_1 \\ a_2 \leq x}} \left(\frac{x - F_n(x)}{1-x}\right) < \varepsilon\right) &= P\left(\sup_{\substack{x \leq 1-a_2 \\ 1-a_1 \leq x}} \left(\frac{F_n(x) - x}{x}\right) < \varepsilon\right) = \\
 &= \sum_{\alpha > n((1+\varepsilon)a_2 - \varepsilon)} \binom{n}{\alpha} a_2^\alpha (1-a_2)^{n-\alpha-1} \left(\frac{\alpha}{n(1+\varepsilon)} + \frac{\varepsilon}{1+\varepsilon} - a_2\right) \times \\
 &\quad \times \left(1 - \frac{\varepsilon}{a_2(1+\varepsilon)} \sum_{j=0}^{[n(a_1(1+\varepsilon) - \varepsilon)]} \binom{\alpha}{j} \left(\frac{\frac{j}{n} + \varepsilon}{a_2(1+\varepsilon)}\right)^{j-1} \left(1 - \frac{\frac{j}{n} + \varepsilon}{a_2(1+\varepsilon)}\right)^{\alpha-j}\right).
 \end{aligned}$$

Bizonyítás. Alkalmazzuk az 1.5. Tételt $c_1 = c_2 = 1 + \varepsilon, b_1 = b_2 = -\varepsilon$ -nal, a második részre azonban, azaz az $\{F_n(x) > (1+\varepsilon)x - \varepsilon, a_2 \leq x < 1\}$ esemény valószínűségének meghatározásánál az 1.4.2. Korolláriumnál $a_2 = 1$ esetén alkalmazott egyszerűsítés lehetséges.

1.5.3. KOROLLÁRIUM. Tekintsük a

$$\sup_{\substack{F_n(x) \leq A_1 \\ A_2 \leq F_n(x)}} (x - F_n(x)), \quad \text{ill. } a \quad \sup_{\substack{F_n(x) \leq 1-A_2 \\ 1-A_1 \leq F_n(x)}} (F_n(x) - x)$$

statisztikákat, melyeknek eloszlása az 1.1. Lemma értelmében azonos. Ekkor

$$\begin{aligned}
 (1.24) \quad P\left(\sup_{\substack{F_n(x) \leq A_1 \\ F_n(x) \geq A_2}} (x - F_n(x)) < \varepsilon\right) &= P\left(\sup_{\substack{F_n(x) \leq 1-A_2 \\ F_n(x) \geq 1-A_1}} (F_n(x) - x) < \varepsilon\right) = \\
 &= \sum_{\alpha > nA_2} \binom{n}{\alpha} (A_2 + \varepsilon)^\alpha (1 - A_2 - \varepsilon)^{n-\alpha} \times \\
 &\quad \times \left(1 - \frac{\varepsilon}{A_2 + \varepsilon} \sum_{j=0}^{[nA_1]} \binom{\alpha}{j} \left(\frac{j}{A_2 + \varepsilon}\right)^{j-1} \left(1 - \frac{j}{A_2 + \varepsilon}\right)^{\alpha-j}\right) \times \\
 &\quad \times \left(1 - \frac{\frac{\alpha}{n} - A_2}{1 - A_2 - \varepsilon} \sum_{i=0}^{[n(1-\varepsilon)]-\alpha} \binom{n-\alpha}{i} \left(\frac{\frac{\alpha+i}{n} - A_2}{1 - A_2 - \varepsilon}\right)^{i-1} \left(\frac{1-\varepsilon - \frac{\alpha+i}{n}}{1 - A_2 - \varepsilon}\right)^{n-\alpha-i}\right).
 \end{aligned}$$

Bizonyítás. Könnyen látható, hogy

$$P\left(\sup_{\substack{F_n(x) \leq A_1 \\ A_2 \leq F_n(x)}} (x - F_n(x)) < \varepsilon\right) = P\left(\sup_{\substack{x \leq a_1 \\ x \geq a_2}} (x - F_n(x)) < \varepsilon\right),$$

ahol $a_1 = A_1 + \varepsilon$, $a_2 = A_2 + \varepsilon$, így a formulát az 1.5.1. Korolláriumból kaphatjuk.

1.5.4. KOROLLÁRIUM. Tekintsük a

$$\sup_{\substack{F_n(x) \leq A_1 \\ A_2 \leq F_n(x)}} \left(\frac{x - F_n(x)}{1 - F_n(x)}\right), \quad \text{ill. a} \quad \sup_{\substack{F_n(x) \leq 1-A_2 \\ 1-A_1 \leq F_n(x)}} \left(\frac{F_n(x) - x}{F_n(x)}\right)$$

statisztikákat, melyek eloszlása az 1.1. Lemma értelmében azonos. Ekkor

$$\begin{aligned}
 (1.25) \quad P\left(\sup_{\substack{F_n(x) \leq A_1 \\ F_n(x) \geq A_2}} \left(\frac{x - F_n(x)}{1 - F_n(x)}\right) < \varepsilon\right) &= P\left(\sup_{\substack{F_n(x) \leq 1-A_2 \\ F_n(x) \geq 1-A_1}} \left(\frac{F_n(x) - x}{F_n(x)}\right) < \varepsilon\right) = \\
 &= \sum_{\alpha > nA_2} \binom{n}{\alpha} (A_2(1-\varepsilon) + \varepsilon)^\alpha (1-\varepsilon)^{n-\alpha} (1 - A_2)^{n-\alpha-1} \left(\frac{\alpha}{n} - A_2\right) \times \\
 &\quad \times \left(1 - \frac{\varepsilon}{A_2(1-\varepsilon) + \varepsilon} \sum_{j=0}^{[nA_1]} \binom{\alpha}{j} \left(\frac{j}{A_2(1-\varepsilon) + \varepsilon}\right)^{j-1} \left(\frac{A_2 - \frac{j}{n}}{A_2(1-\varepsilon) + \varepsilon}\right)^{\alpha-j}\right).
 \end{aligned}$$

Bizonyítás. Könnyen látható, hogy

$$P\left(\sup_{\substack{F_n(x) \leq A_1 \\ A_2 \leq F_n(x)}} \left(\frac{x - F_n(x)}{1 - F_n(x)}\right) < \varepsilon\right) = P\left(\sup_{\substack{x \leq a_1 \\ a_2 \leq x}} \left(\frac{x - F_n(x)}{1 - x}\right) < \frac{\varepsilon}{1 - \varepsilon}\right),$$

ahol $a_1 = A_1(1-\varepsilon) + \varepsilon$ és $a_2 = A_2(1-\varepsilon) + \varepsilon$, így az eloszlást az 1.5.2. Korolláriumból közvetlenül megkaphatjuk.

1.5.5. KOROLLÁRIUM. *Tekintsük a*

$$\max \left(\sup_{0 \leq x \leq \frac{1}{2}} \left(\frac{x - F_n(x)}{1-x} \right), \sup_{\frac{1}{2} \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) \right),$$

ill. a

$$\max \left(\sup_{0 \leq x \leq \frac{1}{2}} \left(\frac{F_n(x) - x}{1-x} \right), \sup_{\frac{1}{2} \leq x \leq 1} \left(\frac{F_n(x) - x}{x} \right) \right)$$

statisztikákat, melyeknek eloszlása az 1.1. Lemma értelmében azonos. Ekkor

$$\begin{aligned} (1.26) \quad & P \left(\max \left(\sup_{0 \leq x \leq \frac{1}{2}} \left(\frac{x - F_n(x)}{1-x} \right), \sup_{\frac{1}{2} \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) \right) < \varepsilon \right) = \\ & = P \left(\max \left(\sup_{0 \leq x \leq \frac{1}{2}} \left(\frac{F_n(x) - x}{1-x} \right), \sup_{\frac{1}{2} \leq x \leq 1} \left(\frac{F_n(x) - x}{x} \right) \right) < \varepsilon \right) = \\ & = \frac{1}{2^{2n}} \sum_{\alpha > \frac{n(1-\varepsilon)}{2}} \binom{n}{\alpha} \left(1 - \frac{2\varepsilon}{1+\varepsilon} \left[\sum_{j=0}^{\lfloor \frac{n(1-\varepsilon)}{2} \rfloor} \binom{\alpha}{j} \left(\frac{\frac{2j}{n} + 2\varepsilon}{1+\varepsilon} \right)^{j-1} \left(1 - \frac{\frac{2j}{n} + 2\varepsilon}{1+\varepsilon} \right)^{\alpha-j} \right] \right) \times \\ & \times \left(1 - \left(\frac{2\alpha}{n(1-\varepsilon)} - 1 \right) \left[\sum_{i=0}^{\lfloor \frac{n(1-\varepsilon)}{2} \rfloor - \alpha} \binom{n-\alpha}{i} \left(\frac{2(\alpha+i)}{n(1-\varepsilon)} - 1 \right)^{i-1} \left(2 - \frac{2(\alpha+i)}{n(1-\varepsilon)} \right)^{n-\alpha-i} \right] \right). \end{aligned}$$

Bizonyítás. Alkalmazzuk az 1.5. Tételt $a_1 = a_2 = 1/2$, $c_1 = 1 + \varepsilon$, $c_2 = 1 - \varepsilon$, $b_1 = -\varepsilon$, $b_2 = 0$ -val.

Tudomásunk szerint a fenti statisztikáknak csupán határeloszlásával foglalkoztak eddig (lásd 2.2. pont), a pontos eloszlások az irodalomban nem találhatók.

1.3. A Ballot-lemma egy kiterjesztése

Az előzőekben olyan statisztikákkal foglalkoztunk, melyek függvények suprémumának eloszlását adják meg. CHANG [20] és TANG [99] megadta az $\inf_{0 < F_n(x) \leq c/n} \frac{F_n(x)}{x}$ statisztika pontos eloszlását. Az eloszlás meghatározása természetesen az 1.2. pontban leírt módszerrel is megy, de az nem adja meg közvetlenül a legegyszerűbb formulát. Az eloszlás meghatározására egyszerűbb út kínálkozik az 1.1. Tétel egy kiterjesztésével. Először ezt adjuk meg.

1.2. LEMMA. Legyen (mint 1.1. Tételnél) $A_0, A_1, \dots, A_n$ teljes eseményrendszer, melyre $P(A_0) = p$, $P(A_i) = q$ ($i = 1, 2, \dots, n$). Természetesen $p + nq = 1$. Végezzünk

$n+k$ független kísérletet, melynek során az A_j esemény v_j -szer következett be ($j=0, 1, 2, \dots, n$). Ekkor

$$(1.27) \quad P(v_1 < 1, v_1 + v_2 < 2, \dots, v_1 + v_2 + \dots + v_n < n) = \\ = \begin{cases} p-kq, & \text{ha } k = -1, -2, \dots, -n, \\ p-kq + kq \sum_{i=0}^{k-1} \binom{n+k}{i} p^i (1-p)^{n+k-1-i}, & \text{ha } k = 0, 1, 2, \dots \end{cases}$$

Bizonyítás. A bizonyítás az 1.1. Tételen, ill. az ott alkalmazott módszeren alapszik. (Egyébként az 1.2. Lemma $k=0$ -ra az 1.1. Tételt adja.) Vizsgáljuk a fenti esemény feltételes valószínűségét v_0 rögzített értéke mellett.

$$\begin{aligned} & P(v_1 < 1, \dots, v_1 + \dots + v_n < n | v_0) = \\ & = P(v_1 < 1, \dots, v_1 + \dots + v_{n+k-v_0} < n+k-v_0 | v_0) = \\ & = P(v_1 < 1, \dots, v_1 + \dots + v_{n+k-v_0} < n+k-v_0 | v_1 + \dots + v_n = n+k-v_0) = \\ & = \frac{v_0-k}{n}, \end{aligned}$$

feltéve, hogy $v_0 \geq k$. Így $k \geq 0$ esetén

$$\begin{aligned} & P(v_1 < 1, \dots, v_1 + \dots + v_n < n) = \\ & = \sum_{v_0=k}^{n+k} \frac{v_0-k}{n} P(v_0) = \sum_{v_0=0}^{n+k} \frac{v_0-k}{n} P(v_0) + \frac{k}{n} \sum_{v_0=0}^{k-1} P(v_0) = \\ & = \frac{E(v_0)-k}{n} + \frac{k}{n} \sum_{v_0=0}^{k-1} \binom{n+k}{v_0} p^{v_0} (1-p)^{n+k-v_0} = \\ & = p-kq + kq \sum_{i=0}^{k-1} \binom{n+k}{i} p^i (1-p)^{n+k-1-i}. \end{aligned}$$

Ha pedig $-n \leq k < 0$, akkor

$$P(v_1 < 1, \dots, v_1 + \dots + v_n < n) = \sum_{v_0=0}^{n+k} \frac{v_0-k}{n} P(v_0) = \frac{(n+k)p-k}{n} = p-kq.$$

Ha $k=1$ és $p=q=\frac{1}{n+1}$, akkor az 1.2. Lemma a következőt adja:

$$P(v_1 < 1, \dots, v_1 + \dots + v_n < n) = \frac{1}{n+1} \left(1 - \frac{1}{n+1} \right)^n = \frac{n^n}{(n+1)^{n+1}}.$$

Ebből megkapható CHANG formulája [20] (lásd még TANG [99]):

1.6. TÉTEL.

$$(1.28) \quad P \left(\inf_{0 < F_n(x) \leq (c/n)} \frac{F_n(x)}{x} < \varepsilon \right) = \\ = \begin{cases} \left(1 - \frac{1}{n\varepsilon}\right)^n + \sum_{j=1}^{[n\varepsilon]} \binom{n}{j} \frac{(j-1)^{j-1} (n\varepsilon - j)^{n-j}}{(n\varepsilon)^n}, & \frac{1}{n} \leq \varepsilon \leq \frac{c}{n} \\ \left(1 - \frac{1}{n\varepsilon}\right)^n + \sum_{j=1}^c \binom{n}{j} \frac{(j-1)^{j-1} (n\varepsilon - j)^{n-j}}{(n\varepsilon)^n}, & \frac{c}{n} < \varepsilon \leq 1. \end{cases}$$

Az $\varepsilon=1, c=n$ eset közvetlenül az 1.2. Lemmából adódik, n helyett $n-1$ -gyel:

$$P \left(\inf_{0 < F_n(x) \leq 1} \frac{F_n(x)}{x} > 1 \right) = \frac{(n-1)^{n-1}}{n^n}.$$

Bizonyítás. $\left\{ \inf_{0 < F_n(x) \leq (c/n)} \frac{F_n(x)}{x} < \varepsilon \right\} = \left\{ F_n(x) < \varepsilon x \text{ vmely } x\text{-re, } \frac{1}{n\varepsilon} \leq x \leq \frac{c}{n\varepsilon} \right\}$, azaz vagy $F_n\left(\frac{1}{n\varepsilon}\right) = 0$, vagy van olyan x , melyre $\varepsilon x = F_n(x)$. Az $\left\{ F_n\left(\frac{1}{n\varepsilon}\right) = 0 \right\}$ esemény valószínűsége $\left(1 - \frac{1}{n\varepsilon}\right)^n$.

Ha $F_n\left(\frac{1}{n\varepsilon}\right) > 0$, legyen $x = \frac{j}{n\varepsilon}$ az első olyan x , melyre $\varepsilon x = F_n(x)$. Annak valószínűsége, hogy $\left\{ F_n\left(\frac{j}{n\varepsilon}\right) = \frac{j}{n} \right\}$, nyilván $\binom{n}{j} \left(\frac{j}{n\varepsilon}\right)^j \left(1 - \frac{j}{n\varepsilon}\right)^{n-j}$. Ezen feltétel mellett tekintsük annak valószínűségét, hogy $\left\{ F_n(x) > \varepsilon x, \frac{1}{n\varepsilon} \leq x < \frac{j}{n\varepsilon} \right\}$. Ha v_i jelöli a $\left(\frac{j-i}{n\varepsilon}, \frac{j-i+1}{n\varepsilon}\right)$ intervallumba eső mintaelemek számát ($i=1, 2, \dots, j-1$), akkor itt éppen a $\{v_1 < 1, \dots, v_1 + \dots + v_{j-1} \leq j-1\}$ eseményről van szó, melynek valószínűsége az 1.2. Lemma szerint $\frac{(j-1)^{j-1}}{j^j}$. Annak valószínűsége tehát, hogy $\frac{j}{n}$ magasságon történjék az első metszés,

$$\binom{n}{j} \left(\frac{j}{n\varepsilon}\right)^j \left(1 - \frac{j}{n\varepsilon}\right)^{n-j} \frac{(j-1)^{j-1}}{j^j} = \binom{n}{j} \frac{(j-1)^{j-1} (n\varepsilon - j)^{n-j}}{(n\varepsilon)^n}$$

és ezek j -re vonatkozó összegezéséből adódik (1.28).

1.4. Nem folytonos alapeloszlás esete

Az 1.3. Tétel számos speciális esetei közül megemlíti még egyet, ti. a nem folytonos alapeloszlás esetét. Ebben az esetben SCHMID [85] határozta meg a határeloszlásokat, míg CONOVER [30], ill. COBERLY—LEWIS [29] rekurziós formulákat adtak meg a pontos eloszlásra. VINCZE [104] vizsgálta a kétmintás esetet. Az alábbi-

akban megadjuk az egyoldali *Kolmogorov*-eltérés pontos eloszlását, explicit alakban, egymintás esetben. Más statisztikák (pl. relatív eltérés stb.) eloszlása is hasonlóképpen adható meg, ezek felsorolásától azonban eltekintünk.

Legyen tehát $F(x)$ egy eloszlásfüggvény, melynek az $x_1, x_2, \dots, x_r$ pontban vannak szakadásai, ahol $-\infty < x_1 < x_2 < \dots < x_r < \infty$, minden más pontban legyen $F(x)$ folytonos.

Legyen

$$p_i = F(x_i) - F(x_{i-1} + 0) \geq 0, \quad i = 1, 2, \dots, r+1$$

$$q_i = F(x_i + 0) - F(x_i), \quad i = 1, 2, \dots, r.$$

Igaz, hogy

$$\sum_{i=1}^{r+1} p_i + \sum_{i=1}^r q_i = 1.$$

Jelölje $F_n(x)$ az ebből az eloszlásból vett n elemű minta empirikus eloszlásfüggvényét, és vizsgáljuk a $\sup_{-\infty < x < \infty} (F(x) - F_n(x))$ statisztikát. Ekkor igaz a következő

1.3. LEMMA. $P\left(\sup_{-\infty < x < \infty} (F(x) - F_n(x)) < \varepsilon\right) = P(\bar{F}_n(x) > x - \varepsilon, 0 \leq x \leq p_1, p_1 + q_1 \leq x \leq p_1 + q_1 + p_2, \dots)$, ahol $\bar{F}_n(x)$ egy $(0, 1)$ -ben egyenletes eloszlásból vett minta empirikus eloszlásfüggvénye.

Bizonyítását lásd NOETHER [72], SCHMID [85], VINCZE [104], WALSH [107]. A fenti lemmából következik, hogy

$$\begin{aligned} P\left(\sup_{-\infty < x < \infty} (F(x) - F_n(x)) < \varepsilon\right) &\geq P(\bar{F}_n(x) > x - \varepsilon, 0 \leq x \leq 1) = \\ &= 1 - \varepsilon \sum_{j=0}^{[n(1-\varepsilon)]} \binom{n}{j} \left(\frac{j}{n} + \varepsilon\right)^{j-1} \left(1 - \varepsilon - \frac{j}{n}\right)^{n-j}. \end{aligned}$$

Az 1.3. lemmát felhasználva kapjuk a következő tételt:

1.7. TÉTEL.

$$\begin{aligned} (1.29) \quad P\left(\sup_{-\infty < x < \infty} (F(x) - F_n(x)) < \varepsilon\right) &= \\ &= n! \sum_{\substack{\alpha_1 + \dots + \alpha_{r+1} = n \\ \beta_i > n(a_i - \varepsilon)}} \prod_{i=1}^{r+1} \frac{(p_i + q_i)^{\alpha_i}}{\alpha_i!} p_n\left(p_i, q_i, \frac{\beta_{i-1}}{n} + \varepsilon - a_{i-1}; \alpha_i\right), \\ \text{ahol} \quad p_n(u_1, u_2, z; \alpha) &= \begin{cases} 1 - \frac{z}{u^\alpha} \sum_{j=0}^{[n(u_1 - z)]} \binom{\alpha}{j} \left(z + \frac{j}{n}\right)^{j-1} \left(u - z - \frac{j}{n}\right)^{\alpha-j}, & u_1 \geq z, \\ 1 & \text{különben,} \end{cases} \\ u &= u_1 + u_2, \quad a_0 = 0, \quad \beta_0 = 0, \quad q_{r+1} = 0, \\ a_i &= \sum_{j=1}^i (p_j + q_j), \quad \beta_i = \alpha_1 + \dots + \alpha_i, \quad i = 1, 2, \dots, r+1. \end{aligned}$$

Bizonyítás. Legyen v_i az (a_{i-1}, a_i) intervallumba eső mintaelemek száma, $i=1, 2, \dots, r+1$. A $\{v_1=\alpha_1, v_2=\alpha_2, \dots, v_{r+1}=\alpha_{r+1}\}$ esemény valószínűségét a polinomiális eloszlásból kaphatjuk. Ezen feltétel mellett az $\{\bar{F}_n(x) > x - \varepsilon, a_{i-1} < x < a_{i-1} + p_i\}$ események $i=1, 2, \dots, r+1$ -re függetlenek.

Nyilván $\bar{F}_n(a_{i-1}) = \frac{\beta_{i-1}}{n}$ és annak valószínűsége, hogy $\bar{F}_n(x)$ a $\frac{\beta_{i-1}+j}{n}$ magasságon áttesse az $y = x - \varepsilon$ egyenest, $\binom{\alpha_i}{j} x_j^j (1-x_j)^{\alpha_i-j}$, ahol $x_j = \frac{\frac{\beta_{i-1}+j}{n} - \varepsilon - a_{i-1}}{a_i - a_{i-1}}$. Annak valószínűsége, hogy itt legyen az első metszés az (a_{i-1}, a_i) intervallumban, az 1.1. Tételből $\frac{x_0}{x_j}$. Így az $x_0 \binom{\alpha_i}{j} x_j^{j-1} (1-x_j)^{\alpha_i-j}$ valószínűségeket kell összegezni olyan j -kre, melyre $\frac{\beta_{i-1}+j}{n} \leq a_{i-1} + p_i - \varepsilon$ és 1-ből levonni. Ezzel kapjuk az 1.7. Tétel állítását.

Amennyiben az 1.3. tételt közvetlenül akarjuk alkalmazni, a $(0, 1)$ intervallumot nem $r+1$, hanem $2r+1$ részre kell bontanunk, így $2r+1$ változós összeget kapunk. Ezt a formulát is megadjuk, ui. célszerűbb ennek alapján számítani a határeloszlást.

$$(1.30) \quad P\left(\sup_{-\infty < x < \infty} (F(x) - F_n(x)) < \varepsilon\right) =$$

$$= \sum_{\substack{\gamma_1 + \dots + \gamma_{r+1} + \delta_1 + \dots + \delta_r = n \\ \gamma_1 + \delta_1 + \dots + \gamma_i + \delta_i > n(a_i - \varepsilon)}} \frac{n!}{\gamma_1! \dots \gamma_{r+1}! \delta_1! \dots \delta_r!} p_1^{\gamma_1} \dots p_{r+1}^{\gamma_{r+1}} q_1^{\delta_1} \dots q_r^{\delta_r} \times$$

$$\times \prod_{i=1}^{r+1} p_n \left(p_i, \frac{\gamma_1 + \delta_1 + \dots + \gamma_{i-1} + \delta_{i-1}}{n} + \varepsilon - a_{i-1}; \gamma_i \right).$$

1.5. A Pyke-féle empirikus eloszlásfüggvény

Legyen, mint az 1.2. pontban $X_1, X_2, \dots, X_n$ egy n -elemű véletlen minta a $(0, 1)$ intervallumban egyenletes eloszlásból. Legyen továbbá $X_0^* = 0 < X_1^* < X_2^* < \dots < X_n^* < X_{n+1}^* = 1$ a megfelelő rendezett minta. PYKE [74] vezette be a következő, módosított empirikus eloszlásfüggvényt:

$$(1.31) \quad F_n^*(x) = \frac{1}{n+1} \left(\frac{x - X_{i-1}^*}{X_i^* - X_{i-1}^*} + i - 1 \right), \quad X_{i-1}^* \leq x \leq X_i^* \quad i = 1, 2, \dots, n+1.$$

Ez az $F_n^*(x)$ függvény egy folytonos törött vonal, amely az $\left(X_i^*, \frac{i}{n+1}\right)$ pontokat köti össze.

A Pyke-féle empirikus eloszlásfüggvényre vonatkozó egyoldali eltérések visszavezethetők a közönséges empirikus eloszlásfüggvényre vonatkozó egyoldali eltérésekre. Igaz ui. a következő lemma:

1.4. LEMMA. Legyen $G(x)$ monoton növekvő, alulról konvex függvény a $(0, 1)$ intervallumban. Ekkor

$$(1.32) P(F_n^*(x) > G(x), 0 \leq x \leq 1) = P\left(F_n(x) > \frac{n+1}{n} G(x) - \frac{1}{n}, 0 \leq x \leq 1\right).$$

Bizonyítás. Legyen tehát $G(x)$ monoton növekvő, alulról konvex függvény. Ekkor az $F_n^*(x) > G(x)$ egyenlőtlenség teljesülését elég megkövetelni az $x = X_i^*$ pontokban, ez már maga után vonja az egyenlőtlenség teljesülését az egész $(0, 1)$ intervallumon. Mivel $F_n^*(X_i^*) = \frac{i}{n+1}$, ezért

$$\begin{aligned} P(F_n^*(x) > G(x), 0 \leq x \leq 1) &= P\left(\frac{i}{n+1} > G(X_i^*), i = 0, 1, \dots, n+1\right) = \\ &= P\left(\frac{i-1}{n} \geq \frac{n+1}{n} G(X_i^*) - \frac{1}{n}, i = 0, 1, \dots, n+1\right) = \\ &= P\left(F_n(X_i^*) > \frac{n+1}{n} G(X_i^*) - \frac{1}{n}, i = 0, 1, \dots, n+1\right) = \\ &= P\left(F_n(x) > \frac{n+1}{n} G(x) - \frac{1}{n}, 0 \leq x \leq 1\right). \end{aligned}$$

E lemmát felhasználva, megadjuk néhány legfontosabb eltérés eloszlását. A részletes felsorolástól itt is eltekintünk.

$$\begin{aligned} (1.33) \quad P\left(\sup_{0 \leq x \leq 1} (x - F_n^*(x)) < \varepsilon\right) &= \\ &= 1 - \left(\varepsilon + \frac{1}{n+1}\right)^{\left[(n+1)(1-\varepsilon)]-1} \sum_{j=0}^n \binom{n}{j} \left(\frac{j+1}{n+1} + \varepsilon\right)^{j-1} \left(1 - \varepsilon - \frac{j+1}{n+1}\right)^{n-j}. \end{aligned}$$

(Lásd PYKE [74], BRUNK [16]).

$$(1.34) \quad P\left(\sup_{0 \leq x \leq 1} \left(\frac{x - F_n^*(x)}{1-x}\right) < \varepsilon\right) = \frac{\varepsilon}{1+\varepsilon} + \frac{1}{(n+1)(1+\varepsilon)}.$$

$$\begin{aligned} (1.35) \quad P\left(\sup_{0 \leq x \leq 1} \left(\frac{x - F_n^*(x)}{x}\right) < \varepsilon\right) &= \\ &= 1 - \frac{1}{(n+1)(1-\varepsilon)} \sum_{j=0}^{[n(1-\varepsilon)]-1} \binom{n}{j} \left(\frac{j+1}{(n+1)(1-\varepsilon)}\right)^{j-1} \left(1 - \frac{j+1}{(n+1)(1-\varepsilon)}\right)^{n-j}. \end{aligned}$$

2. §. HATÁRELOSZLÁS TÉTELEK

2.1. Néhány segédteétel

Az 1. §-ban megadott véges eloszlástételekhez szorosan kapcsolódnak a határeloszlás tételek, midőn a mintaelemszám, n a végtelenhez tart.

Mint a Bevezetőben is említettük, az irodalomban a határeloszlás tételek megadása történetileg megelőzte a véges eloszlástételek megadását. KOLMOGOROV [64] úttörő jellegű munkájában a $\sup_{-\infty < x < \infty} |F_n(x) - F(x)|$ statisztika határeloszlását adja meg. A későbbiekben sokan foglalkoztak határeloszlások meghatározásával, ezek közül a legfontosabb és legérdekesebb cikkek: SZMIRNOV [91], FELLER [51], RÉNYI [77], ANDERSON—DARLING [5], DOOB [43], DONSKER [41], SCHMID [85], LAUWERIER [66].

E cikkekben különféle módszereket dolgoztak ki a határeloszlások meghatározására. Viszonylag kevés azon cikkek száma, melyekben a határeloszlásokat a véges eloszlástételek folyományaként, határátmenettel nyerték. Ezek közül elsősorban SZMIRNOV [92], [93], cikkeit kell említenünk. Disszertációnkban ezt az utat követjük, azaz az 1. §-ban nyert véges eloszlástételekből származtatunk határeloszlás tételeket.

Szükségünk lesz az alábbi lemmákra:

2.1. LEMMA. Ha $0 < K \leq 1$, $0 < B$, $0 < A + B$, akkor

$$(2.1) \quad \frac{1}{\sqrt{2\pi}} \int_0^K \frac{B}{z\sqrt{z(1-z)}} e^{-\frac{1}{2} \frac{(Az+B)^2}{z(1-z)}} dz = \\ = e^{-2B(A+B)} \Phi\left(\frac{AK-B+2BK}{\sqrt{K(1-K)}}\right) + 1 - \Phi\left(\frac{AK+B}{\sqrt{K(1-K)}}\right),$$

ahol itt és a továbbiakban mindenütt

$$\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^u e^{-\frac{v^2}{2}} dv.$$

MEGJEGYZÉS. A $K=0$ eset érdektelen (bár a formula ebben az esetben is igaz), $K=1$ -re (2.1) a következőt adja:

$$(2.2) \quad \frac{1}{\sqrt{2\pi}} \int_0^1 \frac{B}{z\sqrt{z(1-z)}} e^{-\frac{1}{2} \frac{(Az+B)^2}{z(1-z)}} dz = e^{-2B(A+B)}.$$

Bizonyítás. Legyen

$$I_1 = \frac{1}{\sqrt{2\pi}} \int_0^K \frac{(A+B)z + B(1-z)}{2(z(1-z))^{3/2}} e^{-\frac{[(A+B)z + B(1-z)]^2}{2z(1-z)}} dz = \\ = \frac{1}{\sqrt{2\pi}} e^{-2B(A+B)} \int_0^K \frac{(A+B)z + B(1-z)}{2(z(1-z))^{3/2}} e^{-\frac{[(A+B)z + B(1-z)]^2}{2z(1-z)}} dz.$$

Elvégezve az $\frac{(A+B)z-B(1-z)}{\sqrt{z(1-z)}}=u$ helyettesítést, kapjuk

$$I_1 = e^{-2B(A+B)} \int_{-\infty}^{\frac{(A+B)K-B(1-K)}{\sqrt{K(1-K)}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du = e^{-2B(A+B)} \Phi \left(\frac{AK-B+2BK}{\sqrt{K(1-K)}} \right).$$

Legyen most

$$I_2 = \frac{1}{\sqrt{2\pi}} \int_0^K \frac{(A+B)z-B(1-z)}{2(z(1-z))^{3/2}} e^{-\frac{[(A+B)z+B(1-z)]^2}{2z(1-z)}} dz.$$

Elvégezve az $\frac{(A+B)z+B(1-z)}{\sqrt{z(1-z)}}=v$ helyettesítést, kapjuk

$$\begin{aligned} I_2 &= \frac{1}{\sqrt{2\pi}} \int_{\infty}^{\frac{(A+B)K+B(1-K)}{\sqrt{K(1-K)}}} e^{-\frac{v^2}{2}} dv = - \left(1 - \int_{-\infty}^{\frac{AK+B}{\sqrt{K(1-K)}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv \right) = \\ &= - \left(1 - \Phi \left(\frac{AK+B}{\sqrt{K(1-K)}} \right) \right). \end{aligned}$$

A keresett integrál:

$$\begin{aligned} &\frac{1}{\sqrt{2\pi}} \int_0^K \frac{B}{z \sqrt{z(1-z)}} e^{-\frac{1}{2} \frac{(Az+B)^2}{z(1-z)}} dz = I_1 - I_2 = \\ &= e^{-2B(A+B)} \Phi \left(\frac{AK-B+2BK}{\sqrt{K(1-K)}} \right) + 1 - \Phi \left(\frac{AK+B}{\sqrt{K(1-K)}} \right), \quad \text{q.e.d.} \end{aligned}$$

A határeloszlások meghatározásához a továbbiakban először

$$\frac{B \sqrt{pn}}{\alpha - A \sqrt{pn}} \binom{\alpha}{j} \left(\frac{j+B \sqrt{pn}}{\alpha - A \sqrt{pn}} \right)^{j-1} \left(1 - \frac{j+B \sqrt{pn}}{\alpha - A \sqrt{pn}} \right)^{\alpha-j}$$

alakú kifejezések aszimptotikus értékét kell megadnunk, ahol A, B, p konstansok és $\alpha \approx np, j = \alpha z$. Erre vonatkozik a következő lemma.

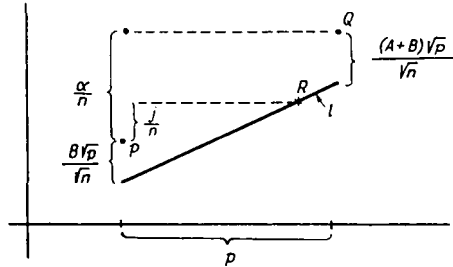
2.2. LEMMA. *Legyenek A, B, p, z n -től független konstansok, $\alpha \approx np, j = \alpha z$. Ekkor $n \rightarrow \infty$ esetén*

$$\begin{aligned} (2.3) \quad &\frac{B \sqrt{pn}}{\alpha - A \sqrt{pn}} \binom{\alpha}{j} \left(\frac{j+B \sqrt{pn}}{\alpha - A \sqrt{pn}} \right)^{j-1} \left(1 - \frac{j+B \sqrt{pn}}{\alpha - A \sqrt{pn}} \right)^{\alpha-j} \sim \\ &\sim \frac{1}{\alpha} \frac{B}{\sqrt{2\pi} z \sqrt{z(1-z)}} e^{-\frac{1}{2} \frac{(Az+B)^2}{z(1-z)}}. \end{aligned}$$

MEGJEGYZÉS. A fenti aszimptotikus érték a 2.1. Lemmában szereplő integrandus, $\frac{1}{\alpha} \approx dz$, így a fenti alakú összegek a 2.1. Lemmában kiszámított integrálba mennek át.

A baloldalon álló valószínűséget a következő ábrával szemléltetjük:

A fenti kifejezés annak valószínűsége, hogy ha $F_n(x)$ a P és Q pontokon keresztül megy, az R pontban messe először az l egyenest.



1. ábra

Bizonyítás. A de Moivre—Laplace-tétel értelmében

$$\binom{N}{j} q^j (1-q)^{N-j} \sim \frac{1}{\sqrt{2\pi N q(1-q)}} e^{-\frac{(j-Nq)^2}{2Nq(1-q)}}.$$

Alkalmazzuk ezt $N = \alpha \approx np$, $q = \frac{j+B\sqrt{pn}}{\alpha-A\sqrt{pn}} \sim z$ -re. Könnyen látható, hogy

$$\frac{(j-\alpha q)^2}{\alpha q(1-q)} = \frac{\left(\frac{B\alpha\sqrt{pn} + A\alpha z\sqrt{pn}}{\alpha - A\sqrt{pn}} \right)^2}{\alpha q(1-q)} \sim \frac{pn(Az+B)^2}{\alpha z(1-z)} \sim \frac{(Az+B)^2}{z(1-z)},$$

így

$$\begin{aligned} \frac{B\sqrt{pn}}{\alpha - A\sqrt{pn}} \binom{\alpha}{j} q^{j-1} (1-q)^{\alpha-j} &= \frac{B\sqrt{pn}}{B\sqrt{pn}+j} \binom{\alpha}{j} q^j (1-q)^{\alpha-j} \sim \\ &\sim \frac{B\sqrt{pn}}{j} \frac{1}{\sqrt{2\pi\alpha z(1-z)}} e^{-\frac{(Az+B)^2}{2z(1-z)}} \sim \frac{B}{\alpha z \sqrt{2\pi z(1-z)}} e^{-\frac{(Az+B)^2}{2z(1-z)}}. \end{aligned}$$

2.2. Határeloszlás tételek

A továbbiakban megadjuk az 1. § tételeinek megfelelő határeloszlás tételeket. A sort azonban nem az 1.1., hanem az 1.2. Tétellel kezdjük.

2.1. TÉTEL. Legyen $0 < a \leq 1$, $v > 0$, $v - ua \geq 0$. Ekkor

$$\begin{aligned} (2.4) \quad \lim_{n \rightarrow \infty} P \left(F_n(x) > \left(1 + \frac{u}{\sqrt{n}} \right) x - \frac{v}{\sqrt{n}}, 0 \leq x \leq a \right) &= \\ &= \Phi \left(\frac{v-ua}{\sqrt{a(1-a)}} \right) - e^{-2v(v-u)} \Phi \left(\frac{2av-v-ua}{\sqrt{a(1-a)}} \right). \end{aligned}$$

MEGJEGYZÉS. $a=1$ -re kapjuk

$$(2.5) \quad \lim_{n \rightarrow \infty} P \left(F_n(x) > \left(1 + \frac{u}{\sqrt{n}} \right) x - \frac{v}{\sqrt{n}}, 0 \leq x \leq 1 \right) = 1 - e^{-2v(v-u)}.$$

Bizonyítás. Alkalmazzuk az 1.2. Tételt $c=1+\frac{u}{\sqrt{n}}$, $b=\frac{v}{\sqrt{n}}$ -el, majd a 2.2. Lemmát $\alpha=n$, $p=1$, $B=v$, $A=-u$ értékekkel. E szerint az integrandus a következő lesz:

$$\frac{v}{\sqrt{2\pi} z \sqrt{z(1-z)}} e^{-\frac{(v-uz)^2}{2z(1-z)}}.$$

Az összegzésben j -re a felső határ $[n(ca-b)]$, így a z -re vonatkozó felső határt a következőképpen kapjuk:

$$j \leq n(ca-b)$$

$$nz \leq n \left(\left(1 + \frac{u}{\sqrt{n}} \right) a - \frac{v}{\sqrt{n}} \right)$$

$$z \leq a + \frac{au-v}{\sqrt{n}} \sim a.$$

Alkalmazzuk tehát a 2.1. Lemmát $K=a$ -val, így kapjuk a tétel állítását.

Az 1.3. Tételnek megfelelően tekintsünk most egy szakaszonként lineáris $G_n(x)$ függvényt, amely tartson az $y=x$ egyeneshez. Legyen tehát

$$G_n(x) = \left(1 + \frac{u_i}{\sqrt{n}} \right) x - \frac{v_i}{\sqrt{n}}, \quad a_{i-1} < x \leq a_i, \quad i = 1, 2, \dots, k,$$

ahol $a_0=0 < a_1 < a_2 < \dots < a_{k-1} < a_k=1$, u_i és v_i rögzítettek.

2.2. TÉTEL. *A fenti feltételek mellett igaz a következő határérték reláció:*

$$(2.6) \quad \lim_{n \rightarrow \infty} P(F_n(x) \leq G_n(x), 0 \leq x \leq 1) =$$

$$= \int \dots \int \frac{1}{\sqrt{2\pi}^{\frac{k-1}{2}}} \frac{1}{\sqrt{p_k}} e^{-\frac{1}{2} \sum_{i=1}^k y_i^2} \times$$

$$\left\{ \sum_{j=1}^i y_j \sqrt{p_j} \leq u_{i+1} a_i - v_{i+1} \right\} \quad \left\{ \begin{matrix} i=1, 2, \dots, k-1 \end{matrix} \right\}$$

$$\times \prod_{i=1}^k \left(1 - e^{-\frac{2 \left(v_i - u_i a_{i-1} + \sum_{j=1}^{i-1} y_j \sqrt{p_j} \right) \left(v_i - u_i a_i + \sum_{j=1}^i y_j \sqrt{p_j} \right)}{p_i}} \right) dy_1 dy_2 \dots dy_{k-1},$$

ahol

$$y_k = -\frac{\sum_{j=1}^{k-1} y_j \sqrt{p_j}}{\sqrt{p_k}}, \quad p_j = a_j - a_{j-1}.$$

Bizonyítás. Alkalmazzuk az 1.3. Tételt $c_i = 1 + \frac{u_i}{\sqrt{n}}$, $b_i = -\frac{v_i}{\sqrt{n}}$ -nel, majd helyettesítsünk $\alpha_i = n(a_i - a_{i-1}) + y_i \sqrt{n(a_i - a_{i-1})}$; $j = \alpha_i z_i$ -t. Ekkor $n \rightarrow \infty$ határátmenet esetén a polinomiális valószínűség többváltozós normális sűrűségfüggvénybe, a $p_n(\cdot)$ függvény pedig a 2.2. és 2.1. Lemma szerint az

$$1 - \exp \left(\frac{-2 \left(v_i - u_i a_{i-1} + \sum_{j=1}^{i-1} y_j \sqrt{a_j - a_{j-1}} \right) \left(v_i - u_i a_i + \sum_{j=1}^i y_j \sqrt{a_j - a_{j-1}} \right)}{a_i - a_{i-1}} \right)$$

függvénybe, az összeg pedig integrálba megy át.

Az alábbi tételek a 2.2. Tétel speciális esetei, de — mint az 1. §-ban — külön tételként mondjuk ki.

2.3. TÉTEL. Legyen $0 \leq a_1 < a_2 \leq 1$ rögzített, $v \geq 0$, $v - u \geq 0$.

Ekkor

$$(2.7) \quad \lim_{n \rightarrow \infty} P \left(F_n(x) > \left(1 + \frac{u}{\sqrt{n}} \right) x - \frac{v}{\sqrt{n}}, a_1 \leq x \leq a_2 \right) =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{ua_1 - v}{\sqrt{a_1(1-a_1)}}}^{\infty} e^{-\frac{y^2}{2}} \left[\Phi \left((v-u) \sqrt{\frac{a_2 - a_1}{(1-a_1)(1-a_2)}} + \left(y \sqrt{a_1} - \frac{a_1 u - v}{\sqrt{1-a_1}} \right) \sqrt{\frac{1-a_2}{a_2 - a_1}} \right) - \right.$$

$$\left. - e^{-2 \left(y \sqrt{a_1} - \frac{a_1 u - v}{\sqrt{1-a_1}} \right) \left(\frac{v-u}{\sqrt{1-a_1}} \right)} \right] \times$$

$$\times \Phi \left((v-u) \sqrt{\frac{a_2 - a_1}{(1-a_1)(1-a_2)}} - \left(y \sqrt{a_1} - \frac{a_1 u - v}{\sqrt{1-a_1}} \right) \sqrt{\frac{1-a_2}{a_2 - a_1}} \right) dy.$$

Bizonyítás. Legyen $b = -\frac{v}{\sqrt{n}}$, $c = 1 + \frac{u}{\sqrt{n}}$, és alkalmazzuk az 1.4. Tétel első felét, tekintve, hogy most $ca_1 + b = \left(1 + \frac{u}{\sqrt{n}} \right) a_1 - \frac{v}{\sqrt{n}} > 0$ elég nagy n -re, feltéve, hogy $a_1 > 0$. Ha $a_1 = 0$, akkor közvetlenül a 2.1. Tételt alkalmazhatjuk.

Írjuk át az (1.13) formulát a következőképpen:

$$\sum_{\alpha > n(ca_1 + b)} \binom{n}{\alpha} a_1^\alpha (1-a_1)^{n-\alpha} \left(1 - \left(\frac{x_0 - a_1}{1-a_1} \right)^{\sum_{j=0}^{n(ca_2+b)-\alpha} (n-\alpha)} \right) \times$$

$$\times \left(\frac{x_j - a_1}{1-a_1} \right)^{j-1} \left(\frac{1-x_j}{1-a_1} \right)^{n-\alpha-j},$$

ahol

$$x_j = \frac{\alpha + j}{nc} - \frac{b}{c}.$$

Legyen $\alpha = na_1 + y\sqrt{na_1(1-a_1)}$, ekkor az $\binom{n}{\alpha} a_1^\alpha (1-a_1)^{n-\alpha}$ binomiális valószínűségekből $\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}$ dy lesz, mikor is az $\alpha > n(ca_1 + b)$ egyenlőtlenség átmegy az $y > \frac{ua_1 - v}{\sqrt{a_1(1-a_1)}}$ egyenlőtlenségbe.

A belső zárójelben levő kifejezés határértékének meghatározásához pedig alkalmazzuk a 2.2. Lemmát, α helyett $n-\alpha$ -val ($n-\alpha \approx n(1-a_1)$), $p=1-a_1$, $B=y\sqrt{a_1} + \frac{v-ua_1}{\sqrt{1-a_1}}$, $A=-y\sqrt{a_1} - u\sqrt{1-a_1}$ helyettesítéssel. Végül a 2.1. Lemma adja a tétel állítását.

2.3.1. KOROLLÁRIUM.

$$(2.8) \lim_{n \rightarrow \infty} P \left(\sup_{a_1 \leq x \leq a_2} (x - F_n(x)) < \frac{t}{\sqrt{n}} \right) = \lim_{n \rightarrow \infty} P \left(\sup_{1-a_2 \leq x \leq 1-a_1} (F_n(x) - x) < \frac{t}{\sqrt{n}} \right) =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{\frac{t}{\sqrt{a_1(1-a_1)}}}^{\infty} e^{-\frac{y^2}{2}} \left[\Phi \left(t \sqrt{\frac{a_2-a_1}{(1-a_1)(1-a_2)}} + \left(y\sqrt{a_1} + \frac{t}{\sqrt{1-a_1}} \right) \sqrt{\frac{1-a_2}{a_2-a_1}} \right) - \right.$$

$$\left. - e^{-2 \left(y\sqrt{a_1} + \frac{t}{\sqrt{1-a_1}} \right) \frac{t}{\sqrt{1-a_1}}} \Phi \left(t \sqrt{\frac{a_2-a_1}{(1-a_1)(1-a_2)}} - \left(y\sqrt{a_1} + \frac{t}{\sqrt{1-a_1}} \right) \sqrt{\frac{1-a_2}{a_2-a_1}} \right) \right] dy.$$

Ha $a_1=0$, akkor

$$(2.9) \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a_2} (x - F_n(x)) < \frac{t}{\sqrt{n}} \right) = \Phi \left(\frac{t}{\sqrt{a_2(1-a_2)}} \right) - e^{-2t^2} \Phi \left(\frac{(2a_2-1)t}{\sqrt{a_2(1-a_2)}} \right),$$

míg ha $a_2=1$, akkor

$$(2.10) \lim_{n \rightarrow \infty} P \left(\sup_{a_1 \leq x \leq 1} (x - F_n(x)) < \frac{t}{\sqrt{n}} \right) = \Phi \left(\frac{t}{\sqrt{a_1(1-a_1)}} \right) - e^{-2t^2} \Phi \left(\frac{(1-2a_1)t}{\sqrt{a_1(1-a_1)}} \right),$$

végül ha $a_1=0$ és $a_2=1$, akkor

$$(2.11) \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \frac{t}{\sqrt{n}} \right) = 1 - e^{-2t^2}, \quad t > 0.$$

Bizonyítás. Helyettesítsünk a 2.3. Tétel (2.7) képletébe $u=0$, $v=t$.

Az $a_1=0$, $a_2=1$ eset SZMIRNOVTÓL [91] származik, míg az általános esetben ekvivalens formulát adott meg MANIJA [68].

2.3.2. KOROLLÁRIUM.

$$\begin{aligned}
 (2.12) \quad \lim_{n \rightarrow \infty} P \left(\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{1-x} \right) < \frac{t}{\sqrt{n}} \right) &= \lim_{n \rightarrow \infty} P \left(\sup_{1-a_2 \leq x \leq 1-a_1} \left(\frac{F_n(x) - x}{x} \right) < \frac{t}{\sqrt{n}} \right) = \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-t\sqrt{\frac{1-a_1}{a_1}}}^{\infty} e^{-\frac{y^2}{2}} \times \\
 &\times \left[\Phi \left((y\sqrt{a_1} + t\sqrt{1-a_1}) \sqrt{\frac{1-a_2}{a_2-a_1}} \right) - \Phi \left(-(y\sqrt{a_1} + t\sqrt{1-a_1}) \sqrt{\frac{1-a_2}{a_2-a_1}} \right) \right] dy.
 \end{aligned}$$

$a_2=1$ esetén a jobboldal 0, míg $a_1=0$ -ra

$$(2.13) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a_2} \left(\frac{x - F_n(x)}{1-x} \right) < \frac{t}{\sqrt{n}} \right) = \Phi \left(t \sqrt{\frac{1-a_2}{a_2}} \right) - \Phi \left(-t \sqrt{\frac{1-a_2}{a_2}} \right).$$

Bizonyítás. Helyettesítsünk a 2.3. Tétel képletébe $u=v=t$.

Ezt a határeloszlást először RÉNYI [77] határozta meg. Lásd még CSÖRGŐ [35], SAHLER [82].

2.3.3. KOROLLÁRIUM.

$$\begin{aligned}
 (2.14) \quad \lim_{n \rightarrow \infty} P \left(\sup_{a_1 \leq x \leq a_2} \left(\frac{x - F_n(x)}{x} \right) < \frac{t}{\sqrt{n}} \right) &= \lim_{n \rightarrow \infty} P \left(\sup_{1-a_2 \leq x \leq 1-a_1} \left(\frac{F_n(x) - x}{1-x} \right) < \frac{t}{\sqrt{n}} \right) = \\
 &= \frac{1}{\sqrt{2\pi}} \int_{-t\sqrt{\frac{a_1}{1-a_1}}}^{\infty} e^{-\frac{y^2}{2}} \left[\Phi \left(t \sqrt{\frac{a_2-a_1}{(1-a_1)(1-a_2)}} + \left(y\sqrt{a_1} + \frac{a_1 t}{\sqrt{1-a_1}} \right) \sqrt{\frac{1-a_2}{a_2-a_1}} \right) - \right. \\
 &\left. - e^{-2 \left(y\sqrt{a_1} + \frac{a_1 t}{\sqrt{1-a_1}} \right) \frac{t}{\sqrt{1-a_1}}} \Phi \left(t \sqrt{\frac{a_2-a_1}{(1-a_1)(1-a_2)}} - \left(y\sqrt{a_1} + \frac{a_1 t}{\sqrt{1-a_1}} \right) \sqrt{\frac{1-a_2}{a_2-a_1}} \right) \right] dy,
 \end{aligned}$$

az $a_1=0$ eset értelmetlen, míg $a_2=1$ esetén

$$(2.15) \quad \lim_{n \rightarrow \infty} P \left(\sup_{a_1 \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \frac{t}{\sqrt{n}} \right) = \Phi \left(t \sqrt{\frac{a_1}{1-a_1}} \right) - \Phi \left(-t \sqrt{\frac{a_1}{1-a_1}} \right).$$

Bizonyítás. Helyettesítsünk a 2.3. Tétel képletébe $v=0, u=-t$.

Ez a határeloszlás ugyancsak RÉNYITŐL [77] származik. Lásd még CSÖRGŐ [35], SAHLER [82].

Itt nem térünk ki a $\sup_{A_1 \leq F_n(x) \leq A_2} (x - F_n(x))$, stb. statisztikák határeloszlásaira (lásd 1.4.4.—1.4.6. Korolláriumok), mert ezek megegyeznek a már tárgyalt $\sup_{a_1 \leq x \leq a_2} (x - F_n(x))$ stb. statisztikák határeloszlásaival.

Most vizsgáljuk az 1.5. Tételnek megfelelő határeloszlásokat.

2.4. TÉTEL. Legyen $0 < a_1 \leq a_2 < 1$, ekkor

(2.16)

$$\begin{aligned} \lim_{n \rightarrow \infty} P \left(F_n(x) > \left(1 + \frac{u_1}{\sqrt{n}} \right) x - \frac{v_1}{\sqrt{n}}, 0 \leq x \leq a_1; F_n(x) > \left(1 + \frac{u_2}{\sqrt{n}} \right) x - \frac{v_2}{\sqrt{n}}, a_2 \leq x \leq 1 \right) = \\ = \frac{1}{\sqrt{2\pi}} \int_{\frac{u_2 a_2 - v_2}{\sqrt{a_2(1-a_2)}}}^{\infty} e^{-\frac{y^2}{2}} \left(1 - e^{-2 \left(\frac{y \sqrt{a_2(1-a_2)} + v_2 - a_2 u_2}{\sqrt{1-a_2}} \right) \cdot \left(\frac{v_2 - u_2}{\sqrt{1-a_2}} \right)} \right) \times \\ \times \left[\Phi \left(\frac{y a_1 \sqrt{1-a_2} + a_2(v_1 - a_2 u_1)}{\sqrt{a_1 a_2(a_2 - a_1)}} \right) - e^{-2 \left(\frac{v_1 - a_2 u_1}{\sqrt{a_2}} \right) \left(\frac{y \sqrt{a_2(1-a_2)} + v_1 - a_2 u_1}{\sqrt{a_2}} \right)} \right] \times \\ \times \Phi \left(\frac{y a_1 \sqrt{1-a_2} + (2a_1 - a_2)(v_1 - a_2 u_1)}{\sqrt{a_1 a_2(a_2 - a_1)}} \right) dy. \end{aligned}$$

Bizonyítás. Alkalmazzuk az 1.5. Tételt és a szokásos határátmeneti eljárást. A részleteket mellőzzük. Speciális esetben ($a_1 = a_2 = 1/2$) MALMQUIST [67] adott meg ekvivalens formulát.

2.4.1. KOROLLÁRIUM.

(2.17)

$$\begin{aligned} \lim_{n \rightarrow \infty} P \left(\sup_{\substack{x \leq a_1 \\ a_2 \leq x}} (x - F_n(x)) < \frac{t}{\sqrt{n}} \right) = \\ = \frac{1}{\sqrt{2\pi}} \int_{\frac{t}{\sqrt{a_2(1-a_2)}}}^{\infty} e^{-\frac{y^2}{2}} \left(1 - e^{-2 \frac{t(y \sqrt{a_2(1-a_2)} + t)}{1-a_2}} \right) \times \\ \times \left[\Phi \left(\frac{y a_1 \sqrt{1-a_2} + t a_2}{\sqrt{a_1 a_2(a_2 - a_1)}} \right) - e^{-2 \frac{t(y \sqrt{a_2(1-a_2)} + t)}{a_2}} \Phi \left(\frac{y a_1 \sqrt{1-a_2} + t(2a_1 - a_2)}{\sqrt{a_1 a_2(a_2 - a_1)}} \right) \right] dy. \end{aligned}$$

Bizonyítás. Helyettesítsünk a 2.4. Tétel (2.16) képletébe $u_1 = u_2 = 0, v_1 = v_2 = t$.

2.4.2. KOROLLÁRIUM.

$$\begin{aligned} (2.18) \quad \lim_{n \rightarrow \infty} P \left(\max \left(\sup_{0 \leq x \leq 1/2} \left(\frac{x - F_n(x)}{1-x} \right), \sup_{1/2 \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) \right) < \frac{t}{\sqrt{n}} \right) = \\ = \lim_{n \rightarrow \infty} P \left(\max \left(\sup_{0 \leq x \leq 1/2} \left(\frac{F_n(x) - x}{1-x} \right), \sup_{1/2 \leq x \leq 1} \left(\frac{F_n(x) - x}{x} \right) \right) < \frac{t}{\sqrt{n}} \right) = \\ = 3\Phi(t) - 2 + e^{4t^2} \Phi(-3t), \quad t > 0. \end{aligned}$$

Bizonyítás. Helyettesítsünk a 2.4. Tétel (2.16) képletébe $a_1=a_2=1/2$, $u_1=v_1=t$, $u_2=-t$, $v_2=0$ -t. Kapjuk a következő integrált:

$$\begin{aligned} & \frac{1}{\sqrt{2\pi}} \int_{-t}^{\infty} e^{-\frac{y^2}{2}} (1 - e^{-2t(y+t)})(1 - e^{-2t(y+t)}) dy = \\ &= \frac{1}{\sqrt{2\pi}} \int_{-t}^{\infty} e^{-\frac{y^2}{2}} dy - 2 \frac{1}{\sqrt{2\pi}} \int_{-t}^{\infty} e^{-\frac{(y+2t)^2}{2}} dy + e^{4t^2} \frac{1}{\sqrt{2\pi}} \int_{-t}^{\infty} e^{-\frac{(y+4t)^2}{2}} dy = \\ &= 1 - \Phi(-t) - 2(1 - \Phi(t)) + e^{4t^2}(1 - \Phi(3t)), \end{aligned}$$

ahonnan adódik az eredmény.

A (2.18) formula egyébként MALMQUIST [67] eredményéből következik.

2.3. Nem folytonos alapeloszlás esete

Tekintsük most a nem folytonos alapeloszlás esetét. Mint említettük, a határeloszlást SCHMID [85] határozta meg először. Lásd még CARNAL [18]. Az alábbi eredményünk azonban inkább VINCZE [104] kétmintás eredményével analóg.

Tekintsük az 1.4. pontban adott $F(x)$ eloszlásfüggvényt.

2.5. TÉTEL.

$$\begin{aligned} (2.19) \quad & \lim_{n \rightarrow \infty} P \left(\sup_{-\infty < x < \infty} (F(x) - F_n(x)) < \frac{t}{\sqrt{n}} \right) = \\ &= \frac{1}{(2\pi)^r \sqrt{p_{r+1}}} \int \dots \int_T e^{-\frac{1}{2} \left(\sum_{i=1}^r (y_i^2 + w_i^2) - \frac{y_{r+1}^2}{p_{r+1}} \right)} \times \\ & \times \prod_{i=1}^{r+1} \left(1 - e^{-\frac{2}{p_i} \left(t + \sum_{j=1}^{i-1} y_j \sqrt{p_j} + \sum_{j=1}^{i-1} w_j \sqrt{q_j} \right) \left(t + \sum_{j=1}^i y_j \sqrt{p_j} + \sum_{j=1}^{i-1} w_j \sqrt{q_j} \right)} \right) dy_1 \dots dy_r dw_1 \dots dw_r, \end{aligned}$$

ahol

$$y_{r+1} = -\frac{1}{\sqrt{p_{r+1}}} \left(\sum_{j=1}^r y_j \sqrt{p_j} + \sum_{j=1}^r w_j \sqrt{q_j} \right)$$

és T az $(y_1, \dots, y_r, w_1, \dots, w_r)$ $2r$ -dimenziós tér azon tartománya, melyre

$$\begin{aligned} t + \sum_{j=1}^{i-1} y_j \sqrt{p_j} + \sum_{j=1}^{i-1} w_j \sqrt{q_j} &> 0 \\ i &= 1, 2, \dots, r. \\ t + \sum_{j=1}^i y_j \sqrt{p_j} + \sum_{j=1}^{i-1} w_j \sqrt{q_j} &> 0 \end{aligned}$$

Bizonyítás. A határeloszlást az 1.7. Tétel (1.29) formulájából is megkaphatnánk, ekkor azonban a 2.1. Lemmát $K < 1$ -re kellene alkalmaznunk. Egyszerűbbnek látszik az (1.30) formulából kiindulni.

Az $n \rightarrow \infty$ határátmenetnél legyen

$$\gamma_i = np_i + y_i \sqrt{np_i}, \quad i = 1, 2, \dots, r+1$$

$$\delta_i = nq_i + w_i \sqrt{nq_i}, \quad i = 1, 2, \dots, r$$

$$j = \gamma_i z.$$

Ekkor a polinomiális valószínűség többváltozós normális sűrűségfüggvénybe, az $1 - \sum_{(j)} \dots$ pedig (rögzített i -re) a szokásos gondolatmenetet (2.2. és 2.1. Lemmát $K=1$ -gyel) alkalmazva az

$$1 - e^{-\frac{2}{p_i}} \left(t + \sum_{j=1}^{i-1} y_j \sqrt{p_j} + \sum_{j=1}^{i-1} w_j \sqrt{q_j} \right) \left(t + \sum_{j=1}^i y_j \sqrt{p_j} + \sum_{j=1}^{i-1} w_j \sqrt{q_j} \right)$$

függvénybe, az összeg pedig integrálba megy át.

2.4. A „nem standard” esetekre vonatkozó határeloszlás tételek

A Chang-féle $\inf_{0 < F_n(x) \leq c/n} \frac{F_n(x)}{x}$ statisztika határeloszlása a következő (lásd CHANG [20], ill. RÉNYI [79]):

$$\lim_{n \rightarrow \infty} P \left(\inf_{0 < F_n(x) \leq (c/n)} \frac{F_n(x)}{x} < \varepsilon \right) = e^{-\frac{1}{\varepsilon}} + \sum_{j=1}^c \frac{(j-1)^{j-1} \left(\frac{1}{\varepsilon} \right)^j}{j!} e^{-\frac{j}{\varepsilon}}.$$

Itt lehet c rögzített, vagy n -től függően ∞ -hez tarthat. Ez utóbbi esetben az összegzés felső határa természetesen ∞ .

Másrészt DWASS [48] bizonyította a következővel ekvivalens formulát:

$$\begin{aligned} \lim_{n \rightarrow \infty} P \left(F_n(x) > cx + 1 - c - \frac{d}{n}, 0 \leq x \leq 1 \right) &= \\ &= \left(1 - \frac{1}{c} \right) \sum_{i=0}^{[d]} \frac{1}{i!} \left(\frac{i-d}{c} \right)^i e^{-\frac{d-i}{c}}, \end{aligned}$$

ahol $c > 1, d > 0$ rögzített.

Mint látjuk, ezek a határeloszlások merőben különböznek az előző „standard” határeloszlásoktól, melyek mind $\sqrt{n}$ -rendűek.

Ezzel kapcsolatban felmerül az a kérdés, hogy milyen esetekben kapunk nem-triviális határértéket a $P(F_n(x) > G_n(x), 0 \leq x \leq 1)$ valószínűségre.

A Glivenko—Cantelli-tételből nyilvánvaló, hogy $G_n(x)$ nem lehet túl távol x -től. Ha pl. $G_n(x) \leq x - \varepsilon$ minden x -re elég nagy n esetén, ε rögzített, akkor a fenti valószínűség nyilván 1-hez tart.

Ha $G_n(x) \rightarrow x, \frac{1}{\sqrt{n}}$ rendben, a megfelelő intervallum pontjaiban, akkor kapjuk a „standard” határeloszlástételeket.

A Chang-féle, ill. Dwass-féle határeloszlásokat abban az esetben kapjuk, ha $G_n(x)$ csak bizonyos x -re, vagy x -ekre tart x -hez $\left(\frac{1}{n}\right)$ rendben), de egyébként a $\max_{(x)} (x - G_n(x))$ -nek pozitív alsó korlátja van.

További vizsgálatok lehetségesek arra az esetre is, mikor $x - G_n(x)$ nagyságrendje $\frac{1}{\sqrt{n}}$ és $\frac{1}{n}$ között van. Tudomásunk szerint az ilyen közbenső esetekkel az irodalomban egyáltalán nem foglalkoztak.

Foglalkozunk először a $P(F_n(x) > cx - b, 0 \leq x \leq a)$ valószínűség különböző határértékeivel, azaz az 1.2. Tételben szereplő formula limeszével.

Szükségünk lesz a következő lemmára:

2.3. LEMMA. Legyen $B > 0, A > 0$, akkor

$$(2.20) \quad \frac{1}{\sqrt{2\pi}} \int_0^\infty \frac{B}{z\sqrt{z}} e^{-\frac{(Az+B)^2}{2z}} dz = e^{-2AB}.$$

Bizonyítás. Visszavezethető a 2.1. Lemma (2.2) formulájára a $t = \frac{z}{1+z}$ helyettesítéssel.

2.6. TÉTEL. (i) Legyen $b_n > 0, b_n \sqrt{n} \rightarrow 0, b_n n \rightarrow \infty$ és $nb_n(1-c_n) \rightarrow w > 0$, midőn $n \rightarrow \infty$, ekkor

$$(2.21) \quad \lim_{n \rightarrow \infty} P(F_n(x) > c_n x - b_n, 0 \leq x \leq a) = 1 - e^{-2w}.$$

(ii) Ha $b_n = \frac{v}{n} > 0, 0 < c < 1$ rögzített, akkor

$$(2.22) \quad \lim_{n \rightarrow \infty} P(F_n(x) > cx - b_n, 0 \leq x \leq a) = 1 - v \sum_{j=0}^{\infty} \frac{(v+j)^{j-1}}{c^j j!} e^{-\frac{v+j}{c}}.$$

Bizonyítás. Alkalmazzuk az 1.2. Tétel (1.3) formuláját és végezzünk határátmenetet a tétel feltételeinek megfelelően és (i)-nél legyen $j = \frac{nb_n}{1-c_n} z$. Ekkor a de Moivre—Laplace-tétel értelmében és némi számolással adódik:

$$\begin{aligned} & \binom{n}{j} \left(\frac{b_n}{c_n} + \frac{j}{nc_n} \right)^j \left(1 - \frac{b_n}{c_n} - \frac{j}{nc_n} \right)^{n-j} \sim \\ & \sim \frac{1}{\sqrt{2\pi n \left(\frac{b_n}{c_n} + \frac{j}{nc_n} \right)}} e^{-\frac{\left(n \left(\frac{b_n}{c_n} + \frac{j}{nc_n} \right) - j \right)^2}{2n \left(\frac{b_n}{c_n} + \frac{j}{nc_n} \right)}} \sim \frac{1}{\sqrt{2\pi n \frac{b_n}{1-c_n} z}} e^{-\frac{(nb_n + (1-c_n)j)^2}{2j}} \sim \\ & \sim \frac{1}{\sqrt{2\pi n \frac{b_n}{1-c_n} z}} e^{-\frac{(\sqrt{w} + z\sqrt{w})^2}{2z}}, \end{aligned}$$

így az (1.3) formulában szereplő összeg $\Delta z = \frac{1-c_n}{b_n}$ lévén, az alábbi integrál közelítő összege lesz:

$$\int_0^\infty \frac{\sqrt{w}}{\sqrt{2\pi z} \sqrt{z}} e^{-\frac{(\sqrt{w}+z\sqrt{w})^2}{2z}} dz,$$

ami a 2.3. Lemma értelmében e^{-2w} .

(ii) esetben rögzített j -re $\binom{n}{j} \sim \frac{n^j}{j!}$, $\left(1 - \frac{b}{c} - \frac{j}{nc}\right)^{n-j} \sim e^{-\frac{v+j}{c}}$, így adódik (2.22).

Mint látjuk, az eredmény a -tól független. Ennek így is kell lenni, hiszen a $(cx-b)$ egyenes olyan, hogy $F_n(x)$ -nek az $(a, 1)$ intervallumban való viselkedése az átmetszés szempontjából már nem sokat számít. Éppen így a következő tételnél a

$$P(F_n(x) > c_n x - b_n, a \leq x \leq 1)$$

valószínűség határértéke független lesz a -tól, ezért egyszerűség kedvéért $a=0$ -t válasszunk.

2.7. TÉTEL. (i) Legyen $b_n > 0$, $c_n > 1$, $1 - c_n + b_n > 0$, $\sqrt{n}(1 - c_n + b_n) \rightarrow 0$, $n(1 - c_n + b_n) \rightarrow \infty$ és $n(1 - c_n + b_n)(c_n - 1) \rightarrow w > 0$ midőn $n \rightarrow \infty$, ekkor

$$(2.23) \quad \lim_{n \rightarrow \infty} P(F_n(x) > c_n x - b_n, 0 \leq x \leq 1) = 1 - e^{-2w}.$$

(ii) Ha $1 - c_n + b_n = \frac{v}{n} > 0$, $b > 0$ rögzített, akkor

$$(2.24) \quad \lim_{n \rightarrow \infty} P(F_n(x) > c_n x - b, 0 \leq x \leq 1) = 1 - \frac{b}{1+b} \sum_{k \geq v} \frac{1}{k!} \left(\frac{k-v}{1+b} \right)^k e^{\frac{v-k}{1+b}}.$$

Bizonyítás. Hasonló, mint a 2.6. Tételé, csak j és $n-j$ szerepét fel kell cserélni. Így (i) esetben $n-j = n \frac{1-c_n+b_n}{c_n-1}$ z -t választva a 2.6. Tételnél követett gondolatmenettel adódik (2.23), míg (ii) esetben $n-j = k$ -t rögzítjük, és tagonként végzünk határártmenetet.

E tétel (ii) része analóg DWASS [48] eredményével.

A továbbiakban vizsgáljuk a $\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right)$ statisztikát, mely $a=0$ esetén értelmetlen. Tekintsünk azonban egy a_n sorozatot, amely 0-hoz tart. Ekkor a következő határeloszlásokat kapjuk:

2.8. TÉTEL. Legyen $a_n > 0$ és $a_n \rightarrow 0$ midőn $n \rightarrow \infty$.

(i) Ha $\lim_{n \rightarrow \infty} n a_n = \infty$, akkor

$$(2.25) \quad \lim_{n \rightarrow \infty} P \left(\sup_{a_n \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \frac{t}{\sqrt{n a_n}} \right) = \Phi(t) - \Phi(-t), \quad t > 0.$$

(ii) Ha $a_n = \frac{v}{n}$, $\varepsilon > 0$ rögzített, akkor

$$(2.26) \quad \lim_{n \rightarrow \infty} P \left(\sup_{a_n \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \varepsilon \right) = \\ = \sum_{\alpha \geq v(1-\varepsilon)} \frac{v^\alpha}{\alpha!} e^{-v} \left(1 - \left(\frac{\alpha}{1-\varepsilon} - v \right) \sum_{j=0}^{\infty} \frac{\left(\frac{\alpha+j}{1-\varepsilon} - v \right)^{j-1}}{j!} e^{-\frac{\alpha+j}{1-\varepsilon} + v} \right).$$

Bizonyítás. Kiindulunk a

$$P \left(\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \varepsilon \right) = \sum_{\alpha > n(1-\varepsilon)a} \binom{n}{\alpha} a^\alpha (1-a)^{n-\alpha} \times \\ \times \left(1 - \left(\frac{\alpha}{n(1-\varepsilon)} - a \right)^{[n(1-\varepsilon)]-\alpha} \sum_{j=0}^{n-\alpha} \binom{n-\alpha}{j} \left(\frac{\frac{\alpha+j}{n(1-\varepsilon)} - a}{1-a} \right)^{j-1} \left(\frac{1 - \frac{\alpha+j}{n(1-\varepsilon)}}{1-a} \right)^{n-\alpha-j} \right)$$

képletből (lásd 1.4.3. Korollárium).

(i) esetben az $\binom{n}{\alpha} a^\alpha (1-a)^{n-\alpha}$ binomiális valószínűség $\frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$ -nal aszimptotikusan egyenlő, ahol $\alpha = na_n + y\sqrt{na_n}$. A j -re vonatkozó összeg pedig $\varepsilon = \frac{t}{\sqrt{na_n}}$,

$= na_n z$ -t behelyettesítve, a *de Moivre—Laplace* lokális tételt alkalmazva, a következő integrál közelítő összege lesz:

$$\int_0^\infty \frac{1}{\sqrt{2\pi}} \frac{y+t}{z\sqrt{z}} e^{-\frac{(y+t+zt)^2}{2z}} dz,$$

ami a 2.3. Lemma értelmében $e^{-2t(y+t)}$ -vel egyenlő. Így

$$\lim_{n \rightarrow \infty} P \left(\sup_{a_n \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \frac{t}{\sqrt{na_n}} \right) = \\ = \int_{-t}^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} (1 - e^{-2t(y+t)}) dy = \Phi(t) - \Phi(-t).$$

Az (ii) esetben a határátmenetet rögzített α -ra, ill. j -re tagonként végezve adódik (2.26).

Hasonlóképpen adódik $\sup_{0 \leq x \leq a_n} \left(\frac{x - F_n(x)}{1-x} \right)$ határeloszlása, midőn $n \rightarrow \infty$, $a_n \rightarrow 1$.

2.9. TÉTEL. Legyen $0 < a_n < 1$ és $\lim_{n \rightarrow \infty} a_n = 1$.

(i) Ha $\lim_{n \rightarrow \infty} n(1-a_n) = \infty$, akkor

$$(2.27) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a_n} \left(\frac{x - F_n(x)}{1-x} \right) < \frac{t}{\sqrt{n(1-a_n)}} \right) = \Phi(t) - \Phi(-t), \quad t > 0.$$

(ii) Ha $1-a_n = \frac{v}{n} > 0$, $v > 0$, rögzített, akkor

$$(2.28) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a_n} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon \right) = 1 - \frac{\varepsilon}{1+\varepsilon} \sum_{k=\lfloor v(1+\varepsilon) \rfloor}^{\infty} \frac{k^k}{(1+\varepsilon)^k k!} e^{-\frac{k}{1+\varepsilon}}.$$

Bizonyítás. Kiindulunk a

$$\begin{aligned} & P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon \right) = \\ &= 1 - \frac{\varepsilon}{1+\varepsilon} \sum_{j=0}^{\lfloor n(a(1+\varepsilon)-\varepsilon) \rfloor} \binom{n}{j} \left(\frac{j}{n(1+\varepsilon)} + \frac{\varepsilon}{1+\varepsilon} \right)^{j-1} \left(1 - \frac{j}{n(1+\varepsilon)} - \frac{\varepsilon}{1+\varepsilon} \right)^{n-j} \end{aligned}$$

formulából (lásd 1.4.2. Korollárium). A j -re vonatkozó összegzésben $n-j=k$ -t helyettesítünk, így

$$P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon \right) = 1 - \frac{\varepsilon}{1+\varepsilon} \sum_{k=\lfloor n(1-a)(1+\varepsilon) \rfloor}^n \binom{n}{k} \left(\frac{k}{n(1+\varepsilon)} \right)^k \left(1 - \frac{k}{n(1+\varepsilon)} \right)^{n-k-1}.$$

Az (i) esetben legyen $\varepsilon = \frac{t}{\sqrt{n(1-a_n)}}$, $k = n(1-a_n)z$ és alkalmazzuk — mint szokásos — a *de Moivre—Laplace* lokális tételt. Könnyen látható, hogy az összeg a következő integrál közelítő összege:

$$\int_1^{\infty} \frac{t}{\sqrt{2\pi z}} e^{-\frac{zt^2}{2}} dz = \int_1^{\infty} \frac{2}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du$$

(az $u = t\sqrt{z}$ helyettesítéssel), amiből adódik a tétel állítása az (i) esetben.

A (ii) esetben a határátmenet az összegben tagonként végezhető.

A határérték tételekkel kapcsolatban még egy problémát vizsgáljunk. Legyen a rögzített, és tekintsük a $P_n = P \left(\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon_n \right)$ valószínűséget, mely $\varepsilon_n \rightarrow 0$ esetén 0-hoz tart (lásd 1.4.2. Korollárium). A következő tételben vizsgáljuk P_n nagyságrendjét.

2.10. TÉTEL.

(i) Ha $\varepsilon_n = \varepsilon > 0$ rögzített, akkor

$$(2.29) \quad \lim_{n \rightarrow \infty} P_n = \frac{\varepsilon}{1+\varepsilon}.$$

(ii) Ha $\varepsilon_n > 0$, $\varepsilon_n \rightarrow 0$, de $\sqrt{n\varepsilon_n} \rightarrow \infty$, midőn $n \rightarrow \infty$, akkor

$$(2.30) \quad \lim_{n \rightarrow \infty} \frac{P_n}{\varepsilon_n} = 1.$$

(iii) Ha $\varepsilon_n = \frac{t}{\sqrt{n}}$, akkor

$$(2.31) \quad \lim_{n \rightarrow \infty} \sqrt{n} P_n = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{a}{1-a}} e^{-\frac{t^2}{2} \left(\frac{1-a}{a} \right)} + t \Phi \left(t \sqrt{\frac{1-a}{a}} \right).$$

(iv) Ha $\sqrt{n\varepsilon_n} \rightarrow 0$, midőn $n \rightarrow \infty$, akkor

$$(2.32) \quad \lim_{n \rightarrow \infty} \sqrt{n} P_n = \frac{1}{\sqrt{2\pi}} \sqrt{\frac{a}{1-a}}.$$

Bizonyítás. Az 1.4.2. Korolláriumból

$$P_n = \sum_{(\alpha)} \binom{n}{\alpha} a^\alpha (1-a)^{n-\alpha} \left(\frac{\frac{\alpha}{n(1+\varepsilon)} + \frac{\varepsilon}{1+\varepsilon} - a \right).$$

Legyen $\alpha = na + y \sqrt{na(1-a)}$, ekkor

$$\binom{n}{\alpha} a^\alpha (1-a)^{n-\alpha} \sim \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy$$

és

$$\frac{\frac{\alpha}{n(1+\varepsilon)} + \frac{\varepsilon}{1+\varepsilon} - a}{1-a} = \frac{y \sqrt{na(1-a)} + n\varepsilon(1-a)}{n(1+\varepsilon)(1-a)},$$

így

$$P_n \sim \int_A \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \frac{y \sqrt{na(1-a)} + n\varepsilon(1-a)}{n(1+\varepsilon)(1-a)} dy,$$

ahol

$$A = -\sqrt{\frac{1-a}{a}} \lim_{n \rightarrow \infty} (\varepsilon_n \sqrt{n}).$$

Az (i) és (ii) esetben $A = -\infty$, így kapjuk

$$P_n \sim \frac{\varepsilon_n}{1+\varepsilon_n}.$$

A (iii) esetben $A = -t \sqrt{\frac{1-a}{a}}$, így

$$\begin{aligned} P_n &\sim \frac{1}{\sqrt{n}} \int_{-t\sqrt{\frac{1-a}{a}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \frac{y\sqrt{a(1-a)} + t(1-a)}{1-a} dy = \\ &= \frac{1}{\sqrt{n}} \left(\frac{1}{\sqrt{2\pi}} \sqrt{\frac{a}{1-a}} e^{-\frac{t^2}{2}\left(\frac{1-a}{a}\right)} + t\Phi\left(t\sqrt{\frac{1-a}{a}}\right) \right), \end{aligned}$$

míg a (iv) eset a (iii)-nak felel meg $t=0$ -val.

Megjegyezzük még, hogy $a=0$ -ra mind a négy esetben $\lim_{n \rightarrow \infty} \frac{P_n}{\varepsilon_n} = 1$ adódik.

Végül megjegyezzük, hogy a Pyke-féle empirikus eloszlásfüggvényre (1.5. pont) vonatkozó határeloszlás tételek megegyeznek a közönséges empirikus eloszlásfüggvényre vonatkozó határeloszlás tételekkel, hiszen a kétféle empirikus eloszlásfüggvény $\frac{1}{n^2}$ rendben különbözik egymástól.

3.§. ERŐFÜGGVÉNYEK VIZSGÁLATA

3.1. Pontos eloszlások alternatívák esetén

3.1.1. Bevezetés

Ebben a paragrafusban erőfüggvényeket vizsgálunk, szakaszonként lineáris alternatívára, azaz VINCZE [103] módszerét követjük. Az erőfüggvény meg van határozva, ha megadjuk az egyes statisztikák eloszlását az alternatíva esetén.

Tekintettel arra, hogy a Kolmogorov—Szmirnov-féle statisztikák monoton statisztikák (lásd CHAPMAN [22]), tetszőleges alternatíva esetére az erőfüggvény alulról és felülről becsülhető az ún. minimum és maximum alternatívákhoz tartozó erőfüggvények segítségével. Ezen szélső alternatívák szakaszonként lineárisak, így tehát vizsgálatainkból ezek a becslések is megkaphatók. (Lásd még BIRNBAUM [11], CHAPMAN [22], MASSEY [69], ROSENBLATT [81]). Normális eloszlásra lásd VAN DER WAERDEN [101], KNOTT [63], míg exponenciális eloszlásra lásd DURBIN [46].

A jelen paragrafusban a véges eloszlásokból levezethető aszimptotikus formulákat is adunk meg az erőfüggvényre. Egyéb aszimptotikus vizsgálatokra vonatkozóan lásd ABRAHAMSON [1], ANDĚL [4], QUADE [75] stb.

Egyszerűség kedvéért csupán a következő statisztikákkal foglalkozunk:

- (i) $\sup_{0 \leq x \leq 1} (x - F_n(x))$
- (ii) $\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right), \quad 0 < a < 1$
- (iii) $\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right), \quad 0 < a < 1.$

A többi statisztikára az erőfüggvény hasonlóképpen vizsgálható.

A null-hipotézisünk mindig

$$(3.1) \quad H_0: F(x) = \begin{cases} 0 & \text{ha } x < 0, \\ x & \text{ha } 0 \leq x \leq 1 \\ 1 & \text{ha } 1 < x \end{cases}$$

és az alternatíva:

$$H_1: F(x) = G(x),$$

melyre $G(0)=0$, $G(1+)=1$ és $G(x)$ szakaszonként lineáris, azaz

$$(3.2) \quad G(x) = \begin{cases} 0, & x \leq 0 \\ c_i x + b_i, & a_{i-1} < x \leq a_i, \quad i = 1, \dots, k \\ 1, & 1 < x \end{cases}$$

ahol $a_0=0 < a_1 < \dots < a_k=1$.

Ha $G(x)$ folytonos, akkor ezen alternatíva esetén $G(X_1), G(X_2), \dots, G(X_n)$ a $(0, 1)$ -ben egyenletes eloszlásból vett minta, melynek empirikus eloszlásfüggvénye $F_n(G^{-1}(x))$. Így az $\{F_n(x) > H(x), 0 \leq x \leq 1\}$ esemény ekvivalens az $\{F_n(G^{-1}(x)) > H(G^{-1}(x)), 0 \leq x \leq 1\}$ eseménnyel. Nyerjük tehát a következő lemmát:

3.1. LEMMA. *A $G(x)$ alternatíva esetén*

$$(3.3) \quad P(F_n(x) > H(x), 0 \leq x \leq 1) = P(\bar{F}_n(x) > H(G^{-1}(x)), 0 \leq x \leq 1)$$

ahol $\bar{F}_n(x)$ egy $(0,1)$ -ben egyenletes eloszlásból vett minta empirikus eloszlásfüggvénye.

Könnyen látható, hogy a lemma nem folytonos $G(x)$ -re is fennáll, ekkor invertálásnál a függőleges szakaszokból (ugrásokból) vízszintes, a vízszintesekből pedig függőleges szakaszok (ugrások) lesznek, azaz G^{-1} -t úgy értelmezzük, mint G tükröképét a 45° -os egyenesre.

Tegyük fel még a továbbiakban, hogy $G(x) \leq x$, tekintve, hogy egyoldali statisztikákat vizsgálunk, ezért az alternatíva is legyen egyoldali.

3.1.2. *A $\sup_{0 \leq x \leq 1} (x - F_n(x))$ statisztika*

3.1. TÉTEL. *Legyen $G(x)$ a (3.2) formulával adott.*

$$(3.4) \quad P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | H_1\right) =$$

$$= n! \sum_{\alpha_1, \dots, \alpha_k} \prod_{i=1}^k \frac{(a_i - a_{i-1})^{\alpha_i} c_i^{\alpha_i}}{\alpha_i!} p_n \left(a_i - a_{i-1}, \frac{\alpha_1 + \dots + \alpha_{i-1}}{n} + \varepsilon - a_{i-1}; \alpha_i \right),$$

ahol $p_n(\cdot)$ definícióját lásd (1.12) és az összegzés azon $\alpha_1, \alpha_2, \dots, \alpha_k$ nem-negatív egész számokra terjed ki, melyekre

$$\alpha_1 + \alpha_2 + \dots + \alpha_k = n,$$

$$\alpha_1 + \alpha_2 + \dots + \alpha_i \geq n \left(\frac{c_i a_i + b_i - b_{i+1}}{c_{i+1}} - \varepsilon \right), \quad i = 1, 2, \dots, k-1.$$

Bizonyítás. A 3.1. Lemma szerint

$$P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | H_1\right) = P(F_n(x) > x - \varepsilon, 0 < x < 1 | H_1) = \\ = P(\bar{F}_n(x) > G^{-1}(x) - \varepsilon, 0 < x < 1),$$

ahol

$$G^{-1}(x) = \frac{x}{c_i} - \frac{b_i}{c_i}, \quad c_i a_{i-1} + b_i \leq x < c_i a_i + b_i.$$

Az 1.3. Tételt alkalmazva adódik az állítás.

Vizsgáljunk most speciális alternatívákat.

Rögzítsük a $\sup_{0 \leq x \leq 1} (x - G(x)) = \Delta > 0$ -t és legyen x_0 az a pont, melyre $x_0 - G(x_0) = \Delta$. Legyen továbbá

$$(3.5) \quad G_1(x) = \begin{cases} x, & 0 \leq x \leq x_0 - \Delta \\ x_0 - \Delta, & x_0 - \Delta \leq x \leq x_0 \\ x, & x_0 < x \leq 1 \end{cases}$$

$$(3.6) \quad G_2(x) = \begin{cases} x - \Delta, & \Delta \leq x \leq 1 \\ 1, & 1 < x. \end{cases}$$

Ekkor igaz az, hogy $G_2(x) \leq G(x) \leq G_1(x)$, és rögzített Δ , x_0 esetén az erőfüggvény $G_1(x)$ -re minimális, $G_2(x)$ -re maximális. Ezért $G_1(x)$ az ún. „minimum alternatíva”, míg $G_2(x)$ az ún. „maximum alternatíva”. (CHAPMAN [22]).

Tekintsük még a két egyenes szakaszból álló alternatívát (VINCZE [103]):

$$(3.7) \quad G_3(x) = \begin{cases} \frac{x_0 - \Delta}{x_0} x, & 0 \leq x \leq x_0 \\ \frac{1 - x_0 + \Delta}{1 - x_0} x - \frac{\Delta}{1 - x_0}, & x_0 \leq x \leq 1, \end{cases}$$

továbbá a $(\Delta, 1)$ intervallumban egyenletes eloszlást:

$$(3.8) \quad G_4(x) = \frac{x - \Delta}{1 - \Delta}, \quad \Delta \leq x \leq 1.$$

Ezekre igazak a következő eloszlástételek:

3.1.1. KOROLLÁRIUM.

$$(3.9) \quad P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | G_1(x)\right) = \sum_{\alpha \geq n(x_0 - \varepsilon)} \binom{n}{\alpha} (x_0 - \Delta)^\alpha (1 - x_0 + \Delta)^{n - \alpha} \times \\ \times \left(1 - \frac{\varepsilon}{x_0 - \Delta} \sum_{j=0}^{[n(x_0 - \Delta - \varepsilon)]} \binom{\alpha}{j} \left(\frac{j}{x_0 - \Delta} + \varepsilon\right)^{j-1} \left(1 - \frac{j}{x_0 - \Delta} + \varepsilon\right)^{\alpha-j}\right) \times \\ \times \left(1 - \frac{\frac{\alpha}{n} + \varepsilon}{1 - x_0 + \Delta} \sum_{i=0}^{[n(1 - \varepsilon)] - \alpha} \binom{n - \alpha}{i} \left(\frac{\frac{\alpha}{n} + i}{1 - x_0 + \Delta} + \varepsilon\right)^{i-1} \left(1 - \frac{\frac{\alpha}{n} + i}{1 - x_0 + \Delta} + \varepsilon\right)^{n - \alpha - i}\right).$$

Bizonyítás. A 3.1. Lemmát alkalmazva

$$P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | G_1(x)\right) = P(F_n(x) > x - \varepsilon, 0 \leq x \leq 1 | G_1(x)) =$$

$$P(\bar{F}_n(x) > x - \varepsilon, 0 \leq x \leq x_0 - \Delta, x_0 \leq x \leq 1; \bar{F}_n(x) > x_0 - \varepsilon, x_0 - \Delta \leq x \leq x_0).$$

Ez utóbbi valószínűség az 1.5. Tételnél alkalmazott gondolatmenettel adódik.

Legyen ui. v a $(0, x_0 - \Delta)$ intervallumba eső pontok száma. A $\{v = \alpha\}$ esemény valószínűsége $\binom{n}{\alpha} (x_0 - \Delta)^\alpha (1 - x_0 + \Delta)^{n-\alpha}$ és nyilván $\frac{\alpha}{n} \geq x_0 - \varepsilon$ kell legyen. A $\{v = \alpha\}$ feltétel mellett az $\{\bar{F}_n(x) > x - \varepsilon, 0 \leq x \leq x_0 - \Delta\}$ és az $\{\bar{F}_n(x) > x - \varepsilon, x_0 \leq x \leq 1\}$ események függetlenek, valószínűségeik meghatározására az 1. § módszerét alkalmazva, adódik az állítás.

3.1.2. KOROLLÁRIUM.

$$(3.10) \quad P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | G_2(x)\right) = \\ = 1 - (\varepsilon - \Delta) \sum_{j=0}^{[n(1-\varepsilon)]} \binom{n}{j} \left(\varepsilon - \Delta + \frac{j}{n}\right)^{j-1} \left(1 - \varepsilon + \Delta - \frac{j}{n}\right)^{n-j}, \quad \varepsilon > \Delta.$$

Bizonyítás. A 3.1. Lemma szerint

$$P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | G_2(x)\right) = \\ = (P(F_n(x) > x - \varepsilon, 0 \leq x \leq 1 | G_2(x)) = P(\bar{F}_n(x) > x - \varepsilon + \Delta, 0 \leq x \leq 1 - \Delta),$$

ez utóbbi valószínűség pedig azonnal adódik az 1.2. tételből.

3.1.3. KOROLLÁRIUM.

$$(3.11) \quad P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | G_3(x)\right) = \sum_{\alpha \geq n(x_0 - \varepsilon)} \binom{n}{\alpha} (x_0 - \Delta)^\alpha (1 - x_0 + \Delta)^{n-\alpha} \times \\ \times \left(1 - \frac{\varepsilon}{x_0} \sum_{j=0}^{[n(x_0 - \varepsilon)]} \binom{\alpha}{j} \left(\left(\frac{j}{n} + \varepsilon\right) x_0\right)^{j-1} \left(1 - \left(\frac{j}{n} + \varepsilon\right) x_0\right)^{n-j}\right) \times \\ \times \left(1 - \frac{\frac{\alpha}{n} + \varepsilon - x_0}{1 - x_0} \sum_{i=0}^{[n(1-\varepsilon)] - \alpha} \binom{n-\alpha}{i} \left(\frac{\frac{\alpha+i}{n} + \varepsilon - x_0}{1 - x_0}\right)^{i-1} \left(1 - \frac{\frac{\alpha+i}{n} + \varepsilon}{1 - x_0}\right)^{n-\alpha-i}\right).$$

Bizonyítás. A 3.1. Lemma értelmében

$$P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | G_3(x)\right) = P(F_n(x) > x - \varepsilon, 0 \leq x \leq 1 | G_3(x)) = \\ = P(\bar{F}_n(x) > G_3^{-1}(x) - \varepsilon, 0 \leq x \leq 1),$$

ahol

$$G_3^{-1}(x) = \begin{cases} \frac{x_0}{x_0 - \Delta} x, & 0 \leq x \leq x_0 - \Delta \\ \frac{1 - x_0}{1 - x_0 + \Delta} x + \frac{\Delta}{1 - x_0 + \Delta}, & x_0 - \Delta \leq x \leq 1, \end{cases}$$

ez a valószínűség pedig az 1.5. Tételből azonnal adódik.

3.1.4. KOROLLÁRIUM.

$$(3.12) \quad P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | G_4(x)\right) = \\ = 1 - \frac{\varepsilon - \Delta}{1 - \Delta} \sum_{j=0}^{[n(1-\varepsilon)]} \binom{n}{j} \left(\frac{j - \Delta + \varepsilon}{1 - \Delta}\right)^{j-1} \left(\frac{1 - \frac{j}{n} - \varepsilon}{1 - \Delta}\right)^{n-j}, \quad \varepsilon > \Delta.$$

Bizonyítás. A 3.1. Lemma szerint

$$P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | G_4(x)\right) = P(\bar{F}_n(x) > (1 - \Delta)x + \Delta - \varepsilon, 0 \leq x \leq 1),$$

ez pedig az 1.2. Tételből azonnal adódik.

$$3.1.3. \text{ A } \sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1 - x}\right) \text{ statisztika}$$

A $G(x)$ függvény definíciójánál (3.2) legyen most $a_{k-1} = a$.

3.2. TÉTEL.

$$(3.13) \quad P\left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1 - x}\right) < \varepsilon | H_1\right) = n! \sum_{\alpha_1, \dots, \alpha_k} \frac{(1-a)^{\alpha_k} c_k^{\alpha_k}}{\alpha_k!} \times \\ \times \prod_{i=1}^{k-1} \frac{(a_i - a_{i-1})^{\alpha_i} c_i^{\alpha_i}}{\alpha_i!} p_n\left((1+\varepsilon)(a_i - a_{i-1}), \frac{\alpha_1 + \dots + \alpha_{i-1}}{n} + \varepsilon - (1+\varepsilon)a_{i+1}; \alpha_i\right),$$

ahol $p_n(\cdot)$ definícióját lásd (1.12), és az összegzés azon $\alpha_1, \alpha_2, \dots, \alpha_k$ nem-negatív egész számokra terjed ki, melyekre

$$\alpha_1 + \alpha_2 + \dots + \alpha_k = n,$$

$$\alpha_1 + \alpha_2 + \dots + \alpha_i \leq n \left(\frac{1+\varepsilon}{c_{i+1}} (c_i a_i + b_i - b_{i+1}) - \varepsilon \right) \quad i = 1, 2, \dots, k-1.$$

Bizonyítás. A 3.1. Lemma szerint

$$P\left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1 - x}\right) < \varepsilon | H_1\right) = P(F_n(x) > (1+\varepsilon)x - \varepsilon, 0 \leq x \leq a | H_1) = \\ = P(\bar{F}_n(x) > (1+\varepsilon)G^{-1}(x) - \varepsilon, 0 \leq x \leq c_{k-1}a + b_{k-1}).$$

Ez a valószínűség pedig az 1.3. Tételből adódik.

Itt is vizsgáljunk speciális alternatívákat. A minimum alternatíva most

$$(3.14) \quad G_b(x) = \begin{cases} x, & 0 \leq x \leq x_0 - \Delta(1 - x_0) \\ x_0 - \Delta(1 - x_0), & x_0 - \Delta(1 - x_0) < x \leq x_0 \\ x, & x_0 < x \leq 1, \end{cases}$$

a maximum alternatíva pedig

$$(3.15) \quad G_6(x) = (1+\Delta)x - \Delta, \quad \frac{\Delta}{1+\Delta} \leq x \leq 1.$$

Legyen továbbá

$$G_7(x) = \begin{cases} \left(1 + \Delta - \frac{\Delta}{x_0}\right)x, & 0 \leq x \leq x_0 \\ (1+\Delta)x - \Delta, & x_0 \leq x \leq 1 \end{cases}$$

és

$$G_8(x) = \begin{cases} x - \Delta(1-a), & \Delta(1-a) \leq x \leq 1 \\ 1, & 1 < x, \end{cases}$$

ahol $\Delta \leq \frac{a}{1-a}$.

Igazak a következő eloszlástételek:

3.2.1. KOROLLÁRIUM.

$$(3.16) \quad P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon | G_5(x) \right) =$$

$$= \sum_{\alpha \equiv n(x_0 - \varepsilon(1-x_0))} \binom{n}{\alpha} (x_0 - \Delta(1-x_0))^\alpha (1-x_0)^{n-\alpha} (1+\Delta)^{n-\alpha} \left(1 - \frac{\varepsilon}{(1+\varepsilon)(x_0 - \Delta(1-x_0))} \right) \times$$

$$\times \sum_{j=0}^{[n((1+\varepsilon)(x_0 - \Delta(1-x_0)) - \varepsilon)]} \binom{\alpha}{j} \left(\frac{\frac{j}{n} + \varepsilon}{(1+\varepsilon)(x_0 - \Delta(1-x_0))} \right)^{j-1} \left(1 - \frac{\frac{j}{n} + \varepsilon}{(1+\varepsilon)(x_0 - \Delta(1-x_0))} \right)^{\alpha-j} \times$$

$$\times \left(1 - \frac{\frac{\alpha}{n} + \varepsilon}{(1+\varepsilon)(1+\Delta)(1-x_0)} \sum_{i=0}^{[n((1+\varepsilon)a - \varepsilon)] - \alpha} \binom{n-\alpha}{i} \left(\frac{\frac{\alpha+i}{n} + \varepsilon}{(1+\varepsilon)(1+\Delta)(1-x_0)} \right)^{i-1} \times \right.$$

$$\left. \times \left(1 - \frac{\frac{\alpha+i}{n} + \varepsilon}{(1+\varepsilon)(1+\Delta)(1-x_0)} \right)^{n-\alpha-i} \right).$$

Bizonyítás. A 3.1. lemmából

$$P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon | G_5(x) \right) =$$

$$= P \left(\begin{array}{l} \bar{F}_n(x) > (1+\varepsilon)x - \varepsilon, \quad 0 \leq x \leq x_0 - \Delta(1-x_0), \quad x_0 \leq x \leq a; \\ \bar{F}_n(x) > (1+\varepsilon)x_0 - \varepsilon, \quad x_0 - \Delta(1-x_0) \leq x \leq x_0. \end{array} \right).$$

Ezt a valószínűséget a 3.1.1. Korolláriumhoz hasonlóan az 1.5. Tétel gondolatmenetével kaphatjuk meg.

3.2.2. KOROLLÁRIUM.

$$\begin{aligned}
 (3.17) \quad & P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon | G_6(x) \right) = \\
 & = 1 - \sum_{j=0}^{[n(a(1+\varepsilon)-\varepsilon)]} \binom{n}{j} \left(\frac{\frac{j}{n}(1+\Delta)+\varepsilon-\Delta}{1+\varepsilon} \right)^{j-1} \left(1 - \frac{\frac{j}{n}(1+\Delta)+\varepsilon-\Delta}{1+\varepsilon} \right)^{n-j} \frac{\varepsilon-\Delta}{1+\varepsilon}, \quad \varepsilon > \Delta.
 \end{aligned}$$

Bizonyítás. A 3.1. Lemma szerint

$$\begin{aligned}
 & P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon | G_6(x) \right) = \\
 & = P \left(\bar{F}_n(x) > x \frac{1+\varepsilon}{1+\Delta} - \frac{\varepsilon-\Delta}{1+\Delta}, \quad 0 \leq x \leq a(1+\Delta)-\Delta \right).
 \end{aligned}$$

Ez a valószínűség az 1.2. Tételből adódik.

3.2.3. KOROLLÁRIUM.

$$\begin{aligned}
 (3.18) \quad & P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon | G_7(x) \right) = \\
 & = \sum_{\alpha \geq n((1+\varepsilon)x_0-\varepsilon)} \binom{n}{\alpha} ((1+\Delta)x_0-\Delta)^\alpha (1-x_0)^{n-\alpha} (1+\Delta)^{n-\alpha} \times \\
 & \times \left(1 - \frac{\varepsilon}{(1+\varepsilon)x_0} \sum_{j=0}^{[n((1+\varepsilon)x_0-\varepsilon)]} \binom{\alpha}{j} \left(\frac{\frac{j}{n}+\varepsilon}{(1+\varepsilon)x_0} \right)^{j-1} \left(1 - \frac{\frac{j}{n}+\varepsilon}{(1+\varepsilon)x_0} \right)^{\alpha-j} \right) \times \\
 & \times \left(1 - \frac{\frac{\alpha}{n}+\frac{\varepsilon-\Delta}{1+\Delta}}{(1+\varepsilon)(1-x_0)} \sum_{i=0}^{[n((1+\varepsilon)a-\varepsilon)]-\alpha} \binom{n-\alpha}{i} \left(\frac{\frac{\alpha+i}{n}+\frac{\varepsilon-\Delta}{1+\Delta}}{(1+\varepsilon)(1-x_0)} \right)^{i-1} \left(1 - \frac{\frac{\alpha+i}{n}+\frac{\varepsilon-\Delta}{1+\Delta}}{(1+\varepsilon)(1-x_0)} \right)^{n-\alpha-i} \right).
 \end{aligned}$$

Bizonyítás. A 3.1. Lemma szerint

$$\begin{aligned}
 & P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon | G_7(x) \right) = \\
 & = P \left(\begin{aligned} & \bar{F}_n(x) > \frac{(1+\varepsilon)x_0}{(1+\Delta)x_0-\Delta} x - \varepsilon, \quad 0 \leq x \leq (1+\Delta)x_0-\Delta; \\ & \bar{F}_n(x) > \frac{1+\varepsilon}{1+\Delta} x - \frac{\varepsilon-\Delta}{1+\Delta}, \quad (1+\Delta)x_0-\Delta \leq x \leq (1+\Delta)a-\Delta \end{aligned} \right).
 \end{aligned}$$

Ez a valószínűség az 1.5. Tételnél alkalmazott gondolatmenettel adódik.

3.2.4. KOROLLÁRIUM. Ha $\Delta \leq \frac{a}{1-a}$, akkor

$$(3.19) \quad P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon | G_\delta(x) \right) = \\ = 1 - \left(\frac{\varepsilon}{1+\varepsilon} - \Delta(1-a) \right)^{\sum_{j=0}^{[n((1+\varepsilon)a-\varepsilon)]} \binom{n}{j} \left(\frac{\frac{j}{n} + \varepsilon}{1+\varepsilon} - \Delta(1-a) \right)^{j-1}} \times \\ \times \left(1 - \frac{\frac{j}{n} + \varepsilon}{1+\varepsilon} + \Delta(1-a) \right)^{n-j}, \quad \varepsilon > \frac{\Delta(1-a)}{1-\Delta(1-a)}.$$

Bizonyítás. A 3.1. Lemma szerint

$$P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \varepsilon | G_\delta(x) \right) = \\ = P(\bar{F}_n(x) > (1+\varepsilon)(x + \Delta(1-a)) - \varepsilon, 0 \leq x \leq a(1+\Delta) - \Delta).$$

Ez a valószínűség az 1.2. Tételből közvetlenül adódik.

3.1.4. $A \sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right)$ statisztika

Legyen most $a_1 = a$.

3.3. TÉTEL.

$$(3.20) \quad P \left(\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \varepsilon | H_1 \right) = n! \sum_{\alpha_1, \dots, \alpha_k} \dots \sum \frac{a^{\alpha_1} c_1^{\alpha_1}}{\alpha_1!} \times \\ \times \prod_{i=2}^k \frac{(a_i - a_{i-1})^{\alpha_i} c_i^{\alpha_i}}{\alpha_i!} p_n \left((a_i - a_{i-1})(1-\varepsilon), \frac{\alpha_1 + \dots + \alpha_{i-1}}{n} - (1-\varepsilon)a_{i-1}; \alpha_i \right),$$

ahol $p_n(\cdot)$ definícióját lásd (1.12), és az összegzés azon $\alpha_1, \alpha_2, \dots, \alpha_k$ nem-negatív egész számokra terjed ki, melyekre

$$\alpha_1 + \alpha_2 + \dots + \alpha_k = n,$$

$$\alpha_1 + \alpha_2 + \dots + \alpha_i \geq n \frac{1-\varepsilon}{c_{i+1}} (c_i a_i + b_i - b_{i+1}), \quad i = 1, 2, \dots, k.$$

Bizonyítás. A 3.1. Lemma értelmében

$$P \left(\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \varepsilon | H_1 \right) = P(F_n(x) > (1-\varepsilon)x, a \leq x \leq 1 | H_1) = \\ = P(\bar{F}_n(x) > (1-\varepsilon)G^{-1}(x), c_1 a + b_1 \leq x \leq 1).$$

Ez a valószínűség az 1.3. Tételből adódik.

Tekintsünk most is speciális alternatívákat, de egyszerűség kedvéért csak a szélső alternatívákkal foglalkozunk.

A minimum alternatíva:

$$G_9(x) = \begin{cases} x, & a \leq x \leq x_0(1-\Delta) \\ x_0(1-\Delta), & x_0(1-\Delta) \leq x \leq x_0 \\ x, & x_0 < x \leq 1, \end{cases}$$

a maximum alternatíva:

$$G_{10}(x) = \begin{cases} x(1-\Delta) & 0 \leq x \leq 1 \\ 1 & 1 < x. \end{cases}$$

Ezekre igazak a következő eloszlástételek:

3.3.1. KOROLLÁRIUM.

$$\begin{aligned} (3.21) \quad & P \left(\sup_{a \leq x \leq 1} \left| \frac{x - F_n(x)}{x} \right| < \varepsilon | G_9(x) \right) = \\ & = \sum_{\substack{\alpha_1 > (1-\varepsilon)a \\ \alpha_1 + \alpha_2 > (1-\varepsilon)x_0}} \dots \sum_{\alpha_2} \frac{n!}{\alpha_1! \alpha_2! (n - \alpha_1 - \alpha_2)!} a^{\alpha_1} (x_0(1-\Delta) - a)^{\alpha_2} (1 - x_0(1-\Delta))^{n - \alpha_1 - \alpha_2} \times \\ & \times \left(1 - \frac{\frac{\alpha_1}{n(1-\varepsilon)} - a}{x_0(1-\Delta) - a} \sum_{j=0}^{[n(1-\varepsilon)(1-\Delta)x_0] - \alpha_1} \binom{\alpha_2}{j} \left(\frac{\frac{\alpha_1 + j}{n(1-\varepsilon)} - a}{x_0(1-\Delta) - a} \right)^{j-1} \left(1 - \frac{\frac{\alpha_1 + j}{n(1-\varepsilon)} - a}{x_0(1-\Delta) - a} \right)^{\alpha_2 - j} \right) \times \\ & \times \left(1 - \frac{\frac{\alpha_1 + \alpha_2}{n(1-\varepsilon)} - x_0(1-\Delta)}{1 - x_0(1-\Delta)} \sum_{i=0}^{[n(1-\varepsilon)] - \alpha_1 - \alpha_2} \binom{n - \alpha_1 - \alpha_2}{i} \left(\frac{\frac{\alpha_1 + \alpha_2 + i}{n(1-\varepsilon)} - x_0(1-\Delta)}{1 - x_0(1-\Delta)} \right)^{i-1} \right) \times \\ & \times \left(\frac{1 - \frac{\alpha_1 + \alpha_2 + i}{n(1-\varepsilon)}}{1 - x_0(1-\Delta)} \right)^{n - \alpha_1 - \alpha_2 - i}. \end{aligned}$$

Bizonyítás. A 3.1. Lemmából

$$\begin{aligned} & P \left(\sup_{a \leq x \leq 1} \left| \frac{x - F_n(x)}{x} \right| < \varepsilon | G_9(x) \right) = \\ & = P \left(\bar{F}_n(x) > (1-\varepsilon)x, a \leq x \leq x_0(1-\Delta), x_0 \leq x \leq 1; \right. \\ & \quad \left. \bar{F}_n(x) > (1-\varepsilon)x_0, x_0(1-\Delta) \leq x \leq x_0. \right) \end{aligned}$$

Ezt a valószínűséget pedig az 1.3. Tételből kaphatjuk meg.

3.3.2. KOROLLÁRIUM

$$\begin{aligned}
 (3.22) \quad & P \left(\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \varepsilon | G_{10}(x) \right) = \\
 & = \sum_{\alpha > a(1-\varepsilon)} \binom{n}{\alpha} (a(1-\Delta))^{\alpha} (1 - a(1-\Delta))^{n-\alpha} \times \\
 & \times \left(1 - \frac{\frac{\alpha}{n(1-\varepsilon)} - a}{1-a} \sum_{j=0}^{[n(1-\varepsilon)]-\alpha} \binom{n-\alpha}{j} \left(\frac{\frac{\alpha+j}{n(1-\varepsilon)} - a}{1-a} \right)^{j-1} \left(\frac{1 - \frac{\alpha+j}{n(1-\varepsilon)}}{1-a} \right)^{n-\alpha-j} \right).
 \end{aligned}$$

Bizonyítás. A 3.1. Lemma szerint

$$\begin{aligned}
 & P \left(\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \varepsilon | G_{10}(x) \right) = \\
 & = P \left(\bar{F}_n(x) > \frac{1-\varepsilon}{1-\Delta} x, a(1-\Delta) \leq x \leq 1-\Delta \right).
 \end{aligned}$$

Ez a valószínűség az 1.4.3. Korolláriumból adódik.

3.2. Erőfüggvények aszimptotikus vizsgálata

3.2.1. Bevezetés

Ismeretes (lásd pl. MASSEY [69]), hogy az egyoldali *Kolmogorov—Szmirnov*-próba konzisztens minden rögzített alternatívára, melyre $\Delta > 0$. Hasonlóképpen belátható, hogy az általunk vizsgált további statisztikán $\left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right), \sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right), \text{ stb.} \right)$ alapuló próbák is konzisztensek minden rögzített alternatívára, melyre $\Delta > 0$. Belátható viszont, hogy a $\sup_{0 \leq x \leq 1} \left(\frac{x - F_n(x)}{1-x} \right)$ statisztikán alapuló próba nem konzisztens. Ennek oka az, hogy a statisztika az előzőekkel ellentétben még nullhipotézis esetén sem tart sztochasztikusan 0-hoz, hiszen (lásd 1.1. Tétel vagy 1.4.2. Korollárium)

$$(3.23) \quad P \left(\sup_{0 \leq x \leq 1} \left(\frac{x - F_n(x)}{1-x} \right) > \varepsilon \right) = \frac{1}{1+\varepsilon},$$

n -től függetlenül. Így rögzítve az $\frac{1}{1+\varepsilon} = \gamma$ elsőfajú hibát és a $G_6(x) = (1+\Delta)x - \Delta$ $\frac{\Delta}{1+\Delta} \leq x \leq 1$ alternatívát (amely pedig maximum alternatíva), a 3.1. Lemma alapján

látható, hogy

$$(3.24) \quad P \left(\sup_{0 \leq x \leq 1} \left(\frac{x - F_n(x)}{1 - x} \right) > \varepsilon | G_\delta(x) \right) = \\ = P \left(\bar{F}_n(x) > x \frac{1 + \varepsilon}{1 + \Delta} - \frac{\varepsilon - \Delta}{1 + \Delta}, 0 \leq x \leq 1 \right) = \frac{\varepsilon - \Delta}{1 + \varepsilon} = 1 - \gamma(1 + \Delta),$$

azaz az erőfüggvény nem tart 1-hez, midőn $n \rightarrow \infty$. Itt fennáll tehát az a paradox helyzet, hogy ha a $\sup \left(\frac{x - F_n(x)}{1 - x} \right)$ statisztikán alapuló próbát akarunk konstruálni, az alternatívák egy igen nagy osztályára jobb próbát kapunk, ha ezt nem az egész $(0, 1)$ intervallumon, hanem csak egy részén, a $(0, a)$ intervallumon vizsgáljuk, azaz információt veszünk.

A $\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1 - x} \right)$ statisztikán alapuló próba azon $G(x)$ alternatívákra nem lesz konzisztens, melyekre $G(x) = x$, a $[0, a]$ intervallumban és csak az $(a, 1]$ intervallumban $G(x) \neq x$. Könnyen tudunk azonban minden rögzített $G(x) \leq x$, $G(x) \neq x$ alternatívára konzisztens próbát konstruálni, ha $a = a_n$ -et veszünk úgy, hogy $\lim_{n \rightarrow \infty} a_n = 1$. A 2.9. Tételből következik ui., hogy ha még $\lim_{n \rightarrow \infty} n(1 - a_n) = \infty$, akkor minden rögzített $\varepsilon > 0$ -ra H_0 esetén

$$(3.25) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a_n} \left(\frac{x - F_n(x)}{1 - x} \right) < \varepsilon \right) = 0,$$

így a $\sup_{0 \leq x \leq a_n} \left(\frac{x - F_n(x)}{1 - x} \right)$ statisztika nullhipotézis esetén sztochasztikusan 0-hoz, alternatíva esetén $\sup_{0 \leq x \leq 1} \left(\frac{x - G(x)}{1 - x} \right) = \Delta > 0$ -hoz tart, amiből már (lásd FRASER [53]) a konzisztencia következik.

Hasonlóképpen a $\sup_{a_n \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right)$ statisztikán alapuló próba minden rögzített $G(x) \leq x$, $G(x) \neq x$ alternatívára konzisztens, ahol $0 < a_n < 1$,

$$\lim_{n \rightarrow \infty} a_n = 0, \quad \text{de} \quad \lim_{n \rightarrow \infty} n a_n = \infty.$$

E rövid kitérő után azonban térjünk vissza rögzített a -ra és vizsgáljuk a tárgyalt statisztikák határeloszlását alternatíva esetén, ill. az efficiencia különböző értelmezéseit a mi esetünkben.

3.2.2. Határeloszlás tételek alternatíva esetén

RAGHAVACHARI [76] megadta a Kolmogorov—Szmirnov-típusú statisztikák határeloszlását rögzített alternatívára, Wiener-folyamat bizonyos funkcionáljainak eloszlásával. Mi explicite megadjuk a határeloszlásokat az általunk vizsgált statisztikákra és alternatívákra.

3.4. TÉTEL.

$$(3.26) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq 1} (x - F_n(x)) - \Delta < \frac{t}{\sqrt{n}} \middle| G_1(x) \right) = \\ = \Phi \left(\frac{t}{\sqrt{(x_0 - \Delta)(1 - x_0 + \Delta)}} \right), \quad i = 1, 3.$$

$$(3.27) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq 1} (x - F_n(x)) - \Delta < \frac{t}{\sqrt{n}} \middle| G_2(x) \right) = \\ = \Phi \left(\frac{t}{\sqrt{\Delta(1 - \Delta)}} \right) - e^{-2t^2} \Phi \left(\frac{(1 - 2\Delta)t}{\sqrt{\Delta(1 - \Delta)}} \right), \quad t > 0.$$

$$(3.28) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq 1} (x - F_n(x)) - \Delta < \frac{v}{n} \middle| G_4(x) \right) = \\ = 1 - v \sum_{j=0}^{\infty} \frac{(v+j)^{j-1}}{(1-\Delta)^j j!} e^{-\frac{v+j}{1-\Delta}}, \quad v > 0.$$

Bizonyítás. A $G_1(x)$ és $G_3(x)$ alternatívánál a $\sup (x - G(x)) = \Delta$ csak az $x = x_0$ pontban éretik el, ezért azonos a határeloszlás.

A 3.1.1. és 3.1.3. Korolláriumból látható ui., hogy mindkét esetben a domináló tag a $\sum \binom{n}{\alpha} (x_0 - \Delta)^\alpha (1 - x_0 + \Delta)^{n-\alpha}$ binomiális valószínűség, amely — mint ismeretes — a normális eloszlásfüggvényhez tart. A zárójelben levő többi összeg — a 2. § módszerét alkalmazva (lásd 2.4. pont) 0-hoz tart.

$G_2(x)$ esetén a határeloszlás a 2.3.1. Korolláriumból, $a_1 = 0, a_2 = 1 - \Delta$ helyettesítéssel adódik.

$G_4(x)$ esetén a $\sup (x - G_4(x))$ az $x = \Delta$ pontban éretik el, és itt $G_4(\Delta) = 0$. Így itt a standard, $\sqrt{n}$ -rendű határátmenet 1-et ad minden rögzített t -re (lásd (3.26) $x_0 = \Delta$ -val). Ezért ez az eset a 2.6. Tétel (ii) részéből adódik $c = 1 - \Delta$ helyettesítéssel.

3.5. TÉTEL.

$$(3.29) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1 - x} \right) - \Delta < \frac{t}{\sqrt{n}} \middle| G_5(x) \right) = \Phi \left(t \sqrt{\frac{1 - x_0}{(1 + \Delta)(x_0 - \Delta(1 - x_0))}} \right).$$

$$(3.30) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1 - x} \right) - \Delta < \frac{t}{\sqrt{n}} \middle| G_6(x) \right) = \\ = \Phi \left(t \sqrt{\frac{1 - a}{(1 + \Delta)(a - \Delta(1 - a))}} \right) - \Phi \left(-t \sqrt{\frac{1 - a}{(1 + \Delta)(a - \Delta(1 - a))}} \right), \quad t > 0.$$

$$\begin{aligned}
 (3.31) \quad & \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) - \Delta < \frac{t}{\sqrt{n}} \middle| G_7(x) \right) = \\
 & = \int_{-t\sqrt{\frac{1-x_0}{(1+\Delta)(x_0-\Delta(1-x_0))}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \left[\Phi \left(\left(y+t \sqrt{\frac{1-x_0}{(1+\Delta)(x_0-\Delta(1-x_0))}} \right) \right) \times \right. \\
 & \times \left. \left[\sqrt{\frac{(1-a)(x_0-\Delta(1-x_0))}{a-x_0+\Delta(1-x_0)}} \right] - \Phi \left(- \left(y+t \sqrt{\frac{1-x_0}{(1+\Delta)(x_0-\Delta(1-x_0))}} \right) \right) \times \right. \\
 & \times \left. \left. \left[\sqrt{\frac{(1-a)(x_0-\Delta(1-x_0))}{a-x_0+\Delta(1-x_0)}} \right] \right] \right] dy.
 \end{aligned}$$

$$(3.32) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) - \Delta < \frac{t}{\sqrt{n}} \middle| G_8(x) \right) = \Phi \left(t \sqrt{\frac{1-a}{(a(1+\Delta)-\Delta)(1+\Delta)}} \right).$$

Bizonyítás. A 3.2.1. Korolláriumnál a domináló tag a $\sum \binom{n}{\alpha} (x_0 - \Delta(1-x_0))^\alpha \times (1-x_0)^{n-\alpha} (1+\Delta)^{n-\alpha}$ binomiális valószínűség. A zárójelben levő összegek ui. 0-hoz tartanak (lásd 2.4. pont). A binomiális összegekre vonatkozó standard módszereket alkalmazva kapjuk (3.29)-et.

$G_6(x)$ esetén a 3.2.2. Korollárium szerint a

$$P \left(\bar{F}_n(x) > \left(1 + \frac{t}{(1+\Delta)\sqrt{n}} \right) x - \frac{t}{(1+\Delta)\sqrt{n}}, 0 \leq x \leq a(1+\Delta) - \Delta \right)$$

valószínűség határértékét kell meghatároznunk, ami közvetlenül adódik a 2.1. Tételből.

$G_7(x)$ esetén a 3.2.3. Korolláriumnál csupán a j -re vonatkozó összegzés hanyagolható el. A

$$P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \frac{t}{\sqrt{n}} \middle| G_7(x) \right)$$

valószínűség így aszimptotikusan egyenlő a

$$P \left(\bar{F}_n(x) > \left(1 + \frac{t}{(1+\Delta)\sqrt{n}} \right) x - \frac{t}{(1+\Delta)\sqrt{n}}, (1+\Delta)x_0 - \Delta \leq x \leq (1+\Delta)a - \Delta \right)$$

valószínűséggel, amelynek határértéke közvetlenül adódik a 2.3. Tételből.

$G_8(x)$ esetén induljunk ki a 3.2.4. Korollárium képletéből, és legyen $\varepsilon = \Delta + \frac{t}{\sqrt{n}}$,
 $\frac{j}{n} = (1+\varepsilon)a - \varepsilon + \frac{z}{\sqrt{n}}$, és alkalmazzuk a 2. § standard módszerét. E szerint

$$\begin{aligned} & \left(\frac{\varepsilon}{1+\varepsilon} - \Delta(1-a) \right) \binom{n}{j} \left(\frac{\frac{j}{n} + \varepsilon}{1+\varepsilon} - \Delta(1-a) \right)^{j-1} \left(1 - \frac{\frac{j}{n} + \varepsilon}{1+\varepsilon} + \Delta(1-a) \right)^{n-j} = \\ & = \frac{\frac{\varepsilon}{1+\varepsilon} - \Delta(1-a)}{\frac{\frac{j}{n} + \varepsilon}{1+\varepsilon} - \Delta(1-a)} \binom{n}{j} \left(\frac{\frac{j}{n} + \varepsilon}{1+\varepsilon} - \Delta(1-a) \right)^j \left(1 - \frac{\frac{j}{n} + \varepsilon}{1+\varepsilon} + \Delta(1-a) \right)^{n-j} \sim \\ & \sim \frac{\Delta}{1+\Delta} \frac{1}{\sqrt{2\pi npq}} e^{-\frac{(np-j)^2}{2npq}}, \end{aligned}$$

ahol $p = \frac{\frac{j}{n} + \varepsilon}{1+\varepsilon} - \Delta(1-a)$ és $q = 1-p$.

Némi számolással adódik:

$$p \sim a - \Delta + \Delta a, \quad np - j \sim \sqrt{n} \left(t(1-a) - \frac{\Delta}{1+\Delta} z \right),$$

így

$$\frac{(np-j)^2}{npq} \sim \frac{\left(\frac{\Delta}{1+\Delta} z - t(1-a) \right)^2}{(a-\Delta+\Delta a)(1+\Delta)(1-a)}.$$

A keresett valószínűség aszimptotikus értékére adódik tehát

$$\begin{aligned} & 1 - \int_{-\infty}^0 \frac{\Delta}{1+\Delta} \frac{1}{\sqrt{2\pi(a-\Delta+\Delta a)(1+\Delta)(1-a)}} e^{-\frac{\left(\frac{\Delta}{1+\Delta} z - t(1-a) \right)^2}{2(a-\Delta+\Delta a)(1+\Delta)(1-a)}} dz = \\ & = 1 - \Phi \left(-t \sqrt{\frac{1-a}{(a-\Delta+\Delta a)(1+\Delta)}} \right) = \Phi \left(t \sqrt{\frac{1-a}{(a-\Delta+\Delta a)(1+\Delta)}} \right). \end{aligned}$$

3.6. TÉTEL.

$$(3.33) \quad \lim_{n \rightarrow \infty} P \left(\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) < \frac{t}{\sqrt{n}} \middle| G_9(x) \right) = \Phi \left(t \sqrt{\frac{x_0}{(1-\Delta)(1-x_0(1-\Delta))}} \right).$$

$$(3.34) \quad \lim_{n \rightarrow \infty} P \left(\sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right) - \Delta < \frac{t}{\sqrt{n}} \middle| G_{10}(x) \right) =$$

$$= \frac{1}{\sqrt{2\pi}} \int_{-t \sqrt{\frac{a}{(1-\Delta)(1-a(1-\Delta))}}}^{\infty} e^{-\frac{y^2}{2}} \left[\Phi \left(y \sqrt{\frac{a\Delta}{1-a}} + t \sqrt{\frac{a\Delta + 1 - a}{\Delta(1-\Delta)(1-a)}} \right) - \right.$$

$$\left. - e^{-2 \left(y \sqrt{a(1-\Delta)} + \frac{at}{\sqrt{1-a(1-\Delta)}} \right) \frac{t}{(1-\Delta)\sqrt{1-a(1-\Delta)}}} \Phi \left(-y \sqrt{\frac{a\Delta}{1-a}} + t \sqrt{\frac{a\Delta + 1 - a}{\Delta(1-\Delta)(1-a)}} \right) \right] dy.$$

Bizonyítás. (3.33) lényegében abból következik, hogy az $\frac{x - G_9(x)}{x}$ csak az $x = x_0$ pontban veszi fel a maximumát, így $G_1(x)$, $G_3(x)$, $G_5(x)$ -hez hasonlóan, $F_n(x)$ viselkedése a többi pontban elhanyagolható.

(3.34) rész a $P \left(\bar{F}_n(x) > x \left(1 - \frac{t}{(1-\Delta)\sqrt{n}} \right), a(1-\Delta) \leq x \leq 1-\Delta \right)$ valószínűség határértéke, ami a 2.3. Tételből közvetlenül adódik.

3.2.3. A Chernoff-féle efficiencia

Rögzített elsőfajú valószínűség és rögzített alternatíva esetén a másodfajú hibavalószínűség általában kb. ϱ^n , azaz exponenciálisan tart 0-hoz. ϱ természetesen függ az alkalmazott próbától és az alternatívától. Egy rögzített alternatívára a legkisebb ϱ a Neyman—Pearson (likelihood-hányados) -próba esetén adódik, hiszen az a legjobb próba. CHERNOFF [26] javasolta a $\frac{\log \varrho_1}{\log \varrho_2}$ mennyiséget efficienciának. SZANOV [90] tételéből következik azonban, hogy a Kolmogorov—Szmirnov-próbánál is elérhető a legjobb ϱ , ilyen értelemben tehát a Kolmogorov—Szmirnov-próba relatív efficienciája a Neyman—Pearson-próbára vonatkozólag 1. Érdemes idézni SZANOV említett tételét:

Legyen $F_n(x)$ az $F(x)$ eloszlásból vett minta empirikus eloszlásfüggvénye és $G(x)$ egy másik eloszlásfüggvény, melyre $\int_{-\infty}^{\infty} \log \frac{dG(x)}{dF(x)} dG(x)$ (I-divergencia) létezik és véges. Ekkor

$$(3.35) \quad P \left(\sup_{(x)} |F_n(x) - G(x)| < \varepsilon \right) = e^{n \left(- \int_{-\infty}^{\infty} \log \frac{dG(x)}{dF(x)} dG(x) + \gamma_\varepsilon + \delta_{n,\varepsilon} \right)},$$

ahol $\lim_{\varepsilon \rightarrow 0} \gamma_\varepsilon = 0$, $\lim_{\varepsilon \rightarrow 0} \lim_{n \rightarrow \infty} \delta_{n,\varepsilon} = 0$.

Egy másik cikkében CHERNOFF [25] a következő módosítást vezette be: Tartson a γ elsőfajú hiba is alkalmas módon 0-hoz, mégpedig azonos rendben mint a másodfajú hiba. Általában elérhető, hogy ez a nagyságrend exponenciális legyen, azaz

$$(3.36) \quad \lim_{n \rightarrow \infty} \sqrt[n]{\gamma} = \lim_{n \rightarrow \infty} \sqrt[n]{\beta} = \varrho,$$

valamilyen ϱ -val. (Ez a ϱ természetesen nem azonos az előző ϱ -val). CHERNOFF az $e = \frac{\log \varrho_1}{\log \varrho_2}$ mennyiséget nevezi relatív efficienciának. A Neyman—Pearson-próbára (lásd CHERNOFF [25])

$$(3.37) \quad \varrho = \inf_{0 < t < 1} \int (f_1(x))^t (f_0(x))^{1-t} dv(x)$$

ahol $v(\cdot)$ mind a nullhipotézist, mind az alternatívát domináló mérték, $f_0(x)$ és $f_1(x)$ az erre vonatkozó sűrűségfüggvény nullhipotézis, ill. alternatíva esetén.

A ϱ meghatározásához szükségünk lesz ABRAHAMSON [1] következő tételére:

Legyen $K_{\psi,n} = \sup_{0 < x < 1} ((x - F_n(x))\psi(x))$, ahol $\psi(x) \geq 0$ folytonos súlyfüggvény, melyre

$$\lim_{x \rightarrow 1^-} (1-x)\psi(x) = 0, \quad \lim_{x \rightarrow 0^+} x\psi(x) = 0.$$

Ekkor

$$(3.38) \quad \lim_{n \rightarrow \infty} \frac{1}{n} \log P(K_{\psi,n} \geq \varepsilon) = \sup_{(x)} \log \varrho \left(x, \frac{\varepsilon}{\psi(x)} \right),$$

ahol

$$\varrho(x, \varepsilon) = \left(\frac{1-x}{1-x+\varepsilon} \right)^{1-x+\varepsilon} \left(\frac{x}{x-\varepsilon} \right)^{x-\varepsilon}, \quad \varepsilon < x < 1.$$

Határozzuk meg tehát ϱ -t az általunk vizsgált esetekben. Nem mindig van azonban értelme a Chernoff-efficienciának. A $\sup_{0 \leq x \leq 1} (x - F_n(x))$ statisztika és $G_2(x)$ alternatíva esetén ui. a 3.1.2. Korolláriumból látható, hogy rögzített ε esetén

$$(3.39) \quad \beta = P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon | G_2(x)\right) = 0,$$

ha $\varepsilon < \Delta$, míg $\lim_{n \rightarrow \infty} \beta = 1$, ha $\varepsilon > \Delta$. Ugyanez a helyzet a $G_4(x)$ alternatívánál, valamint a $G_6(x)$, $G_8(x)$, $G_{10}(x)$ alternatívánál is. A többi esetet vegyük most egybe:

3.7. TÉTEL.

(i) A $\sup_{0 \leq x \leq 1} (x - F_n(x))$ statisztika, $G_1(x)$ és $G_3(x)$ alternatívák esetén

$$(3.40) \quad \varrho = \left(\frac{x_0 - \Delta}{x_0 - \varepsilon} \right)^{x_0 - \varepsilon} \left(\frac{1 - x_0 + \Delta}{1 - x_0 + \varepsilon} \right)^{1 - x_0 + \varepsilon},$$

ahol ε a

$$\sup_{\varepsilon < x < 1} \left(\left(\frac{1-x}{1-x+\varepsilon} \right)^{1-x+\varepsilon} \left(\frac{x}{x-\varepsilon} \right)^{x-\varepsilon} \right) = \left(\frac{x_0 - \Delta}{x_0 - \varepsilon} \right)^{x_0 - \varepsilon} \left(\frac{1 - x_0 + \Delta}{1 - x_0 + \varepsilon} \right)^{1 - x_0 + \varepsilon}$$

egyenlet megoldása.

(ii) $A \sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1 - x} \right)$ statisztika esetén a $G_5(x)$ alternatívánál

$$(3.41) \quad \varrho = \left(\frac{x_0 - \Delta(1 - x_0)}{x_0 - \varepsilon} \right)^{x_0 - \varepsilon} \left(\frac{(1 - x_0)(1 + \Delta)}{1 - x_0 + \varepsilon} \right)^{1 - x_0 + \varepsilon},$$

ahol ε az

$$\left(\frac{1}{1 + \varepsilon} \right)^{(1 + \varepsilon)(1 - a)} \left(\frac{a}{a(1 + \varepsilon) - \varepsilon} \right)^{a(1 + \varepsilon) - \varepsilon} = \left(\frac{x_0 - \Delta(1 - x_0)}{x_0 - \varepsilon} \right)^{x_0 - \varepsilon} \left(\frac{(1 - x_0)(1 + \Delta)}{1 - x_0 + \varepsilon} \right)^{1 - x_0 + \varepsilon}$$

egyenlet megoldása.

(iii) $A \sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right)$ statisztika esetén a $G_9(x)$ alternatívánál

$$(3.42) \quad \varrho = \left(\frac{x_0(1 - \Delta)}{x_0 - \varepsilon} \right)^{x_0 - \varepsilon} \left(\frac{1 - x_0 + \Delta x_0}{1 - x_0 + \varepsilon} \right)^{1 - x_0 + \varepsilon},$$

ahol ε az

$$\left(\frac{1}{1 - \varepsilon} \right)^{a(1 - \varepsilon)} \left(\frac{1 - a}{1 - a + a\varepsilon} \right)^{1 - a + a\varepsilon} = \left(\frac{x_0(1 - \Delta)}{x_0 - \varepsilon} \right)^{x_0 - \varepsilon} \left(\frac{1 - x_0 + \Delta x_0}{1 - x_0 + \varepsilon} \right)^{1 - x_0 + \varepsilon}$$

egyenlet megoldása.

Bizonyítás. Az (i) esetben Abrahamson idézett tétele szerint rögzített ε -ra

$$\lim_{n \rightarrow \infty} \sqrt[n]{\bar{\gamma}} = \sup_{(x)} \left(\left(\frac{1 - x}{1 - x + \varepsilon} \right)^{1 - x + \varepsilon} \left(\frac{x}{x - \varepsilon} \right)^{x - \varepsilon} \right).$$

A 3.1.1. és 3.1.3. Korolláriumból

$$\beta = P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) | G_i(x)\right) = \sum_{\alpha \geq n(x_0 - \varepsilon)} \binom{n}{\alpha} (x_0 - \Delta)^\alpha (1 - x_0 + \Delta)^{n - \alpha}, \quad i = 1, 3$$

és itt a jobboldal a pontos nagyságrendet is megadja. (lásd 3.2.2. pont). Így (lásd CHERNOFF [25])

$$\lim_{n \rightarrow \infty} \sqrt[n]{\bar{\beta}} = \left(\frac{x_0 - \Delta}{x_0 - \varepsilon} \right)^{x_0 - \varepsilon} \left(\frac{1 - x_0 + \Delta}{1 - x_0 + \varepsilon} \right)^{1 - x_0 + \varepsilon}.$$

ε -t megválasztva úgy, hogy $\lim_{n \rightarrow \infty} \sqrt[n]{\bar{\gamma}} = \lim_{n \rightarrow \infty} \sqrt[n]{\bar{\beta}}$ legyen, kapjuk a tétel állítását.

Az (ii) esetben ABRAHAMSON tétele szerint

$$\begin{aligned} \lim_{n \rightarrow \infty} \sqrt[n]{\bar{\gamma}} &= \sup_{0 \leq x \leq a} \left(\left(\frac{1}{1 + \varepsilon} \right)^{(1 + \varepsilon)(1 - x)} \left(\frac{x}{x(1 + \varepsilon) - \varepsilon} \right)^{x(1 + \varepsilon) - \varepsilon} \right) = \\ &= \left(\frac{1}{1 + \varepsilon} \right)^{(1 + \varepsilon)(1 - a)} \left(\frac{a}{a(1 + \varepsilon) - \varepsilon} \right)^{a(1 + \varepsilon) - \varepsilon} \end{aligned}$$

és (CHERNOFF [25])

$$\lim_{n \rightarrow \infty} \sqrt[n]{\beta} = \left(\frac{x_0 - \Delta(1-x_0)}{x_0 - \varepsilon} \right)^{x_0 - \varepsilon} \left(\frac{(1-x_0)(1+\Delta)}{1-x_0 + \varepsilon} \right)^{1-x_0 + \varepsilon},$$

amiből következik az állítás.

Az (iii) esetben ABRAHAMSON szerint

$$\begin{aligned} \lim_{n \rightarrow \infty} \sqrt[n]{\gamma} &= \sup_{a \leq x \leq 1} \left(\left(\frac{1-x}{1-x+\varepsilon x} \right)^{1-x+\varepsilon x} \left(\frac{1}{1-\varepsilon} \right)^{x(1-\varepsilon)} \right) = \\ &= \left(\frac{1}{1-\varepsilon} \right)^{a(1-\varepsilon)} \left(\frac{1-a}{1-a+\varepsilon a} \right)^{1-a+\varepsilon a}, \end{aligned}$$

míg (CHERNOFF [25])

$$\lim_{n \rightarrow \infty} \sqrt[n]{\beta} = \left(\frac{x_0(1-\Delta)}{x_0 - \varepsilon} \right)^{x_0 - \varepsilon} \left(\frac{1-x_0 + \Delta x_0}{1-x_0 + \varepsilon} \right)^{1-x_0 + \varepsilon},$$

amiből következik az állítás.

3.2.4. A Bahadur-féle efficiencia

A Bahadur-féle efficiencia [6] előnye a könnyebb kezelhetősége.

Legyen T_n ($n=1, 2, \dots$) statisztikák egy végtelen sorozata, melyre

(i) nullhipotézis esetén

$$(3.43) \quad \lim_{n \rightarrow \infty} P(T_n < t) = H(t),$$

folytonos eloszlásfüggvény,

$$(3.44) \quad (ii) \quad \log(1-H(t)) = -\frac{At^2}{2}(1+o(1)), \quad t \rightarrow \infty.$$

(iii) A P_θ ellenhipotézis esetén

$$(3.45) \quad P \left(\lim_{n \rightarrow \infty} \frac{T_n}{\sqrt{n}} = b(\theta) \right) = 1.$$

Ekkor a $c(\theta) = A(b(\theta))^2$ jelöléssel a Bahadur-efficiencia az

$$(3.46) \quad e(\theta) = \frac{c_1(\theta)}{c_2(\theta)}$$

mennyiség, ahol $c_1(\theta)$ és $c_2(\theta)$ a két összehasonlítandó statisztika-sorozathoz tartozó $c(\theta)$ -érték.

Határozzuk meg a továbbiakban $c(\theta)$ értékét az általunk vizsgált esetekben.

Megjegyezzük, hogy ABRAHAMSON [1] nem a fenti efficienciát vizsgálja, hanem egy módosított formáját, az ún. egzakt Bahadur-efficienciát. Ez a módosítás azonban nehezebben kezelhető, mint az eredeti definíció. A Chernoff- és Bahadur-efficiencia részletes tárgyalását lásd BAHADUR [7].

3.8. TÉTEL.

(i) $A \sup_{0 \leq x \leq 1} (x - F_n(x))$ statisztika esetén minden $\sup_{(x)} (x - G(x)) = \Delta$ alternatívára $c = 4\Delta^2$.

(ii) $A \sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1 - x} \right)$ statisztika esetén minden $\sup_{0 \leq x \leq a} \left(\frac{x - G(x)}{1 - x} \right) = \Delta$ alternatívára $c = \frac{1-a}{a} \Delta^2$.

(iii) $A \sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right)$ statisztika esetén minden $\sup_{a \leq x \leq 1} \left(\frac{x - G(x)}{x} \right) = \Delta$ alternatívára $c = \frac{a}{1-a} \Delta^2$.

Bizonyítás. Legyen (i)-nél $T_n = \sqrt{n} \sup_{0 \leq x \leq 1} (x - F_n(x))$. Ekkor nullhipotézis esetén (2.3.1. Korollárium)

$$\lim_{n \rightarrow \infty} P(T_n < t) = 1 - e^{-2t^2}, \quad t > 0,$$

így $A=4$. A Glivenko—Cantelli-tétel szerint alternatíva esetén

$$\frac{T_n}{\sqrt{n}} \rightarrow \Delta \quad 1 \text{ valószínűséggel,}$$

tehát $c = 4\Delta^2$.

(ii)-nél legyen $T_n = \sqrt{n} \sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1 - x} \right)$. Nullhipotézis esetén (2.3.2. Korollárium)

$$\lim_{n \rightarrow \infty} P(T_n < t) = \Phi \left(t \sqrt{\frac{1-a}{a}} \right) - \Phi \left(-t \sqrt{\frac{1-a}{a}} \right),$$

és tekintve, hogy $1 - \Phi(x) = \frac{1}{\sqrt{2\pi}x} e^{-\frac{x^2}{2}} (1 + o(1))$, $x \rightarrow \infty$ (lásd pl. RÉNYI [78]),

ezért $A = \frac{1-a}{a}$. Ellenhipotézis esetén a Glivenko—Cantelli-tétel szerint

$$\frac{T_n}{\sqrt{n}} \rightarrow \Delta \quad 1 \text{ valószínűséggel,}$$

így $c = \frac{1-a}{a} \Delta^2$.

(iii)-nál legyen $T_n = \sqrt{n} \sup_{a \leq x \leq 1} \left(\frac{x - F_n(x)}{x} \right)$. Nullhipotézis esetén (2.3.3. Korollárium)

$$\lim_{n \rightarrow \infty} P(T_n < t) = \Phi \left(t \sqrt{\frac{a}{1-a}} \right) - \Phi \left(-t \sqrt{\frac{a}{1-a}} \right),$$

így $A = \frac{a}{1-a}$. Ellenhipotézis esetén ugyancsak

$$\frac{T_n}{\sqrt{n}} \rightarrow A \quad 1 \text{ valószínűséggel,}$$

tehát $c = \frac{a}{1-a} \Delta^2$.

Ismerve a c értékét a fenti esetekre, meghatározhatjuk az egyes próbák *Bahadur*-efficienciáját egy másik próbához viszonyítva.

$G_1(x)$ és $G_3(x)$ alternatívák esetén pl. tekintsük az x_0 pontban alkalmazott binomiális próbát, azaz az $F_n(x_0)$ statisztikán alapuló próbát. Könnyű látni, hogy

az erre vonatkozó c -érték $c = \frac{\Delta^2}{x_0(1-x_0)}$, így a *Kolmogorov—Szmirnov*-próba efficienciája $G_1(x)$ és $G_3(x)$ esetén $e = 4x_0(1-x_0)$. Ha tehát $x_0 = 1/2$, akkor $e = 1$.

$G_2(x)$ és $G_4(x)$ alternatívák esetén a legjobb próba X_1^* -on alapul, és ez nem elégíti ki a *Bahadur*-efficiencia definíciójánál előírt feltételeket. U. i. nX_1^* -nak van határeloszlása, de alternatíva esetén $X_1^* \rightarrow \Delta$, így azt is mondhatnánk, hogy a *Bahadur*-efficiencia 0. A legjobb próba ezekben az esetekben lényegesen jobb, mint a *Kolmogorov—Szmirnov*-próba. Hasonló a helyzet $G_6(x)$, $G_8(x)$ és $G_{10}(x)$ alternatíva esetén.

$G_5(x)$ és $G_7(x)$ lényegében azonos a $G_1(x)$ és $G_3(x)$ alternatívával, csak Δ helyébe $\Delta(1-x_0)$ -at kell tenni. Így az $F_n(x_0)$ -on alapuló binomiális próbára $c = \frac{1-x_0}{x_0} \Delta^2$, tehát erre $e = \frac{(1-a)x_0}{a(1-x_0)}$. Ha $x_0 = a$, akkor $e = 1$.

Az előzőkhöz hasonlóan $G_9(x)$ -re a binomiális próbánál $c = \frac{x_0}{1-x_0} \Delta^2$, így $e = \frac{a(1-x_0)}{(1-a)x_0}$.

3.2.5. A Pitman-féle efficiencia

Az előzőkben az alternatívákat rögzítettük, abban az értelemben, hogy határátmenetnél nem függtek a mintaelemszámtól. Egy próba érzékenységét azonban tulajdonképpen az méri, hogy a nullhipotézishez közeli alternatívákat is minél jobban meg tud különböztetni. Ha rögzítjük a γ elsőfajú hibavalószínűséget, az alternatívát a nullhipotézishez tartatjuk, ($\Delta_n \rightarrow 0$, ha $n \rightarrow \infty$), az erőfüggvénynek általában valamilyen γ és 1 közé eső határértéke lesz. Erre az esetre vonatkozik az efficiencia legrégebb, PITMANTÓL származó definíciója (lásd pl. FRASER [53]).

Tekintsünk két próbát (statisztika-sorozatot) és az alternatíva függjön egy Δ_n paramétertől, melyre $\Delta_n > 0$, $\lim_{n \rightarrow \infty} \Delta_n = 0$ (nullhipotézis). Jelölje $\beta_n(\Delta)$ az egyik, $\beta_m^*(\Delta)$ a másik próbához tartozó másodfajú hibavalószínűséget, n , illetve m mintaelemszám, Δ alternatíva esetén. Tegyük fel, hogy az n_i és m_i részsorozatok olyanok, hogy

$$(3.47) \quad \lim_{i \rightarrow \infty} \beta_{n_i}(\Delta_i) = \lim_{i \rightarrow \infty} \beta_{m_i}^*(\Delta_i),$$

ahol ez a határérték 0 és $1-\gamma$ közé esik, de azoktól különböző. Ha $\lim_{i \rightarrow \infty} \frac{n_i}{m_i}$ létezik és azonos minden, a fenti feltételnek eleget tevő részsorozatra, akkor ezt nevezzük a két próba relatív efficienciájának.

Valójában a *Pitman*-féle efficiencia akkor kezelhető jól, ha a szóban forgó határeloszlások normálisak. Az erre vonatkozó tételek megtalálhatók FRASER [53] könyvében. Esetünkben azonban nem ez a helyzet.

Meg kell határoznunk tehát a másodfajú hibavalószínűség határértékét rögzített elsőfajú hibavalószínűség és $\Delta_n \rightarrow 0$ esetén. De csak akkor kapunk 0-tól és $1-\gamma$ -tól különböző határértékeket, ha $\Delta_n = \frac{w}{\sqrt{n}}$. Ennél távolabbi alternatíva esetén ui.

az erőfüggvény 1-hez tart, közelebbi alternatíva esetén pedig γ -hoz tart.

Megadjuk a határeloszlásokat az általunk vizsgált statisztikák és alternatívák esetén, ahol legyen mindenütt $\Delta = \Delta_n = \frac{w}{\sqrt{n}}$.

3.9. TÉTEL.

$$(3.48) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \frac{t}{\sqrt{n}} \middle| G_1(x) \right) = \\ = 1 - \Phi \left(\frac{t+w}{\sqrt{x_0(1-x_0)}} \right) + \Phi \left(\frac{t-w}{\sqrt{x_0(1-x_0)}} \right) - \\ - e^{-2t^2} \left(1 - \Phi \left(\frac{t(1-2x_0)+w}{\sqrt{x_0(1-x_0)}} \right) + \Phi \left(\frac{t(1-2x_0)-w}{\sqrt{x_0(1-x_0)}} \right) \right).$$

$$(3.49) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \frac{t}{\sqrt{n}} \middle| G_2(x) \right) = 1 - e^{-2(t-w)^2}, \quad w < t.$$

$$(3.50) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \frac{t}{\sqrt{n}} \middle| G_3(x) \right) = \\ = \Phi \left(\frac{t-w}{\sqrt{x_0(1-x_0)}} \right) + e^{-\frac{2tw}{x_0(1-x_0)}} \left(1 - \Phi \left(\frac{t+w}{\sqrt{x_0(1-x_0)}} \right) \right) - \\ - e^{-2t^2 + \frac{2wt}{x_0}} \Phi \left(\frac{t(2x_0-1)-w}{\sqrt{x_0(1-x_0)}} \right) - e^{-2t^2 + \frac{2wt}{1-x_0}} \Phi \left(\frac{t(1-2x_0)-w}{\sqrt{x_0(1-x_0)}} \right).$$

$$(3.51) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq 1} (x - F_n(x)) < \frac{t}{\sqrt{n}} \middle| G_4(x) \right) = 1 - e^{-2t(t-w)}, \quad w < t.$$

Bizonyítás. A 2. § módszereinek alkalmazásával kaphatjuk a fenti határeloszlásokat, a megfelelő véges eloszlásokból. A részleteket egyszerűség kedvéért elhagyjuk.

A fenti formulák megadják a másodfajú hibavalószínűségek határértékét. Ui. az elfogadási tartomány

$$\left\{ \sup_{0 \leq x \leq 1} (x - F_n(x)) < \varepsilon \right\},$$

és rögzített γ elsőfajú hiba azt jelenti, hogy $\varepsilon \sim \frac{1}{\sqrt{n}} \sqrt{\frac{\log 1/\gamma}{2}}$. (A (2.3.1) Korolláriumból ui.

$$\lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq 1} (x - F_n(x)) > \frac{\sqrt{\frac{\log 1/\gamma}{2}}}{\sqrt{n}} \right) = e^{-2 \left(\sqrt{\frac{\log 1/\gamma}{2}} \right)^2} = \gamma.$$

A másodfajú hibavalószínűségek határértékét tehát

$$t = \sqrt{\frac{\log 1/\gamma}{2}}$$

helyettesítéssel kaphatjuk.

3.10. TÉTEL.

$$\begin{aligned} (3.52) \quad & \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \frac{t}{\sqrt{n}} \middle| G_5(x) \right) = \\ & = \frac{1}{\sqrt{2\pi}} \int_{(w-t)\sqrt{\frac{1-x_0}{a-x_0}}}^{\infty} e^{-\frac{y^2}{2}} (1 - e^{-2t \left(y \sqrt{\frac{1-x_0}{a-x_0}} + t \sqrt{\frac{1-x_0}{a-x_0}} \right)}) \times \\ & \times \left(\Phi \left(\sqrt{\frac{1-a}{a-x_0}} (y \sqrt{x_0} + t \sqrt{1-x_0}) \right) - \Phi \left(-\sqrt{\frac{1-a}{a-x_0}} (y \sqrt{x_0} + t \sqrt{1-x_0}) \right) \right) dy. \end{aligned}$$

$$\begin{aligned} (3.53) \quad & \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \frac{t}{\sqrt{n}} \middle| G_6(x) \right) = \\ & = \Phi \left((t-w) \sqrt{\frac{1-a}{a}} \right) - \Phi \left(-(t-w) \sqrt{\frac{1-a}{a}} \right), \quad w < t. \end{aligned}$$

$$\begin{aligned} (3.54) \quad & \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \frac{t}{\sqrt{n}} \middle| G_7(x) \right) = \\ & = \frac{1}{\sqrt{2\pi}} \int_{(w-t)\sqrt{\frac{1-x_0}{a-x_0}}}^{\infty} e^{-\frac{y^2}{2}} (1 - e^{-2t \left(y \sqrt{\frac{1-x_0}{a-x_0}} + (t-w) \sqrt{\frac{1-x_0}{a-x_0}} \right)}) \times \\ & \times \left(\Phi \left(\sqrt{\frac{1-a}{a-x_0}} (y \sqrt{x_0} + (t+w) \sqrt{1-x_0}) \right) - \Phi \left(-\sqrt{\frac{1-a}{a-x_0}} (y \sqrt{x_0} + (t+w) \sqrt{1-x_0}) \right) \right) dy. \end{aligned}$$

$$\begin{aligned} (3.55) \quad & \lim_{n \rightarrow \infty} P \left(\sup_{0 \leq x \leq a} \left(\frac{x - F_n(x)}{1-x} \right) < \frac{t}{\sqrt{n}} \middle| G_8(x) \right) = \\ & = \Phi \left((t-w) \sqrt{\frac{1-a}{a}} \right) - e^{2w(1-a)(t-w+aw)} \Phi \left((w-t-2aw) \sqrt{\frac{1-a}{a}} \right), \quad w(1-a) < t. \end{aligned}$$

A bizonyítást itt is mellőzzük.

Megjegyezzük, hogy a másodfajú hibavalószínűségek határértékének meghatározásához most t értékét a

$$\Phi\left(t\sqrt{\frac{1-a}{a}}\right) - \Phi\left(-t\sqrt{\frac{1-a}{a}}\right) = 1 - \gamma$$

egyenlet gyökeként kell meghatározni.

3.11. TÉTEL.

$$\begin{aligned} (3.56) \quad \lim_{n \rightarrow \infty} P\left(\sup_{a \leq x \leq 1} \left|\frac{x - F_n(x)}{x}\right| < \frac{t}{\sqrt{n}} \mid G_9(x)\right) = \\ = \frac{1}{\sqrt{2\pi}} \int_{(w-t)\sqrt{\frac{x_0}{1-x_0}}}^{\infty} e^{-\frac{y^2}{2}} (1 - e^{-2t\left(y\sqrt{\frac{x_0}{1-x_0}} + t\sqrt{\frac{x_0}{1-x_0}}\right)}) \times \\ \times \left(\Phi\left(\sqrt{\frac{a}{x_0-a}}(y\sqrt{1-x_0} + t\sqrt{x_0})\right) - \Phi\left(-\sqrt{\frac{a}{x_0-a}}(y\sqrt{1-x_0} + t\sqrt{x_0})\right)\right) dy. \end{aligned}$$

$$\begin{aligned} (3.57) \quad \lim_{n \rightarrow \infty} P\left(\sup_{a \leq x \leq 1} \left|\frac{x - F_n(x)}{x}\right| < \frac{t}{\sqrt{n}} \mid G_{10}(x)\right) = \\ = \Phi\left((t-w)\sqrt{\frac{a}{1-a}}\right) - \Phi\left(-(t-w)\sqrt{\frac{a}{1-a}}\right), \quad w < t. \end{aligned}$$

A bizonyítást mellőzzük.

A másodfajú hibavalószínűségeket a fenti formulákból kaphatjuk, ahol most t a

$$\Phi\left(t\sqrt{\frac{a}{1-a}}\right) - \Phi\left(-t\sqrt{\frac{a}{1-a}}\right) = 1 - \gamma$$

egyenlet megoldása.

A fenti határértékek lehetőséget adnak a vizsgált próbák összehasonlításához a *Pitman*-efficiencia alapján — legalábbis elvileg. Gyakorlatilag azonban a formulák nehezen kezelhetők.

Megjegyezzük még, hogy nullhipotézishez $1/\sqrt{n}$ -rendben tartó alternatívák esetén erőfüggvények határértéke a *Doob*-módszerrel, *Wiener*-folyamatra vonatkozó megfelelő valószínűségek meghatározásával is kapható. Ezt az utat követte QUADB [75].

4. §. AZ ITERÁLT LOGARITMUS TÉTEL EMPIRIKUS ELOSZLÁSFÜGGVÉNYRE

4.1. Történeti áttekintés és néhány segédteétel

Az $F_n(x)$ empirikus eloszlásfüggvény binomiális eloszlású független valószínűségi változók átlaga, s mint ilyen, igaz rá a nagy számok erős törvénye, azaz

$$(4.1) \quad P(\lim_{n \rightarrow \infty} F_n(x) = x) = 1.$$

A Glivenko—Cantelli-tétel kimondja, (lásd [54]), hogy ez a konvergencia x -ben egyenletes, azaz

$$(4.2) \quad P\left(\lim_{n \rightarrow \infty} \sup_{0 \leq x \leq 1} |x - F_n(x)| = 0\right) = 1.$$

E tétel általánosítására és egyéb kiterjesztéseire vonatkozólag utalunk a [88], [100] dolgozatokra.

Független változók összegére azonban — a szórás létezésének feltételezése mellett — több is igaz, nevezetesen fennáll az iterált logaritmus tétel (lásd pl. HARTMAN—WINTNER [57]). Ebből következik, hogy

$$(4.3) \quad P\left(\lim_{n \rightarrow \infty} \left(\sqrt{\frac{n}{\log \log n}} |x - F_n(x)|\right) = \sqrt{2x(1-x)}\right) = 1.$$

SZMIRNOV [92] bizonyította, hogy az iterált logaritmus tétel fennáll az eltérés maximumára is, azaz

$$(4.4) \quad P\left(\lim_{n \rightarrow \infty} \left(\sqrt{\frac{n}{\log \log n}} \sup_{0 \leq x \leq 1} |x - F_n(x)|\right) = \frac{1}{\sqrt{2}}\right) = 1.$$

Itt az $\frac{1}{\sqrt{2}}$ konstans onnan adódik, hogy $x = \frac{1}{2}$ -re $\sqrt{2x(1-x)} = \frac{1}{\sqrt{2}}$, más x -ekre természetesen $\sqrt{2x(1-x)} < \frac{1}{\sqrt{2}}$. Lásd még CHUNG [28].

CASSELS [19] bizonyította, hogy (4.3) egyenletesen áll fenn:

$$(4.5) \quad P\left(\lim_{n \rightarrow \infty} \sup_{0 \leq x \leq 1} \left(\sqrt{\frac{n}{\log \log n}} |x - F_n(x) - \sqrt{2x(1-x)}|\right) = 0\right) = 1.$$

CASSELS valamivel többet is bizonyított, nevezetesen, hogy

(4.6)

$$P\left(\lim_{n \rightarrow \infty} \sup_{0 \leq y < x \leq 1} \left(\sqrt{\frac{n}{\log \log n}} |x - y - F_n(x) + F_n(y) - \sqrt{2(x-y)(1-x+y)}|\right) = 0\right) = 1.$$

CASSELS (4.5) eredményéből következik SZMIRNOV (4.4) eredménye, de fordítva nem.

Az iterált log tétel funkcionális (Strassen-féle) verziójára vonatkozólag utalunk FINKELSTEIN [52] és WICHURA [109] dolgozataira.

Felmerül a kérdés, hogy igaz-e az iterált log tétel az általánosabb

$$(4.7) \quad K_n = \sup_{0 \leq x \leq 1} (|x - F_n(x)| \tau(x))$$

statisztikára, ahol $\tau(x)$ alkalmas súlyfüggvény. Rögzített x -re nyilván

$$(4.8) \quad P\left(\overline{\lim}_{n \rightarrow \infty} \left(\sqrt{\frac{n}{\log \log n}} |x - F_n(x)| \tau(x) = \tau(x) \sqrt{2x(1-x)} \right) = 1,\right.$$

így

$$(4.9) \quad P\left(\overline{\lim}_{n \rightarrow \infty} \left(\sqrt{\frac{n}{\log \log n}} K_n \right) \cong A\right) = 1,$$

ahol

$$A = \sup_{0 \leq x \leq 1} (\tau(x) \sqrt{2x(1-x)}).$$

(4.5) egyszerű következményeként belátható, hogy ez az A konstans alkalmas súlyfüggvényre el is érhető. Ez nem lesz igaz azonban minden $\tau(x)$ -re, nevezetesen

$$\tau(x) = \frac{1}{\sqrt{x(1-x)}} \text{-re}$$

$$(4.10) \quad P\left(\overline{\lim}_{n \rightarrow \infty} \left(\sqrt{\frac{n}{\log \log n}} \sup_{0 < x < 1} \frac{|x - F_n(x)|}{\sqrt{x(1-x)}} \right) = \infty\right) = 1.$$

Először egy új bizonyítást adunk CASSELS (4.5) tételére a [31] dolgozat módszerével. Úgy gondoljuk, hogy ez a bizonyítás több szempontból sem érdektelen. Először is, kimondunk egy iterált log tételt szemi-martingálokra, ami önmagában is érdekes lehet és más esetre is alkalmazható. Másrészt belátjuk, hogy az $nK_n^+ = n \sup_{0 < x < 1} (x - F_n(x)) \tau(x)$ sorozat szemi-martingál. Továbbá explicit becslést adunk K_n^+ eloszlására, ill. momentum generáló függvényére. Ezek a tulajdonságok alkalmazhatók pl. a Darling—Robbins-féle [39] szekvenciális próbák konstruálására (lásd STANLEY [86]).

4.1. LEMMA. Legyen $\mathcal{F}_1 \subset \mathcal{F}_2 \subset \dots \subset \mathcal{F}_n \subset \dots$ egy növekvő σ -algebra sorozat, $\zeta_1, \zeta_2, \dots, \zeta_n, \dots$ valószínűségi változók sorozata, ζ_i legyen mérhető $\mathcal{F}_i$ -re vonatkozólag ($i=1, 2, \dots$), és

$$(4.11) \quad E(\zeta_n | \mathcal{F}_{n-1}) \cong \zeta_{n-1}, \quad n = 2, 3, \dots$$

azaz a fenti sorozat alkotson szemi-martingált a σ -algebra sorozatra vonatkozólag. Létezzen a változók $E(e^{t\zeta_n})$ momentum generáló függvénye elég nagy n -re és $0 \leq t < t_0$ -ra, továbbá

$$(4.12) \quad E(e^{t\zeta_n}) \leq (\psi(t))^n, \quad 0 \leq t < t_0,$$

ahol

$$\psi(t) = 1 + \frac{B^2}{2} t^2 + O(t^3) \quad t \rightarrow 0.$$

Ekkor

$$(4.13) \quad P\left(\overline{\lim}_{n \rightarrow \infty} \frac{\zeta_n}{\sqrt{n \log \log n}} \leq B\sqrt{2}\right) = 1.$$

Bizonyítás. Legyen $P_k = P\left(\max_{n_k < i \leq n_{k+1}} \zeta_i \geq (B\sqrt{2} + \varepsilon) \sqrt{n_k \log \log n_k}\right)$. Ekkor a Borel—Cantelli-lemma értelmében elég bizonyítani, hogy $\sum_{(k)} P_k$ konvergens egy alkalmas n_k sorozatra és minden $\varepsilon > 0$ -ra.

Az $e^{t\zeta_{n_1}}, \dots, e^{t\zeta_{n_k}}, \dots$ sorozat ugyancsak szemi-martingál (Doob [42], 295. o. 1.1. Tétel) és a Kolmogorov-egyenlőtlenséget (Doob [42] 314. o. 3.2. Tétel) és a (4.12) feltételt alkalmazva, kapjuk

$$\begin{aligned} P_k &= P\left(\max_{n_k < i \leq n_{k+1}} \zeta_i \geq (B\sqrt{2} + \varepsilon) \sqrt{n_k \log \log n_k}\right) = \\ &= P\left(\max_{n_k < i \leq n_{k+1}} e^{t\zeta_i} \geq e^{t(B\sqrt{2} + \varepsilon) \sqrt{n_k \log \log n_k}}\right) \leq \\ &\leq \frac{E(e^{t\zeta_{n_{k+1}}})}{e^{t(B\sqrt{2} + \varepsilon) \sqrt{n_k \log \log n_k}}} \leq \frac{(\psi(t))^{n_{k+1}}}{e^{t(B\sqrt{2} + \varepsilon) \sqrt{n_k \log \log n_k}}}. \end{aligned}$$

Legyen $n_k = \left\lceil \left(1 + \frac{\varepsilon}{B\sqrt{2}}\right)^k \right\rceil$ ($\lceil \cdot \rceil$: egész rész) és $t = \frac{\sqrt{2}}{B} \sqrt{1 + \frac{\varepsilon}{B\sqrt{2}}} \sqrt{\frac{\log \log n_{k+1}}{n_{k+1}}}$, ekkor

$$\begin{aligned} n_{k+1} \log \psi(t) &= n_{k+1} \log \left(1 + \frac{B^2}{2} t^2 + O(t^3)\right) = \\ &= n_{k+1} \frac{B^2 t^2}{2} + o(1) = \left(1 + \frac{\varepsilon}{B\sqrt{2}}\right) \log \log n_{k+1} + o(1) \quad (k \rightarrow \infty), \end{aligned}$$

és

$$\begin{aligned} t(B\sqrt{2} + \varepsilon) \sqrt{n_k \log \log n_k} &= \\ &= (B\sqrt{2} + \varepsilon) \frac{\sqrt{2}}{B} \sqrt{1 + \frac{\varepsilon}{B\sqrt{2}}} \sqrt{\frac{n_k}{n_{k+1}}} \sqrt{\log \log n_k \log \log n_{k+1}} = \\ &= 2 \left(1 + \frac{\varepsilon}{B\sqrt{2}}\right) \log \log n_{k+1} + o(1). \end{aligned}$$

Innen

$$P_k \leq \frac{c}{(\log n_{k+1})^{1 + \frac{\varepsilon}{B\sqrt{2}}}} \leq \frac{c}{(k+1)^{1 + \frac{\varepsilon}{B\sqrt{2}}}}.$$

Ezzel a 3.1. Lemmát bebizonyítottuk.

Megjegyezzük, hogy a (4.12) feltétel gyengíthető, elegendő ui. egy

$$(4.14) \quad E(e^{t\zeta_n}) \leq c_n(t) (\psi(t))^n$$

egyenlőtlenség teljesülésének megkövetelése, ahol a $c_n(t)$ tényező a $\sum_{(k)} P_k$ konvergenciáját nem befolyásolja. Ilyen tényező pl. $c_n(t) = (t\sqrt{n})^\kappa$, ahol κ egy pozitív konstans.

4.2. LEMMA. Legyen $X_1, X_2, \dots, X_n, \dots$ független, $(0,1)$ -ben egyenletes valószínűségi változók sorozata, $F_n(x)$ az $X_1, X_2, \dots, X_n$ változók empirikus eloszlásfüggvénye, és legyen $\mathcal{F}_n$ az a σ -algebra, melyet az $X_1, X_2, \dots, X_n$ változók generálnak. Legyen továbbá $\tau(x)$ szakaszonként folytonos és korlátos függvény. Ekkor a

$$(4.15) \quad \zeta_n = n \sup_{0 \leq x \leq 1} ((x - F_n(x))\tau(x)), \quad n = 1, 2, \dots$$

sorozat szemi-martingál az $\mathcal{F}_n$, $n=1, 2, \dots$ σ -algebra sorozatra vonatkozóan, azaz

$$(4.16) \quad E(\zeta_{n+1} | \mathcal{F}_n) \cong \zeta_n.$$

Bizonyítás. Legyen q_n a maximum hely, amely a tett feltevések mellett létezik és $\mathcal{F}_n$ -re mérhető, azaz

$$\zeta_n = n(q_n - F_n(q_n))\tau(q_n).$$

Nyilván

$$\zeta_{n+1} \cong (n+1)(q_n - F_{n+1}(q_n))\tau(q_n) = \zeta_n + q_n\tau(q_n) + (nF_n(q_n) - (n+1)F_{n+1}(q_n))\tau(q_n).$$

Ha most $X_{n+1} < q_n$, amelynek $\mathcal{F}_n$ feltétel mellett q_n a valószínűsége, akkor $(n+1)F_{n+1}(q_n) = nF_n(q_n) + 1$, és így $\zeta_{n+1} \cong \zeta_n - (1 - q_n)\tau(q_n)$.

Ha pedig $X_{n+1} > q_n$, amelynek $\mathcal{F}_n$ feltétel mellett $1 - q_n$ a valószínűsége, akkor $(n+1)F_{n+1}(q_n) = nF_n(q_n)$ és így $\zeta_{n+1} \cong \zeta_n + q_n\tau(q_n)$. Tehát

$$P(\zeta_{n+1} - \zeta_n \cong -(1 - q_n)\tau(q_n) | \mathcal{F}_n) = q_n,$$

$$P(\zeta_{n+1} - \zeta_n \cong q_n\tau(q_n) | \mathcal{F}_n) = 1 - q_n,$$

így

$$E(\zeta_{n+1} - \zeta_n | \mathcal{F}_n) \cong -q_n(1 - q_n)\tau(q_n) + (1 - q_n)q_n\tau(q_n) = 0.$$

4.2. Az iterált log tétel

4.1. TÉTEL. Legyen $\tau(x) \geq 0$, $0 \leq x \leq 1$ korlátos függvény. Legyen továbbá $X_1, X_2, \dots, X_n, \dots$ független, $(0,1)$ -ben egyenletes eloszlású valószínűségi változók sorozata, és jelölje $F_n(x)$ az $X_1, X_2, \dots, X_n$ változók empirikus eloszlásfüggvényét. Ekkor

$$(4.17) \quad P\left(\lim_{n \rightarrow \infty} \left(\sqrt{\frac{n}{\log \log n}} \sup_{0 \leq x \leq 1} (|x - F_n(x)|\tau(x))\right) = A\right) = 1,$$

ahol

$$(4.18) \quad A = \sup_{0 \leq x \leq 1} (\tau(x) \sqrt{2x(1-x)}).$$

Bizonyítás. Tekintsük először a

$$(4.19) \quad \tau_1(x) = \frac{1}{c \sqrt{x(1-x)} + b}, \quad 0 \leq x \leq 1$$

súlyfüggvényt, ahol $c \geq 0$, $b > 0$. Legyen

$$(4.20) \quad \zeta_n = n \sup_{0 \leq x \leq 1} ((x - F_n(x))\tau_1(x)).$$

A 4.2. Lemma értelmében $\{\zeta_n\}$ szemi-martingál, így a 4.1. Lemma alkalmazásához elegendő egy (4.14) egyenlőtlenséget bizonyítani.

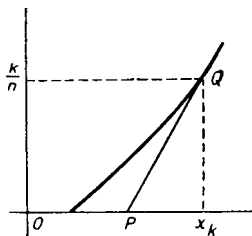
Könnyű látni, hogy

$$(4.21) \quad E(e^{t\zeta_n}) = 1 + tE(\zeta_n) + t \int (e^{tz} - 1)(1 - H(z))dz,$$

ahol

$$H(z) = P(\zeta_n < z) = P\left(F_n(x) > x - \frac{z}{n\tau_1(x)}, \quad 0 \leq x \leq 1\right),$$

tehát $1 - H(z)$ annak valószínűsége, hogy $F_n(x)$ átmetszi az $y = x - \frac{z}{n\tau_1(x)} = x - \frac{z}{n}(c\sqrt{x(1-x)} + b)$ görbét. Erre a valószínűségre egy egyszerű becslést nyerhetünk a következőképpen:



2. ábra

Átmetszés csak $\frac{k}{n}$ magasságon történhet, aminek valószínűsége

$$\binom{n}{k} x_k^k (1 - x_k)^{n-k},$$

ahol

$$x_k - \frac{z}{n\tau_1(x_k)} = \frac{k}{n}.$$

Ezen feltétel mellett annak valószínűsége, hogy $\frac{k}{n}$ előtt ne legyen átmetszés, felülről becsülhető annak valószínűségével, hogy $F_n(x)$ a $\overline{PQ}$ szakaszt ne messe, ahol $\overline{PQ}$ az $y = x - \frac{z}{n\tau_1(x)}$ görbe $Q: \left(x_k, \frac{k}{n}\right)$ pontbeli érintőjének a Q pont és az x -tengely közé eső szakasza. Ez a valószínűség az 1.1. Tétel szerint

$$(4.22) \quad \frac{\overline{OP}}{x_k} = \frac{x_k - \frac{k}{n\left(1 + \frac{z}{n} \frac{\tau_1'(x_k)}{\tau_1^2(x_k)}\right)}}{x_k} = \frac{(nx_k - k)(\tau_1(x_k) + x_k \tau_1'(x_k))}{x_k(n\tau_1(x_k) + (nx_k - k)\tau_1'(x_k))}.$$

Tehát

$$(4.23) \quad 1 - H(z) \cong \sum_{k=0}^{[n-zb]} \binom{n}{k} \frac{(nx_k - k)(\tau_1(x_k) + x_k \tau'_1(x_k))}{n\tau_1(x_k) + (nx_k - k)\tau'_1(x_k)} x_k^{k-1} (1 - x_k)^{n-k},$$

így

$$\begin{aligned} & \int_0^{n/b} (e^{tz} - 1)(1 - H(z)) dz \cong \\ & \cong \int_0^{n/b} (e^{tz} - 1) \sum_{k=0}^{[n-zb]} \binom{n}{k} \frac{(nx_k - k)(\tau_1(x_k) + x_k \tau'_1(x_k))}{n\tau_1(x_k) + (nx_k - k)\tau'_1(x_k)} x_k^{k-1} (1 - x_k)^{n-k} dz = \\ & = \sum_{k=0}^n \binom{n}{k} \int_0^{(n-k)/b} (e^{tz} - 1) \frac{(nx_k - k)(\tau_1(x_k) + x_k \tau'_1(x_k))}{n\tau_1(x_k) + (nx_k - k)\tau'_1(x_k)} x_k^{k-1} (1 - x_k)^{n-k} dz. \end{aligned}$$

Helyettesítsünk az integrálba $x_k = y \cdot t$, azaz $z = ny\tau_1(y) - k\tau_1(y) \cdot t$, ekkor $\frac{dz}{dy} = n\tau_1(y) + (ny - k)\tau'_1(y)$ és

$$\begin{aligned} & \int_0^{n/b} (e^{tz} - 1)(1 - H(z)) dz \cong \\ & \cong \sum_{k=0}^n \binom{n}{k} \int_{k/n}^1 (e^{t(ny-k)\tau_1(y)-1} - 1)(ny - k)(\tau_1(y) + y\tau'_1(y)) y^{k-1} (1 - y)^{n-k} dy. \end{aligned}$$

Tekintve, hogy az integrandus minden $0 < y < 1$ -re nem-negatív, az összeget növeljük, ha az integrált 0-tól 1-ig vesszük, majd k szerint összegezve a következőt kapjuk:

$$\begin{aligned} & \int_0^{n/b} (e^{tz} - 1)(1 - H(z)) dz \cong \sum_{k=0}^n \binom{n}{k} \int_0^1 (\dots) dy = \\ & = \int_0^1 (\tau_1(y) + y\tau'_1(y)) \sum_{k=0}^n \binom{n}{k} (e^{t(ny-k)} - 1)(ny - k)y^{k-1} (1 - y)^{n-k} dy = \\ & = \int_0^1 (\tau_1(y) + y\tau'_1(y)) n e^{tny\tau_1(y)} (1 - y)(1 - e^{-t\tau_1(y)})(ye^{-t\tau_1(y)} + (1 - y))^{n-1} dy \cong \\ & \cong n(1 - e^{-t/b}) \int_0^1 (\tau_1(y) + y\tau'_1(y))(ye^{t(y-1)\tau_1(y)} + (1 - y)e^{ty\tau_1(y)})^n dy. \end{aligned}$$

Taylor-sorfejtéssel kapjuk:

$$\begin{aligned} & ye^{t(y-1)\tau_1(y)} + (1 - y)e^{ty\tau_1(y)} = \\ & = 1 + \frac{t^2}{2} \tau_1^2(y)y(1 - y) + \frac{t^3}{6} (y - 1)^3 y \tau_1^3(y) e^{\theta(y-1)\tau_1(y)} + y^3 (1 - y) \tau_1^3(y) e^{\theta y \tau_1(y)} \cong \\ & \cong 1 + \frac{B^2 t^2}{2} + \frac{t^3}{6b^3} e^{t/b}, \quad \text{ahol} \quad B = \frac{1}{c + 2b}. \end{aligned}$$

Innen

$$\begin{aligned} & \int_0^{n/b} (e^{tz} - 1)(1 - H(z)) dz \leq \\ & \leq n(1 - e^{-t/b}) \left(1 + \frac{B^2 t^2}{2} + \frac{t^3}{6b^3} e^{t/b} \right)^n \int_0^1 (\tau_1(y) + y\tau_1'(y)) dy = \\ & = n(1 - e^{-t/b}) \left(1 + \frac{B^2 t^2}{2} + \frac{t^3}{6b^3} e^{t/b} \right)^n \frac{1}{b}. \end{aligned}$$

Végül felhasználva az $1 - e^{-t/b} \leq \frac{t}{b}$ egyenlőtlenséget, (4.21)-ből kapjuk

$$(4.24) \quad E(e^{t\zeta_n}) \leq 1 + tE(\zeta_n) + \frac{nt^2}{b^2} \left(1 + \frac{B^2 t^2}{2} + \frac{t^3}{6b^3} e^{t/b} \right)^n,$$

ami a 4.1. Lemmánál kívánt becslés, tekintve, hogy $E(\zeta_n) = O(\sqrt{n})$, így az első két tag a harmadikhoz képest elhanyagolható.

Hasonlóképpen vizsgálható $n \sup_{0 \leq x \leq 1} ((F_n(x) - x)\tau_1(x))$ (vagy azt is mondhatnánk, hogy az $x'_i = 1 - x_i$ transzformációval az előző esetre visszavezethető), így kapjuk a tétel állítását a $\tau_1(x)$ súlyfüggvényre.

Általános esetben legyen $0 \leq \tau(x) \leq K, 0 \leq x \leq 1$; igaz a következő becslés:

$$(4.25) \quad \tau(x)|x - F_n(x)| \leq \frac{|x - F_n(x)|}{\sqrt{2x(1-x)} + \frac{\varepsilon}{K}} (A + \varepsilon), \quad 0 \leq x \leq 1,$$

így a már bizonyítottak szerint

$$\begin{aligned} & \overline{\lim}_{n \rightarrow \infty} \left(\sqrt{\frac{n}{\log \log n}} \sup_{0 \leq x \leq 1} (|x - F_n(x)| \tau(x)) \right) \leq \\ & \leq (A + \varepsilon) \overline{\lim}_{n \rightarrow \infty} \left(\sqrt{\frac{n}{\log \log n}} \sup_{0 \leq x \leq 1} \frac{|x - F_n(x)|}{\sqrt{2x(1-x)} + \frac{\varepsilon}{K}} \right) = (A + \varepsilon) \left(1 + \frac{\varepsilon}{K\sqrt{2}} \right) \end{aligned}$$

és tekintve, hogy ε tetszőleges, igaz a Tétel állítása.

A 4.1. Tétel $\tau_1(x)$ -re ekvivalens (4.5)-tel, azaz CASSELS tételét adja.

A 4.1. Tétel bizonyításában fontos szerepet játszott a $K_n^+ = n \sup (x - F_n(x)) \tau_1(x)$ momentumgeneráló függvényének becslése. Ez a $P(K_n^+ > z)$ valószínűségre adott becslés segítségével történt. A momentumgeneráló függvényre adott becslés és a Markov-egyenlőtlenség segítségével viszont megbecsülhető a $P(K_n^+ > z)$ valószínűség. $\tau_1(x) = 1$ -re DVORETZKY—KIEFER—WOLFOWITZ [47] bizonyította, hogy létezik $c > 0$, hogy

$$P\left(\sup_{0 \leq x \leq 1} (x - F_n(x)) > \varepsilon\right) < ce^{-2n\varepsilon^4}, \quad 0 < \varepsilon < 1,$$

de c numerikus értékét nem adták meg. Becsléseikből az világos, hogy az $\hat{c}$ értékük igen messze van a BIRNBAUM—MCCARTY [13] által sejtett $c=1$ értéktől. Valamit javított a becslésen BUTLER—MCCARTY [17], de tudomásunk szerint a $c=1$ sejtést sem igazolni, sem megcáfolni eddig még nem sikerült. További becsléseket DARLING—ROBBINS [39], STANLEY [86], WHITTLE [108] adtak meg.

4.3. A $\tau(x) = \frac{1}{\sqrt{x(1-x)}}$ súlyfüggvény

Könnyen látható, hogy a 4.2. pont módszere nem alkalmazható a $\tau(x) = \frac{1}{\sqrt{x(1-x)}}$ ($0 < x < 1$) súlyfüggvényre, hacsak nem zárunk ki az $x=0$ és $x=1$ pont körül egy kis intervallumot, azaz tekintjük a

$$(4.26) \quad \tau(x) = \begin{cases} 0 & 0 \leq x < a_1 \\ \frac{1}{\sqrt{x(1-x)}} & a_1 \leq x \leq a_2 \\ 0 & a_2 < x \leq 1 \end{cases}$$

súlyfüggvényt. A módszerből látható, hogy a_1 és $1-a_2$ még — nem túl gyorsan — 0-hoz is tarthat. EICKER [49] és KIEFER [62] bizonyította, hogy ha $x_n \rightarrow 0$, de

$\frac{nx_n}{\log \log n} \rightarrow \infty, n \rightarrow \infty$, akkor

$$(4.27) \quad P \left(\lim_{n \rightarrow \infty} \frac{|nF_n(x_n) - nx_n|}{\sqrt{x_n(1-x_n)n \log \log n}} \leq \sqrt{2} \right) = 1.$$

Felmerül a kérdés, hogy a_n milyen választása mellett igaz még a

$$\sup_{a_n \leq x \leq 1-a_n} \left(\frac{|x - F_n(x)|}{\sqrt{x(1-x)}} \right)$$

statisztikára az iterált log tétel. KIEFER [62] cikke nagyon valószínűsíti azt a sejtést (bár eredményeiből nem következik), hogy $a_n = \frac{\log \log n}{n}$ még választható, de

$a_n = \frac{c \log \log n}{n}$ tetszőleges c -vel már nem. Pontosabban szólva, az iterált log tétel-

ben előforduló nagyságrend még jó lesz, csak a konstans függni fog c -től, és elég kis c -re a szokásos $\sqrt{2}$ -nél nagyobb lesz, sőt $c \rightarrow 0$ esetén ∞ -hez tart.

Ezt a sejtést azonban a 4.2. pont módszerével jelenleg nem tudjuk bizonyítani, de remélhetőleg az ott alkalmazott becslések finomításával már belátható.

$x_n = \frac{1}{n}$ -re még BAXTER [9] bizonyította, hogy

$$(4.28) \quad P \left(\lim_{n \rightarrow \infty} \left(\frac{n \log \log \log n}{\log \log n} F_n \left(\frac{1}{n} \right) \right) = 1 \right) = 1,$$

amiből már következik, hogy

$$(4.29) \quad P \left(\overline{\lim}_{n \rightarrow \infty} \sqrt{\frac{n}{\log \log n}} \sup_{(1/n) \leq x \leq 1 - (1/n)} \frac{|x - F_n(x)|}{\sqrt{x(1-x)}} = \infty \right) = 1.$$

KIEFER [62] szerint, ha $\sum a_n = \infty$, akkor $\min(X_1, \dots, X_n) < a_n$ végtelen sokszor következik be 1 valószínűséggel, így tehát $F_n(a_n) \cong \frac{1}{n}$ végtelen sok n -re. Ha tehát $a_n = o\left(\frac{1}{n}\right)$, de $\sum a_n = \infty$, akkor

$$\frac{|a_n - F_n(a_n)|}{\sqrt{a_n(1-a_n)}} \cong \frac{c}{n\sqrt{a_n}}$$

1 valószínűséggel végtelen sok n -re, azaz

$$\sup_{0 < x < 1} \frac{|x - F_n(x)|}{\sqrt{x(1-x)}} \cong \frac{c}{n\sqrt{a_n}}.$$

Tekintve, hogy minden a_n sorozathoz, melyre $\sum a_n = \infty$, található olyan a'_n , hogy $\sum a'_n = \infty$, de $\frac{a'_n}{a_n} \rightarrow 0$, $n \rightarrow \infty$, azért

$$P \left(\overline{\lim}_{n \rightarrow \infty} \left(n\sqrt{a_n} \sup_{0 < x < 1} \frac{|x - F_n(x)|}{\sqrt{x(1-x)}} \right) = \infty \right) = 1.$$

Ha viszont már $\sum b_n < \infty$, be fogjuk bizonyítani, hogy

$$P \left(\lim_{n \rightarrow \infty} \left(n\sqrt{b_n} \sup_{0 < x < 1} \frac{|x - F_n(x)|}{\sqrt{x(1-x)}} \right) = 0 \right) = 1.$$

Igaz lesz tehát a következő tétel:

4.2. TÉTEL.

$$(4.30) \quad \text{Ha } \sum a_n = \infty, \text{ akkor } P \left(\overline{\lim}_{n \rightarrow \infty} \left(n\sqrt{a_n} \sup_{0 < x < 1} \frac{|x - F_n(x)|}{\sqrt{x(1-x)}} \right) = \infty \right) = 1.$$

$$(4.31) \quad \text{Ha } \sum b_n < \infty, \text{ akkor } P \left(\lim_{n \rightarrow \infty} \left(n\sqrt{b_n} \sup_{0 < x < 1} \frac{|x - F_n(x)|}{\sqrt{x(1-x)}} \right) = 0 \right) = 1.$$

Bizonyítás. A tétel első része, mint fentebb láttuk, következik KIEFER [62] eredményéből. Itt csupán a második részt bizonyítjuk.

Legyen tehát $\sum b_n < \infty$. Elegendő a $\sup_{0 < x < 1/2}$ -al foglalkozni, a $\sup_{1/2 < x < 1}$ hasonlóan megy. Továbbá

$$(4.32) \quad \sup_{0 < x < 1/2} \frac{|x - F_n(x)|}{\sqrt{x(1-x)}} \leq \sqrt{2} \sup_{0 < x < 1/2} \frac{|x - F_n(x)|}{\sqrt{x}}$$

ezért elegendő ez utóbbi supremummal foglalkozni.

Legyen $U_n = \sup_{0 < x < 1/2} \left(\frac{F_n(x) - x}{\sqrt{x}} \right)$, és tekintsük az $\left\{ U_n > \frac{\varepsilon}{n\sqrt{b_n}} \right\}$ eseményt. Legyen $C_n = \left\{ n \text{ az első olyan index, hogy } U_n > \frac{\varepsilon}{n\sqrt{b_n}} \right\}$. Ekkor a *Borel—Cantelli-lemma* értelmében $P(\lim_{n \rightarrow \infty} n\sqrt{b_n} U_n = 0) = 1$ -hez elegendő bizonyítani, hogy $\sum P(C_n) < \infty$, minden $\varepsilon > 0$ -ra. Tekintve, hogy $C_n \subset \{ \text{van olyan } 0 < x < 1/2, \text{ hogy}$

$$F_n(x) > x + \frac{\varepsilon}{n\sqrt{b_n}} \sqrt{x}, \quad F_{n-1}(x) < x + \frac{\varepsilon}{(n-1)\sqrt{b_{n-1}}} \sqrt{x} \}$$

$\subset \{ \text{van olyan } 0 < x < \frac{1}{2}, \text{ hogy}$

$$nF_n(x) > nx + \frac{\varepsilon}{\sqrt{b_n}} \sqrt{x}, \quad (n-1)F_{n-1}(x) < nx + \frac{\varepsilon}{\sqrt{b_n}} \sqrt{x} \} = \sum_{k=1}^{[n/2]} B_{nk},$$

ahol

$B_{nk} = \{ \text{van olyan } x_{k-1} \leq x < x_k, \text{ hogy}$

$$nF_n(x) > nx + \frac{\varepsilon}{\sqrt{b_n}} \sqrt{x}, \quad (n-1)F_{n-1}(x) < nx + \frac{\varepsilon}{\sqrt{b_n}} \sqrt{x} \},$$

és x_k az $nx + \frac{\varepsilon}{\sqrt{b_n}} \sqrt{x} = k$ egyenlet megoldása, azaz

$$x_k = \frac{4k^2}{4nk + \frac{\varepsilon^2}{b_n} \left(1 + \sqrt{1 + \frac{4nk b_n}{\varepsilon^2}} \right)}.$$

$$(4.33) \quad P(C_n) \leq \sum_{k=1}^{[n/2]} P(B_{nk}) = \sum_{k=1}^{L_n} P(B_{nk}) + \sum_{k=L_n+1}^{[n/2]} P(B_{nk}),$$

ahol

$$L_n = \frac{\varepsilon^2}{64nb_n}.$$

A $\sum_{k=1}^{L_n} P(B_{nk})$ összeget a következőképpen becsülhetjük:

$B_{nk} = \{ \text{van olyan } x_{k-1} \leq x < x_k, \text{ hogy}$

$$(n-1)F_{n-1}(x) = k-1, \quad X_n \leq x \},$$

így

$$P(B_{nk}) \leq x_k \binom{n-1}{k-1} x_k^{k-1} = \frac{k}{n} \binom{n}{k} x_k^k \leq \frac{k}{n} \frac{(nx_k)^k}{k!}.$$

Könnyen látható, hogy $x_k < \frac{2k^2 b_n}{\varepsilon^2}$, így

$$(4.34) \quad P(B_{nk}) \leq \frac{k}{n} \left(\frac{2nk^2 b_n}{\varepsilon^2} \right)^k \frac{1}{k!} = \frac{2b_n}{\varepsilon^2} \left(\frac{2nb_n}{\varepsilon^2} \right)^{k-1} \frac{k^{2k}}{(k-1)!}.$$

Legyen

$$\alpha_k = \left(\frac{2nb_n}{\varepsilon^2} \right)^{k-1} \frac{k^{2k}}{(k-1)!},$$

akkor

$$\frac{\alpha_{k+1}}{\alpha_k} = \frac{2nb_n}{\varepsilon^2} \frac{(k+1)^{2k+2}}{k^{2k+2}} k \leq \frac{2nb_n}{\varepsilon^2} 16L_n = \frac{1}{2},$$

és így

$$\alpha_{k+1} \leq \frac{1}{2} \alpha_k \leq \left(\frac{1}{2} \right)^2 \alpha_{k-1} \leq \dots \leq \left(\frac{1}{2} \right)^k,$$

tehát

$$(4.35) \quad \sum_{k=1}^{L_n} P(B_{nk}) \leq \frac{2b_n}{\varepsilon^2} \sum_{k=1}^{L_n} \left(\frac{1}{2} \right)^{k-1} < \frac{4b_n}{\varepsilon^2}.$$

Az összeg másik részére pedig a következő becslést nyerhetjük:

$$B_{nk} \subset \{ \text{van olyan } x_{k-1} \leq x < x_k, \text{ hogy } nF_n(x) = k \},$$

és tekintve, hogy $x_k < \frac{k}{n}$, ezért belátható, hogy

$$(4.36) \quad P(B_{nk}) \leq \binom{n}{k} x_k^k (1-x_k)^{n-k}.$$

A Stirling-formula alkalmazásával kapjuk

$$(4.37) \quad \binom{n}{k} x_k^k (1-x_k)^{n-k} < \text{const } e^{-\frac{1}{2} \frac{(k-nx_k)^2}{k}}.$$

De ha $k > L_n$, akkor

$$\begin{aligned} k - nx_k &= \frac{\frac{\varepsilon^2}{b_n} \left(1 + \sqrt{1 + \frac{4nk b_n}{\varepsilon^2}} \right)}{4n + \frac{\varepsilon^2}{k b_n} \left(1 + \sqrt{1 + \frac{4nk b_n}{\varepsilon^2}} \right)} > \\ &> \frac{\frac{\varepsilon^2}{b_n} \sqrt{\frac{4nk b_n}{\varepsilon^2}}}{4n + \frac{\varepsilon^2}{L_n b_n} \left(1 + \sqrt{1 + \frac{4n L_n b_n}{\varepsilon^2}} \right)} = \frac{\varepsilon}{34 + 8\sqrt{17}} \sqrt{\frac{k}{n b_n}}, \end{aligned}$$

így

$$\frac{1}{2} \frac{(k - nx_k)^2}{k} > \frac{1}{2} \frac{\varepsilon^2}{(34 + 8\sqrt{17})^2} \frac{1}{nb_n} > 3 \log n,$$

ha n elég nagy. (Ui. $\sum b_n < \infty$ lévén, $b_n < \frac{c}{n \log n}$).

Tehát

$$(4.38) \quad P(B_{nk}) < c' e^{-3 \log n} < \frac{c'}{n^3}, \quad L_n < k \leq \frac{n}{2},$$

és így

$$\sum_{k=L_n+1}^{[n/2]} P(B_{nk}) < n \frac{c'}{n^3} = \frac{c'}{n^2},$$

ha n elég nagy. Összevetve (4.35)-tel,

$$(4.39) \quad P(C_n) < \frac{4}{\varepsilon^2} b_n + \frac{c'}{n^2},$$

azaz $\sum P(C_n) < \infty$ minden $\varepsilon > 0$ -ra.

Hasonlóképpen bizonyítható, hogy

$$(4.40) \quad P(\lim_{n \rightarrow \infty} (n \sqrt{b_n} V_n) = 0) = 1,$$

$V_n = \sup_{0 < x < 1/2} \left(\frac{x - F_n(x)}{\sqrt{x}} \right)$ -re. Valójában ez az eset még egyszerűbb, mert belátható, hogy

$$(4.41) \quad P\left(V_n > \frac{\varepsilon}{n \sqrt{b_n}}\right) \leq \sum_{k=0}^{[n/2]} \binom{n}{k} x_k^k (1-x_k)^{n-k},$$

ahol most

$$x_k = \frac{k}{n} + \frac{\varepsilon^2}{2n^2 b_n} + \frac{\varepsilon}{2n \sqrt{b_n}} \sqrt{\frac{\varepsilon^2}{n^2 b_n} + \frac{4k}{n}},$$

és itt az egész összegre alkalmazható a fenti, normális sűrűségfüggvénnyel történő becslés:

$$(4.42) \quad \binom{n}{k} x_k^k (1-x_k)^{n-k} \leq \text{const } e^{-\frac{1}{2} \frac{(nx_k - k)^2}{nx_k}}.$$

De

$$\frac{(nx_k - k)^2}{nx_k} = \frac{\left(\frac{\varepsilon^2}{2nb_n} + \frac{\varepsilon}{2\sqrt{b_n}} \sqrt{\frac{\varepsilon^2}{n^2 b_n} + \frac{4k}{n}} \right)^2}{k + \frac{\varepsilon^2}{2nb_n} + \frac{\varepsilon}{2\sqrt{b_n}} \sqrt{\frac{\varepsilon^2}{n^2 b_n} + \frac{4k}{n}}} = \frac{\varepsilon^2}{nb_n} > 6 \log n,$$

így

$$(4.43) \quad P\left(V_n > \frac{\varepsilon}{n \sqrt{b_n}}\right) \leq \sum_{k=0}^{[n/2]} \frac{c''}{n^3} < \frac{c''}{n^2}.$$

Véleményünk szerint V_n -re még sokkal több is igaz, valószínűleg fennáll az ite-
rált $\log$ tétel, de ezt nem tudjuk bizonyítani.

A 4.2. Tételt ezzel bizonyítottuk.

A 4.2. Tételt $a_n = \frac{1}{n \log n}$, ill. $b_n = \frac{1}{n(\log n)^{1+\varepsilon}}$ -ra alkalmazva, kapjuk a követ-
kező korolláriumot:

4.2.1. KOROLLÁRIUM.

$$(4.44) \quad P \left(\lim_{n \rightarrow \infty} \left(\sqrt{n} \sup_{0 < x < 1} \frac{|x - F_n(x)|}{\sqrt{x(1-x)}} \right)^{\frac{1}{\log \log n}} = \sqrt{e} \right) = 1.$$

4.4. További megjegyzések és problémák

a) A 4.1. és 4.2. pontban ismertetett módszer segítségével aránylag pontos
felső becslés adható átmetszések valószínűségére (lásd STANLEY [86]). Ez viszont
alkalmazható a *Darling—Robbins*-féle szekvenciális próbák (lásd [38], [39]) konst-
ruálására.

Legyen $H_0: F(x) = x, \quad 0 \leq x \leq 1$

$H_1: F(x) = G(x),$

ahol $G(x) \geq x$ és $\sup_{0 < x < 1} (G(x) - x) = \Delta > 0$.

Definiáljuk N -et úgy, hogy az első olyan $n \geq m$ index, melyre

$$(4.45) \quad \sqrt{n} \sup_{0 < x < 1} (F_n(x) - x) > \psi_n,$$

ahol m adott egész szám és ψ_n alkalmas sorozat. A [86] cikk alapján (amely lénye-
gében a [31] cikk módszerét követi) megadható m és $\psi_n = O(\sqrt{\log \log n})$ úgy, hogy

$$(4.46) \quad P(N < \infty | H_0) \leq \gamma$$

előre adott γ -ra.

A *Darling—Robbins*-féle (D—R) próba a következő: elfogadjuk H_0 -t, ha
 $N = \infty$, elfogadjuk H_1 -t, ha $N < \infty$. Könnyen látható, hogy a próba ereje 1.

DARLING—ROBBINS [39] módszerét követve továbbá belátható, hogy H_1 esetén

$$(4.47) \quad E(N) = O \left(\frac{1}{\Delta^2} \log \log \frac{1}{\Delta} \right),$$

ami nagyságrendben a lehető legjobb és megegyezik egy paraméteres próbára vonat-
kozó nagyságrenddel. Itt nem volna érdektelen nem csak a nagyságrendet, hanem
a pontos konstans meghatározni, és ennek alapján összehasonlítani a D—R próbákat
efficiencia szempontjából. Ez azonban nem látszik könnyen keresztülvihetőnek.

Megjegyezzük még, hogy DARLING és ROBBINS [39] csak $\psi_n = O(\sqrt{\log \log n})$ -et tudtak
választani, és ennek megfelelően $E(N) = O \left(\frac{1}{\Delta^2} \log \Delta \right)$, ami valamivel rosszabb nagy-
ságrend, mint (4.47).

b) Bár az irodalomban többnyire $\limsup$ -ot szoktak vizsgálni, érdekes volna bizonyos $\liminf$ -re is valamit mondani.

CHUNG [27] vizsgálta $S_n^* = \max_{1 \leq i \leq n} |S_i|$ -t, ahol S_i független (azonos eloszlású) standard valószínűségi változók részletösszegei, és bizonyította, hogy

$$(4.48) \quad P \left(\lim_{n \rightarrow \infty} \left(\sqrt{\frac{\log \log n}{n}} S_n^* \right) = \frac{\pi}{\sqrt{8}} \right) = 1.$$

Ennek analógiájára — tekintve, hogy $\sqrt{n} \sup_{0 < x < 1} |x - F_n(x)|$ nagyságrendben $S_n^*/\sqrt{n}$ -el azonos határeloszlással rendelkezik — azt sejtjük, hogy

$$(4.49) \quad P \left(\lim_{n \rightarrow \infty} \left(\sqrt{n \log \log n} \sup_{0 < x < 1} |x - F_n(x)| \right) = \frac{\pi}{\sqrt{8}} \right) = 1.$$

c) FINKELSTEIN [52] bizonyította a következő tételt: Legyen

$$(4.50) \quad G_n(x) = \frac{\sqrt{n}(F_n(x) - x)}{\sqrt{2 \log \log n}}$$

és K a $[0, 1]$ -ben értelmezett $f(x)$ függvények halmaza, melyre

- (i) $f(0) = f(1) = 0$,
- (ii) $f(x)$ a Lebesgue-mértékre nézve abszolút folytonos,
- (iii) $\int_0^1 (f'(x))^2 dx \leq 1$.

Ekkor 1 valószínűséggel fennáll, hogy $G_n(x)$ torlódás-függvényeinek a halmaza K .

Ezt alkalmazva a $G_n(x)\tau(x)$ súlyozott eltérésre, ha $\tau(x)$ mind alulról, mind felülről korlátos, $0 < c_1 \leq \tau(x) \leq c_2$, kapjuk, hogy $G_n(x)\tau(x)$ torlódásfüggvényei $f(x)\tau(x)$, ahol $f(x) \in K$.

5. §. NÉHÁNY EGYÉB PROBLÉMA AZ EMPIRIKUS ELOSZLÁSFÜGGVÉNYEKEL KAPCSOLATBAN

5.1. Az empirikus eloszlásfüggvénnyel kapcsolatos egyéb statisztikák

Az egyoldali eltérésekkel kapcsolatban még számos egyéb statisztikát vizsgálhatunk. Teljességre nem törekedve, felsorolunk néhányat, melynek eloszlása a disszertációnkban ismertetett eljáráshoz hasonlóan, egyszerű kombinatorikus módszerrel meghatározható.

a) Ismeretes, hogy $(x - F_n(x))$ a maximumát egy rendezett mintaelemnél, azaz valamely $x = X_k^*$ pontban veszi fel. BIRNBAUM—PYKE [14], ill. TAKÁCS [96] foglalkozott a maximum hely és a k index eloszlásával és együttes eloszlásával. Lásd még DEMPSTER [40]. (Ők tulajdonképpen $(F_n(x) - x)$ -el foglalkoztak, de ez a fentivel ekvivalens.)

Legyen $0 < X_1^* < X_2^* < \dots < X_n^* < 1$ a $(0, 1)$ intervallumban egyenletes eloszlásból vett rendezett minta, $F_n(x)$ az empirikus eloszlásfüggvény és $y = cx - b$ egy tetszőleges egyenes, melyre $b > 0$, $c - b \leq 1$. Tekintsük a következő eseményt:

$$A_k(u) = \{X_k^* = u, F_n(u) = cu - b; F_n(x) > cx - b, \text{ ha } 0 \leq x < u, \text{ ill. } u < x \leq 1\}.$$

Kapjuk a következő tételt:

5.1. TÉTEL. Legyen $1 \leq k \leq [n(c-b)], 0 < u < 1$. Ekkor

$$(5.1) \quad \frac{P(A_k(u))}{du} = \binom{n-1}{k-1} u^{k-2} \left(\frac{nb}{c} (1-u)^{n-k} - \frac{b}{c^2} \sum_{i=0}^{[n(c-b)]-k} \binom{n-k}{i} \times \right. \\ \left. \times \left(\frac{i+1}{nc} \right)^{i-1} \left(1 - \frac{k+i}{nc} - \frac{b}{c} \right)^{n-k-i} \right).$$

Bizonyítás. Az $\{X_k^* = u\}$ esemény valószínűsége, mint ismeretes, $n \binom{n-1}{k-1} u^{k-1} (1-u)^{n-k} du$. Ezen feltétel mellett az $\{F_n(x) > cx - b, 0 \leq x < u\}$ és az $\{F_n(x) > cx - b, u < x \leq 1\}$ események függetlenek. Az 1.1. Tételből közvetlenül adódik, hogy

$$P\left(F_n(x) > cx - b, 0 \leq x < u \mid F_n(u) = \frac{k-1}{n} = cu - b\right) = \frac{b}{cu},$$

míg az 1.2. Tétel alapján

$$P\left(F_n(x) > cx - b, u < x \leq 1 \mid F_n(u+) = \frac{k}{n} = cu - b + \frac{1}{n}\right) = \\ = 1 - \sum_{i=0}^{[n(c-b)]-k} \binom{n-k}{i} \left(\frac{\frac{k+i}{nc} + \frac{b}{c} - u}{1-u} \right)^{i-1} \left(1 - \frac{\frac{k+i}{nc} + \frac{b}{c} - u}{1-u} \right)^{n-k-i} \frac{\frac{k}{nc} + \frac{b}{c} - u}{1-u} = \\ = 1 - \frac{1}{(1-u)^{n-k}} \cdot \frac{1}{nc} \sum_{i=0}^{[n(c-b)]-k} \binom{n-k}{i} \left(\frac{i+1}{nc} \right)^{i-1} \left(1 - \frac{k+i}{nc} - \frac{b}{c} \right)^{n-k-i},$$

tekintve, hogy $u = \frac{k-1}{nc} + \frac{b}{c}$.

Ezek után közvetlenül adódik (5.1).

Tekintsük a $\sup_{0 < x < 1} (x - F_n(x)) = \max_{1 \leq i \leq n} \left(X_i^* - \frac{i-1}{n} \right)$ statisztikát. Jelölje κ azt az indexet, melyre ez a maximum elérik, és legyen $X_\kappa^* = u^*$. Ekkor az 5.1. Tétel

folymányaként kapjuk $c=1$ és $b=u-\frac{k-1}{n}$ helyettesítéssel, (lásd BIRNBAUM—PYKE [14] és TAKÁCS [96])

$$\frac{dP(x=k, u^* < u)}{du} = n \binom{n-1}{k-1} \left(u - \frac{k-1}{n}\right) u^{k-2} (1-u)^{n-k} -$$

$$- \binom{n-1}{k-1} \left(u - \frac{k-1}{n}\right) u^{k-2} \sum_{i=0}^{[n(1-u)]-1} \binom{n-k}{i} \left(\frac{i+1}{n}\right)^{i-1} \left(1 - u - \frac{i+1}{n}\right)^{n-k-i},$$

ha $nu \geq k-1$; vagy az Abel-féle azonosságot (lásd (1.8)) alkalmazva

$$(1-u)^{n-k} = \sum_{i=0}^{n-k} \binom{n-k}{i} \frac{1}{n} \left(\frac{i+1}{n}\right)^{i-1} \left(1 - u - \frac{i+1}{n}\right)^{n-k-i},$$

kapjuk, hogy

$$(5.2) \quad \frac{dP(x=k, u^* < u)}{du} = \binom{n-1}{k-1} \left(u - \frac{k-1}{n}\right) u^{k-2} \times$$

$$\times \sum_{i=[n(1-u)]}^{n-k} \binom{n-k}{i} \left(\frac{i+1}{n}\right)^{i-1} \left(1 - u - \frac{i+1}{n}\right)^{n-k-i}, \quad k=1, 2, \dots, n, \quad \frac{k-1}{n} < u < 1.$$

(5.2)-ből k -ra vonatkozó összegzéssel, ill. u -ra vonatkozó integrálással nyerhetők az ismert marginális eloszlások (BIRNBAUM—PYKE [14]):

$$(5.3) \quad \frac{dP(u^* < u)}{du} = 1, \quad 0 < u < 1,$$

$$(5.4) \quad P(x=k) = \frac{1}{n^n} \sum_{i=k}^{n-1} \frac{1}{i+1} \binom{n}{i} i^i (n-i)^{n-i-1}, \quad k=1, 2, \dots, n.$$

Tekintsük most az $\frac{x-F_n(x)}{1-x}$ mennyiséget. Könnyen látható, hogy ez is egy $x=X_k^*$ mintaelemnél veszi fel a maximumát. Legyen $\tilde{x}$ az az index, ahol ez eléretik és $X_{\tilde{x}}^* = \tilde{u}^*$. Ekkor az 5.1. Tételben végezzük el a következő helyettesítéseket:

$$b = \frac{u - \frac{k-1}{n}}{1-u}, \quad c = \frac{1 - \frac{k-1}{n}}{1-u}.$$

Itt még az (5.1) formula némileg egyszerűsíthető, ui. az 1.1. Tételből közvetlenül adódik, hogy a fenti b és c -vel

$$P\left(F_n(x) > cx - b, u < x < 1/F_n(u+)\right) = \frac{k}{n} = cu - b + \frac{1}{n} = \frac{1}{nc(1-u)}.$$

Így kapjuk (lásd DEMPSTER [40]) a

$$(5.5) \quad \frac{dP(\tilde{x} = k, \tilde{u}^* < u)}{du} = \frac{n}{n-k+1} \binom{n}{k-1} \left(u - \frac{k-1}{n}\right) u^{k-2} (1-u)^{n-k},$$

ill. a

$$P(\tilde{x} = k, \tilde{u}^* < u) = \frac{1}{n-k+1} \binom{n}{k-1} \left[\left(\frac{k-1}{n}\right)^{k-1} \left(1 - \frac{k-1}{n}\right)^{n-k+1} - u^{k-1} (1-u)^{n-k+1} \right]$$

összefüggéseket, ha $\frac{k-1}{n} < u < 1$, $k = 1, 2, \dots, n$.

Az $u=1$ helyettesítéssel,

$$(5.6) \quad P(\tilde{x} = k) = \frac{1}{n-k+1} \binom{n}{k-1} \left(\frac{k-1}{n}\right)^{k-1} \left(1 - \frac{k-1}{n}\right)^{n-k+1}, \quad k = 1, 2, \dots, n.$$

$\tilde{x}$ fenti eloszlásából látható, hogy az n -hez közeli értékeket veszi fel nagy valószínűséggel. Nem igaz azonban az, hogy a $P(\tilde{x}=k)$ értékek monoton sorozatot alkotnak, mert

$$P(\tilde{x} = 1) = \frac{1}{n}$$

$$P(\tilde{x} = 2) = \frac{1}{n} \left(1 - \frac{1}{n}\right)^{n-2} < \frac{1}{n}, \quad \text{ha } n > 2.$$

A valószínűségek először monoton fogynak, és csak később növekszenek. A maximum természetesen $k=n$ -nél van. Igaz továbbá a következő határérték reláció, mely az (5.6) formulából adódik, tagonkénti határártmenettel.

Legyen $s=n-k$ rögzített, akkor

$$(5.7) \quad \lim_{n \rightarrow \infty} P(\tilde{x} = n-s) = \frac{(s+1)^{s-1}}{s!} e^{-(s+1)}, \quad s = 0, 1, 2, \dots$$

b) CHENG [24] és TAKÁCS [96] foglalkozott az $\frac{r}{n} - X_r^*$ ($r=1, 2, \dots, n$) mennyiségek közül a pozitívok számával. Ezzel kapcsolatban felmerül a kérdés, hogy mi azon $\left(X_i^*, \frac{i-1}{n}\right)$ pontok számának az eloszlása, melyek egy adott $y=cx-b$ egyenes alatt vannak. Sajnos, az általános esetre nem sikerült egyszerű formulát találni, ezért csak olyan egyenessel foglalkozunk, mely átmegy az $(1, 1)$ ponton.

Tekintsük tehát az $y = \frac{x-p}{1-p}$ egyenest ($0 < p < 1$) és jelölje τ azon $\left(X_i^*, \frac{i-1}{n}\right)$ pontok számát, melyek a fenti egyenes alatt vannak, azaz melyekre $\frac{X_i^* - p}{1-p} > \frac{i-1}{n}$, $i=1, 2, \dots, n$. τ eloszlását adja meg a következő tétel:

5.2. TÉTEL. Legyen $q = \frac{1-p}{n}$, ekkor

(5.8)

$$P(\tau = j) = \sum_{t=0}^{n-j} (p-q) \binom{n}{t} q^{n-t} (p+(t-1)q)^{t-1} (n-t+1)^{n-t-1}, \quad j = 0, 1, 2, \dots, n.$$

Bizonyítás. Először is megjegyezzük, hogy $t=0$ -nak megfelelő tag $(p-q)q^n(p-q)^{-1}(n+1)^{n-1} = q^n(n+1)^{n-1}$, és így a fenti formula $p=q = \frac{1}{n+1}$ -re a következőt adja:

$$(5.9) \quad P(\tau = j) = \frac{1}{n+1}, \quad j = 0, 1, 2, \dots, n.$$

Ez a formula — melyet a továbbiakban felhasználunk — következik TAKÁCS alábbi lemmájából [96]:

Ha $\gamma_1, \gamma_2, \dots, \gamma_N$ egész értékeket felvevő felcserélhető változók, és Δ_N jelöli a $\gamma_1 + \dots + \gamma_i$ ($i=1, 2, \dots, N$) összegek közül a pozitívok számát, akkor

$$P(\Delta_N = j | \gamma_1 + \dots + \gamma_N = 1) = \frac{1}{N}, \quad j = 1, 2, \dots, N,$$

feltéve, hogy a feltételes valószínűség értelmezhető.

A bizonyításban legyen először $p > q$. Ekkor tekintsük az $y = \frac{x-(p-q)}{1-p} = \frac{x-p}{1-p} + \frac{1}{n}$ egyenest. Az empirikus eloszlásfüggvény ezt valamely $\frac{k}{n}$ magasságon feltétlenül metszi ($0 \leq k \leq n$). Jelölje t az első ilyen k -t, és legyen $x_t = p + \frac{t-1}{n}(1-p) = p + (t-1)q$. Az $F_n(x_t) = \frac{t}{n}$ esemény akkor és csak akkor következik be, ha a $(0, x_t)$ intervallumon pontosan t pont van, ennek valószínűsége

$$\begin{aligned} \binom{n}{t} x_t^t (1-x_t)^{n-t} &= \binom{n}{t} (p+(t-1)q)^t (1-p-(t-1)q)^{n-t} = \\ &= \binom{n}{t} (p+(t-1)q)^t q^{n-t} (n-t+1)^{n-t}. \end{aligned}$$

Az $F_n(x_t) = \frac{t}{n}$ feltétel mellett annak valószínűsége, hogy előbb nem történik metszés, az 1.1. Tétel szerint $\frac{x_0}{x_t} = \frac{p-q}{p+(t-1)q}$, és annak valószínűsége, hogy az $(x_t, 1)$ intervallumban éppen j darab pont van az $y = \frac{x-p}{1-p}$ egyenes alatt, a fent idézett Takács-lemma szerint $\frac{1}{n-t+1}$. (A $(0, x_t)$ intervallumban ui. minden pont az egye-

nes felett van.) t -re való összegzéssel kapjuk a Tétel állítását a $p > q$ esetben. (t -re nyilván teljesülnie kell az $n-t \geq j$ egyenlőtlenségnek.)

A $p < q$ esetre a fenti okoskodást némileg módosítani kell. Tekintsük a $(p-q, 1)$ intervallumot (amely magában foglalja a $(0, 1)$ intervallumot), és tegyük fel, hogy most ezen az intervallumon veszünk n egyenletes eloszlású pontot: $\bar{X}_1^* < \bar{X}_2^* < \dots < \bar{X}_n^*$.

Legyen $\bar{\tau}$ azon $\left(\bar{X}_i^*, \frac{i-1}{n}\right)$ pontok száma, melyek az $y = \frac{x-p}{1-p}$ egyenes alatt vannak. Ekkor a Takács-lemma szerint $P(\bar{\tau}=j) = \frac{1}{n+1}$, $j=0, 1, \dots, n$. Ezt a valószínűséget másképpen is előállíthatjuk. Jelölje μ a $(p-q, 0)$ intervallumba eső mintaelemek számát. Ekkor

$$(5.10) \quad P(\bar{\tau} = j) = P(\mu = 0)P(\bar{\tau} = j | \mu = 0) + P(\bar{\tau} = j, \mu \geq 1).$$

Nyilván $P(\mu=0) = \frac{1}{(1+q-p)^n} = \frac{1}{q^n(n+1)^n}$, és $P(\bar{\tau}=j | \mu=0)$ éppen a keresett valószínűség. Ha $j=n$, akkor $P(\bar{\tau}=n, \mu \geq 1) = 0$, így kapjuk

$$P(\bar{\tau} = n | \mu = 0) = q^n(n+1)^{n-1},$$

azaz ebben az esetben igaz a tétel állítása. Ha $j < n$, akkor $\mu \geq 1$ esetén létezik olyan $1 \leq t \leq n-j$, hogy $F_n(x_t) = \frac{t}{n}$ és $F_n(x_i) > \frac{i}{n}$ $i = 1, 2, \dots, t-1$, ahol $x_i = p + (i-1)q$. Az $F_n(x_t) = \frac{t}{n}$ esemény valószínűsége most

$$\binom{n}{t} \left(\frac{tq}{1+q-p} \right)^t \left(1 - \frac{tq}{1+q-p} \right)^{n-t} = \binom{n}{t} \frac{(tq)^t q^{n-t} (n+1-t)^{n-t}}{(1+q-p)^n}.$$

Ezen feltétel mellett annak valószínűsége, hogy az $(x_t, 1)$ intervallumon j pont van az $y = \frac{x-p}{1-p}$ egyenes alatt, ugyancsak a Takács-lemma szerint $\frac{1}{n-t+1}$. Meg kell határoznunk még az $F_n(x_t) = \frac{t}{n}$ feltétel mellett annak valószínűségét, hogy $\mu \geq 1$ és t az első olyan index, melyre $F_n(x_t) = \frac{t}{n}$. Az 1.1. Tételt alkalmazva, erre a valószínűségét kapjuk:

$$\begin{aligned} & \sum_{m=1}^t \binom{t}{m} \left(\frac{q-p}{tq} \right)^m \left(1 - \frac{q-p}{tq} \right)^{t-m} \frac{p+(m-1)q}{p+(t-1)q} = \\ &= \frac{1}{p+(t-1)q} \sum_{m=1}^t \binom{t}{m} \left(\frac{q-p}{tq} \right)^m \left(1 - \frac{q-p}{tq} \right)^{t-m} (p-q+mq) = \\ &= \frac{1}{p+(t-1)q} \left[(p-q) \left(1 - \left(1 - \frac{q-p}{tq} \right)^t \right) + qt \frac{q-p}{tq} \right] = \frac{(q-p)(p+(t-1)q)^{t-1}}{(tq)^t}. \end{aligned}$$

Innen

$$P(\bar{\tau} = j, \mu \equiv 1) = \sum_{t=1}^{n-j} \binom{n}{t} (q-p) \frac{(p+(t-1)q)^{t-1} q^{n-t} (n-t+1)^{n-t-1}}{(1+q-p)^n},$$

és (5.10)-ből $P(\bar{\tau}=j|\mu=0)$ -ra éppen a Tétel állítását kapjuk. Ezzel az 5.2. Tételt minden esetre bizonyítottuk.

MEGJEGYZÉSEK. 1. $j=0$ -ra az 1.1. Tételből $P(\tau=0)=p$, tehát

$$(5.11) \quad \sum_{t=0}^n (p-q) \binom{n}{t} q^{n-t} (p+(t-1)q)^{t-1} (n-t+1)^{n-t-1} = p,$$

ami egyébként némi számolással az *Abel*-azonosságból is belátható. Ezt felhasználva kaphatjuk az alábbi formulát:

$$(5.12) \quad P(\tau = j) = p + \sum_{s=0}^{j-1} (q-p) \binom{n}{s} q^s (1-(s+1)q)^{n-s-1} (s+1)^{s-1}.$$

2. $p=0$ -ra kapjuk CHENG [24] eredményét:

$$(5.13) \quad P(\tau = j) = \sum_{s=0}^{j-1} \binom{n}{s} \frac{1}{n^{s+1}} (s+1)^{s-1} \left(1 - \frac{s+1}{n}\right)^{n-s-1}.$$

3. $p>q$ -ra a valószínűségek monoton fogyók:

$$P(\tau = 0) > P(\tau = 1) > \dots > P(\tau = n),$$

míg $p<q$ -ra a valószínűségek monoton növekednek:

$$P(\tau = 0) < P(\tau = 1) < \dots < P(\tau = n).$$

$p=q$ -ra természetesen

$$P(\tau = 0) = P(\tau = 1) = \dots = P(\tau = n) = \frac{1}{n+1}.$$

4. Ha p rögzített, az 5.2. Tételből közvetlenül adódik, hogy

$$(5.14) \quad \lim_{n \rightarrow \infty} P(\tau = j) = p \sum_{s=j}^{\infty} \frac{(1-p)^s (s+1)^{s-1}}{s!} e^{-(1-p)(s+1)}, \quad j = 0, 1, 2, \dots$$

c) $F_n(x)$ és valamely $y=cx-b$ egyenes átmetszéseiinek számával foglalkozott CHANG [21], NEF [70], TAKÁCS [98], CSÁKI—TUSNÁDY [32]. Itt csupán ez utóbbi cikkből idézzük az eredményeket.

5.3. TÉTEL. Az 1.1. Tétel feltételei mellett jelölje λ azon m indexek számát, melyekre

$$\sum_{i=1}^m v_i = m, \quad m = 1, 2, \dots, n.$$

Ekkor

$$(5.15) \quad P(\lambda = l) = n(n-1)\dots(n-l+1)q^l, \quad l = 1, 2, \dots, n.$$

Az 5.3. Tétel egyébként megadja az $F_n(x)$ és az $y = \frac{x-p}{1-p}$ metszései számának az eloszlását. Általánosabban, tekintsük az $F_n(x)$ és $y = cx - b$ egyenes átmetszései-nek számát, azaz azon x pontok számát, melyekre $F_n(x) = cx - b$. Jelölje ezt β . Az 5.3. Tétel következményeként kapjuk:

5.4. TÉTEL.

$$(5.16) \quad P(\beta \geq l+1) = \frac{n(n-1)\dots(n-l+1)}{(nc)^l} \left(\frac{l}{nc} + \frac{b}{c} \right) \times \\ \times \sum_{k=l}^{[n(c-b)]} \binom{n-l}{k-l} \left(\frac{k}{nc} + \frac{b}{c} \right)^{k-l-1} \left(1 - \frac{k}{nc} - \frac{b}{c} \right)^{n-k}, \quad l = 0, 1, \dots, [n(c-b)].$$

Ha $c = 1 + b$, akkor az 5.3. Tételből közvetlenül

$$P(\beta \geq l+1) = n(n-1)\dots(n-l+1) \frac{1}{(n(1+b))^l}.$$

Így pl. $b=0, l=2$ -re

$$(5.17) \quad P(\beta \geq 3) = \frac{n(n-1)}{n^2} = 1 - \frac{1}{n},$$

ami annak valószínűségét adja, hogy az $x=0$ és $x=1$ -en kívül legyen még legalább egy olyan x_0 , ($0 < x_0 < 1$), hogy $F_n(x_0) = x_0$ (lásd az 1970. évi SCHWEITZER MIKLÓS emlékverseny 12. feladatát).

Az említett [32] cikkben határeloszlások is találhatók, valamint a véletlen mintaelemszám esetével is foglalkozunk.

5.2. Véletlen mintaelemszám

Gyakran előfordul, különösen élettartam vizsgálatoknál, biológiában stb. hogy a mintaelemszám nem előre rögzített, hanem valószínűségi változó. Sokszor feltehető, hogy a mintaelemszám Poisson eloszlású, és független a mintaelemektől. Mi csupán ezzel az esettel foglalkozunk. Tegyük fel, hogy

$$(5.18) \quad P(n = i) = \frac{\lambda^i}{i!} e^{-\lambda}, \quad i = 0, 1, 2, \dots$$

és n független az $(X_1, X_2, \dots)$ változóktól.

KAC [60] ebben az esetben a következőképpen módosította az empirikus eloszlásfüggvényt:

$$(5.19) \quad F_\lambda^*(x) = \frac{1}{\lambda} \sum_{\substack{X_i < x \\ 1 \leq i \leq n}} 1, \quad \text{ha } n \geq 1,$$

$$F_\lambda^*(x) = 0, \quad \text{ha } n = 0.$$

Fennáll a következő összefüggés:

$$F_\lambda^*(x) = \frac{n}{\lambda} F_n(x).$$

$F_\lambda^*(x)$ -en alapuló egyoldali eltérések eloszlásával foglalkozott TAKÁCS [95], ALLEN—BEEKMAN [2], ALVO—CSÖRGŐ [3].

A következő tétel — amely formuláikat speciális esetként tartalmazza — az 1.4. Tételnek megfelelője.

5.5. TÉTEL. Legyen $0 \leq a_1 < a_2 \leq 1$, $c > 0$, ekkor

$$(5.20) \quad P(F_\lambda^*(x) \geq cx + b, a_1 \leq x \leq a_2) = \\ = \sum_{\alpha > \lambda(ca_1+b)} \frac{(\lambda a_1)^\alpha}{\alpha!} e^{-\lambda a_1} - \sum_{\alpha > \lambda(ca_1+b)} \frac{(\lambda a_1)^\alpha}{\alpha!} \left(\frac{\alpha - b}{c} - a_1 \right)^{[\lambda(ca_2+b)] - \alpha} \frac{1}{j!} \times \\ \times \left(\frac{\alpha + j - \lambda b}{c} - \lambda a_1 \right)^{j-1} e^{-\frac{\alpha + j - \lambda b}{c}}, \quad ca_1 + b > 0, \\ + \frac{\lambda b}{c} \sum_{j=0}^{[\lambda(ca_2+b)]} \frac{(j - \lambda b)^{j-1}}{c^{j-1} j!} e^{-\frac{j - \lambda b}{c}}, \quad ca_1 + b \leq 0.$$

Bizonyítás. Nyilvánvalóan fennáll a következő összefüggés:

$$(5.21) \quad P(F_\lambda^*(x) \geq cx + b, a_1 \leq x \leq a_2) = \\ = \sum_{n=0}^{\infty} P\left(F_n(x) \geq \frac{c\lambda}{n}x + \frac{b\lambda}{n}, a_1 \leq x \leq a_2 | n\right) \frac{\lambda^n}{n!} e^{-\lambda}.$$

Az összegzésben szereplő feltételes valószínűséget az 1.4. Tétel adja. A megfelelő formulát behelyettesítve, majd az összegzést felcserélve és n -re elvégezve kapjuk a Tétel állítását.

Az $a_1 = 0$, $a_2 = 1$, $c = 1$, $b = -\varepsilon$ esetet lásd TAKÁCS [95] és ALLEN—BEEKMAN [2].

Az $a_1 = 0$, $c = 1$, $b = -\varepsilon$, ill. az $a_1 = 0$, $c = 1 + \varepsilon$, $b = -\varepsilon$ esetet lásd ALVO—CSÖRGŐ [3].

Az 1. § többi formuláinak megfelelőjét is hasonlóképpen kaphatnánk meg.

5.3. Kétdimenziós eset

Ismeretes, hogy többdimenzióban a *Kolmogorov*-statisztika nem eloszlásmentes. Mindig elérhető azonban (lásd pl. DURBIN [45]), hogy az eloszlást az egységnyezetben (ill. egység kockában) egyenletes eloszlásba transzformáljuk, így a feladat arra redukálódik, hogy ellenőrizzük, hogy az eloszlás az egységnyezetben egyenletes-e. Ez a transzformáció azonban nem egyértelmű, másrészt — az egydimenziós esettel ellentétben — a *Kolmogorov*-statisztika erre a transzformációra nem invariáns. További problémát jelent, hogy a kapott eloszlások nehezen kezelhetők. A többdimenziós esetet egyébként részletesen tárgyalja PURI-SEN [73] könyve.

Itt csupán két problémára kívánunk kitérni. DURBIN [45] vezette be a következő statisztikát:

$$(5.22) \quad G_n^- = \sup_{0 \leq x \leq 1} \left(\frac{x}{2} - G_n(x) \right),$$

ahol

$$G_n(x) = \int_0^1 F_n(x, y) dy,$$

$F_n(x, y)$ a kétdimenziós empirikus eloszlásfüggvény. A Doob-féle heurisztikát alkalmazva határozta meg G_n^- határeloszlását. Az alábbiakban meghatározzuk G_n^- pontos eloszlását, amiből ugyancsak megkapható a határeloszlás.

Legyen $(X_1, Y_1), \dots, (X_n, Y_n)$ a kétdimenziós minta. DURBIN [45] szerint

$$(5.23) \quad G_n(x) = \frac{1}{n} \left(r - \sum_{Y_j < x} Y_j \right) = \frac{1}{n} \sum_{X_j < x} (1 - Y_j), \quad X_r^* < x \leq X_{r+1}^*.$$

Jelölje A_r az $\{X_r^* < x \leq X_{r+1}^*\}$ eseményt. Így $nG_n(x)$ az A_r feltétel mellett r darab, független, $(0, 1)$ -ben egyenletes eloszlású valószínűségi változó összege.

Nyilván

$$E(G_n(x)|A_r) = \frac{r}{2n},$$

tehát

$$E(G_n(x)) = E\left(\frac{r}{2n}\right) = \frac{x}{2}.$$

A

$$P(G_n(x) > cx - b, 0 \leq x < 1)$$

valószínűség meghatározása az 1.2. Tételnél alkalmazott módszerrel történhet. Nem lényeges, hogy $nG_n(x)$ éppen egyenletes eloszlású változók összege, tekinthetjük általánosabban:

$$(5.24) \quad G_n(x) = \frac{1}{n} \sum_{j=1}^r Z_j, \quad X_r^* < x \leq X_{r+1}^*,$$

ahol $Z_1, Z_2, \dots, Z_n$ független azonos eloszlású, pozitív valószínűségi változók.

Az 1.1. Tétel átvihető ilyen „véletlen ugrású” empirikus eloszlásfüggvényre: legyen $0 < p < s$, $S_0 = p$, $S_i = p + \sum_{j=1}^i U_j$, $i = 1, 2, \dots, n$, ahol $U_1, \dots, U_n$ független, azonos eloszlású, pozitív valószínűségi változók. Tekintsük az $\{S_n = s\}$ feltételt és vegyünk a $(0, s)$ intervallumon egymástól és U_i -ktől függetlenül n egyenletes eloszlású pontot. Jelölje v_i az (S_{i-1}, S_i) véletlen intervallumba eső pontok számát, $i = 1, \dots, n$. Ekkor az 1.1. Tételhez hasonlóan

$$(5.25) \quad P(v_1 < 1, v_1 + v_2 < 2, \dots, v_1 + v_2 + \dots + v_n < n | S_n = s) = \frac{p}{s}.$$

Az állítás következik TAKÁCS [97] 3. fej. 1. Tételéből, de közvetlenül is belátható az 1.1. Tételhez hasonlóan. Jelölje v_0 a $(0, p)$ intervallumba eső pontok számát, és alkalmazzunk teljes indukciót:

$$P(v_1 < 1, v_1 + v_2 < 2, \dots, v_1 + \dots + v_n < n | S_n = s, v_0, S_{v_0}) = \frac{S_{v_0}}{s}.$$

Belátható, hogy $E\left(\frac{S_{v_0}}{s} \middle| S_n = s, v_0\right) = \frac{v_0}{n}$, így

$$P(v_1 < 1, v_1 + v_2 < 2, \dots, v_1 + \dots + v_n < n | S_n = s, v_0) = \frac{v_0}{n},$$

végül

$$P(v_1 < 1, \dots, v_1 + \dots + v_n < n | S_n = s) = E \left(\frac{v_0}{n} \middle| S_n = s \right) = \frac{p}{s}.$$

(5.25) segítségével meghatározható az 1.2. stb. Tételeknél tekintett valószínűségek megfelelője a „véletlen ugrású” empirikus eloszlásfüggvényre. Exponenciális eloszlású U_i -kre érdekes, egyszerű formulákat adott meg LÁSZLÓ [65] egy készletgazdálkodási problémával kapcsolatban.

Az alábbi tételünk általános eloszlású U_i -kre vonatkozik, mikor is a formula általában valószínűleg nem egyszerűsíthető.

5.6. TÉTEL. Jelölje $H_r(u)$ a $\sum_{j=1}^r Z_j$ eloszlásfüggvényét, legyen $0 < c < 1$, $0 < b < 1$.

Ekkor

$$(5.26) \quad P(G_n(x) > cx - b, 0 < x < 1) = \\ = 1 - \left(1 - \frac{b}{c}\right)^n - \frac{b}{c} \sum_{r=1}^n \binom{n}{r} \int_0^{n(c-b)} \left(\frac{u}{n} + b\right)^{r-1} \left(1 - \frac{\frac{u}{n} + b}{c}\right)^{n-r} dH_r(u).$$

Bizonyítás. Annak valószínűségét határozzuk meg, hogy $G_n(x)$ átmetszi az $y = cx - b$ egyenest. Annak valószínűsége, hogy ez az $u=0$ magasságon történjék, nyilván $\left(1 - \frac{b}{c}\right)^n$. Tekintsük most a $\left\{\sum_{j=1}^r Z_j = u\right\}$ eseményt, melynek valószínűsége

$dH_r(u)$, ezen feltétel mellett annak valószínűsége, hogy $\left\{G_n\left(\frac{u}{n} + b\right) = \frac{u}{n}\right\}$, nyilván $\binom{n}{r} \left(\frac{u}{n} + b\right)^r \left(1 - \frac{\frac{u}{n} + b}{c}\right)^{n-r}$, és annak valószínűsége, hogy r az első olyan index,

amelyre igaz a fenti egyenlőség, (5.25)-öt $p = \frac{b}{c}$, $s = \frac{\frac{u}{n} + b}{c}$ -vel alkalmazva $\frac{b}{\frac{u}{n} + b}$,

Ezzel szorozva, r -re összegezve, és u szerint integrálva adódik (5.26).

Visszatérve a Durbin-féle problémára, $H_r(u)$ r darab $(0, 1)$ -ben egyenletes eloszlású változó összegének eloszlásfüggvénye, $h_r(u) = H'_r(u)$; $c = 1/2$, $b = \varepsilon$ -nal kapjuk:

$$(5.27) \quad P\left(\sup_{0 < x < 1} \left(\frac{x}{2} - G_n(x)\right) < \varepsilon\right) = P\left(G_n(x) > \frac{x}{2} - \varepsilon, 0 < x < 1\right) = \\ = 1 - (1 - 2\varepsilon)^n - 2\varepsilon \sum_{r=1}^n \binom{n}{r} \int_0^{n(1/2 - \varepsilon)} \left(\frac{2u}{n} + 2\varepsilon\right)^{r-1} \left(1 - \frac{2u}{n} - 2\varepsilon\right)^{n-r} h_r(u) du.$$

Ebből határeloszlást a következőképpen számíthatunk:

Legyen $\varepsilon = \frac{t}{\sqrt{n}}$, $r = nz$, és helyettesítsünk az integrálban u helyére

$$\frac{u - \frac{r}{2}}{\sqrt{r}} \sqrt{12} = v - t.$$

Ezzel nyilván standardizáltuk a változót, és így a centrális határeloszlás tétel értelmében

$$h_r(u) du \sim \frac{1}{\sqrt{2\pi}} e^{-\frac{v^2}{2}} dv,$$

és a de Moivre—Laplace-tételt alkalmazva

$$\binom{n}{r} \left(\frac{2u}{n} + 2\varepsilon \right)^r \left(1 - \frac{2u}{n} - 2\varepsilon \right)^{n-r} \sim \frac{1}{\sqrt{2\pi n z(1-z)}} e^{-\frac{\left(v \sqrt{\frac{z}{3} + 2t} \right)^2}{2z(1-z)}}.$$

Így v szerint a következő integrált kell kiszámítani:

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-\frac{\left(v \sqrt{\frac{z}{3} + 2t} \right)^2}{2z(1-z)}} e^{-\frac{v^2}{2}} dv = \sqrt{\frac{3(1-z)}{4-3z}} e^{-\frac{1}{2} \frac{12t^2}{z(4-3z)}}.$$

Tekintve, hogy

$$\frac{2\varepsilon}{\frac{2u}{n} + 2\varepsilon} \sim \frac{2t}{z\sqrt{n}} \quad \text{és} \quad dz = \frac{1}{n},$$

az r -re vonatkozó összegzésből határértékben a következő integrált kapjuk:

$$\begin{aligned} & \int_0^1 \frac{1}{\sqrt{2\pi}} \frac{2t}{z\sqrt{z(1-z)}} \sqrt{\frac{3(1-z)}{4-3z}} e^{-\frac{1}{2} \frac{12t^2}{z(4-3z)}} dz = \\ &= \int_0^{3/4} \frac{1}{\sqrt{2\pi}} \frac{\frac{3}{2}t}{w\sqrt{w(1-w)}} e^{-\frac{\left(\frac{3}{2}t\right)^2}{2w(1-w)}} dw = e^{-\frac{9}{2}t^2} \Phi(t\sqrt{3}) + 1 - \Phi(2t\sqrt{3}), \end{aligned}$$

ahol közben a $z = \frac{4w}{3}$ helyettesítést végeztük el, és a 2.1. Lemmát alkalmaztuk.

Kapjuk tehát a DUBIN által is nyert határeloszlást:

$$(5.28) \quad \lim_{n \rightarrow \infty} P \left(\sup_{0 < x < 1} \left| \frac{x}{2} - G_n(x) \right| < t \sqrt{n} \right) = \Phi(2\sqrt{3}t) - e^{-\frac{9}{2}t^2} \Phi(t\sqrt{3}), \quad t > 0.$$

A másik probléma, amivel itt röviden kívánunk foglalkozni, VINCZE [102] kétféle problémájának egymintás megfelelője:

Mi a valószínűsége annak, hogy egy véletlenül választott y -ra

$$\left\{ \sup_{0 < x < 1} (xy - F_n(x, y)) < \varepsilon \right\}?$$

Ezt adja meg a következő tétel:

5.7. TÉTEL.

$$(5.29) \quad E_y(P(\sup_{0 < x < 1} (xy - F_n(x, y)) < \varepsilon)) = \int_0^1 P(\sup_{0 < x < 1} (xy - F_n(x, y)) < \varepsilon) dy = \\ = 1 - \varepsilon \sum_{j=0}^{[n(1-\varepsilon)]} \binom{n}{j} \left(\frac{j}{n} + \varepsilon\right)^{j-1} \left(1 - \frac{j}{n} - \varepsilon\right)^{n-j+1}.$$

Bizonyítás. Belátható, hogy minden rögzített y -ra

$$(5.30) \quad P(\sup_{0 < x < 1} (xy - F_n(x, y)) < \varepsilon) = P(\sup_{0 < x < y} (x - F_n(x)) < \varepsilon),$$

így az 1.2. Tételt alkalmazva, $a=y$, $c=1$, $b=\varepsilon$ helyettesítéssel,

$$(5.31) \quad P(\sup_{0 < x < 1} (xy - F_n(x, y)) < \varepsilon) = \\ = \begin{cases} 1, & \text{ha } 0 < y \leq \varepsilon \\ 1 - \varepsilon \sum_{j=0}^{[n(y-\varepsilon)]} \binom{n}{j} \left(\frac{j}{n} + \varepsilon\right)^{j-1} \left(1 - \frac{j}{n} - \varepsilon\right)^{n-j}, & \text{ha } \varepsilon < y < 1. \end{cases}$$

(5.31)-et y szerint integrálva adódik (5.29).

A 2.1. Tételből megkapható a VINCZE [102] által is nyert határérték:

$$(5.32) \quad \lim_{n \rightarrow \infty} E_y \left(P \left(\sup_{0 < x < 1} (xy - F_n(x, y)) < \frac{t}{\sqrt{n}} \right) \right) = \\ = \int_0^1 \Phi \left(\frac{t}{\sqrt{y(1-y)}} \right) dy - e^{-2t^2} \int_0^1 \Phi \left(\frac{2y-1}{\sqrt{y(1-y)}} t \right) dy, \quad t > 0.$$

IRODALOM

- [1] ABRAHAMSON, J. G.: "Exact Bahadur efficiencies for the Kolmogorov—Smirnov and Kuiper one- and two-sample statistics", *Ann. Math. Stat.* **38** (1967), 1475—1490.
- [2] ALLEN, J. L.—BEEKMAN, J. A.: "A statistical test involving a random number of random variables" *Ann. Math. Stat.* **37** (1966), 1305—1311.
- [3] ALVO, M.—CSÖRGŐ, M.: "Distribution results and power functions for Kac statistics", *Ann. Inst. Stat. Math.* **22** (1970), 257—260.
- [4] ANDÉL, J.: "Local asymptotic power and efficiency of tests of Kolmogorov—Smirnov type", *Ann. Math. Stat.* **38** (1967), 1705—1725.
- [5] ANDERSON, T. W.—DARLING, D. A.: "Asymptotic theory of certain 'Goodness of Fit' criteria based on stochastic processes", *Ann. Math. Stat.* **23** (1952), 193—212.
- [6] BAHADUR, R. R.: "Stochastic comparison of tests", *Ann. Math. Stat.* **31** (1960), 276—295.
- [7] BAHADUR, R. R.: "Some limit theorems in statistics", Regional conference series in applied mathematics, Philadelphia, 1971.
- [8] BARTON, D. E.—MALLOWS, C. L.: "Some aspects of the random sequence", *Ann. Math. Stat.* **36** (1965), 236—260.

- [9] BAXTER, G.: "An analogue of the law of the iterated logarithm", *Proc. Amer. Math. Soc.* **6** (1955), 177—181.
- [10] BERTRAND, J.: «Solution d'un problème», *C. R. Acad. Sci. Paris* **105** (1887), 369.
- [11] BIRNBAUM, Z. W.: "On the power of a one sided test of fit for continuous probability functions", *Ann. Math. Stat.* **24** (1953), 484—489.
- [12] BIRNBAUM, Z. W.—LIENTZ, B. P.: "Exact distributions for some Rényi-type statistics", *Zastosowania Matematyki* (Applicationes Mathematicae), **10** (1969) (Hugo Steinhaus Jubilee Volume), 179—192.
- [13] BIRNBAUM, Z. W.—MCCARTY, R. C.: "A distribution-free upper confidence bound for $Pr(Y < X)$ based on independent samples of X and Y ", *Ann. Math. Stat.* **29** (1958) 558—562.
- [14] BIRNBAUM, Z. W.—PYKE, R.: "On some distributions related to the statistic D_n^+ ", *Ann. Math. Stat.* **29** (1958), 179—187.
- [15] BIRNBAUM, Z. W.—TINGEY F. H.: "One sided confidence contours for probability distribution functions", *Ann. Math. Stat.* **22** (1951), 592—596.
- [16] BRUNK, H. D.: "On the range of the difference between hypothetical distribution function and Pyke's modified empirical distribution function", *Ann. Math. Stat.* **33** (1962), 525—532.
- [17] BUTLER, J. B.—MCCARTY, R. C.: "A lower bound for the distribution of the statistic D_n^+ " (Abstract), *Notices Amer. Math. Soc.* **7** (1960), 80—81.
- [18] CARNAL, H.: «Sur les théorèmes de Kolmogorov et Smirnov dans le cas d'une distribution discontinue», *Comm. Math. Helv.* **37** (1962), 19—35.
- [19] CASSELS, J. W. S.: "An extension of the law of the iterated logarithm", *Proc. Cambridge Philos. Soc.* **47** (1961), 55—64.
- [20] CHANG, LI CHIEN: "On the ratio of an empirical distribution function to the theoretical distribution function", *Acta Math. Sinica* **5** (1955), 437—468.
- [21] CHANG, LI CHIEN: "Relative positions of the empirical and theoretical distribution functions", *Academic Records of Peking Univ.* **2** (1956), 129—157.
- [22] CHAPMAN, D. G.: "A comparative study of several one-sided goodness of fit tests", *Ann. Math. Stat.* **29** (1958), 655—674.
- [23] CHAPMAN, D. G.: "On a limiting distribution due to Rényi", *Ann. Math. Stat.* **29** (1958), 1282.
- [24] CHENG, P.: "Non-negative jump points of an empirical distribution function relative to a theoretical distribution function", *Acta Math. Sinica* **8** (1958), 333—347.
- [25] CHERNOFF, H.: "A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations", *Ann. Math. Stat.* **23** (1952), 213—225.
- [26] CHERNOFF, H.: "Large sample theory: parametric case", *Ann. Math. Stat.* **27** (1956), 1—22.
- [27] CHUNG, K. L.: "On the maximum partial sums of sequences of independent random variables", *Trans. Amer. Math. Soc.* **64** (1948), 205—233.
- [28] CHUNG, K. L.: "An estimate concerning the Kolmogorov limit distribution", *Trans. Amer. Math. Soc.* **67** (1949), 36—50.
- [29] COBERLY, W. A.—LEWIS, T. O.: "A note on a one-sided Kolmogorov—Smirnov test of fit for discrete distribution function", *Ann. Inst. Stat. Math.* **24** (1972), 183—188.
- [30] CONOVER, W. J.: "A Kolmogorov-type goodness of fit test for discontinuous distributions", *Journ. Amer. Stat. Ass.* **67** (1972), 591—596.
- [31] CSÁKI, E.: "An iterated logarithm law for semimartingales and its application to empirical distribution function", *Studia Sci. Math. Hung.* **3** (1968), 287—292.
- [32] CSÁKI, E.—TUSNÁDY, G.: "On the number of intersections and the ballot theorem", *Periodica Math. Hung.* **2** (1972), 5—13.
- [33] CSÖRGÖ, M.: "Exact and limiting probability distributions of some Smirnov type statistics", *Canad. Math. Bull.* **8** (1965), 93—103.
- [34] CSÖRGÖ, M.: "Exact probability distribution functions of some Rényi type statistics", *Proc. Amer. Math. Soc.* **16** (1965), 1158—1167.
- [35] CSÖRGÖ, M.: "Some Rényi-type limit theorems of empirical distribution function", *Ann. Math. Stat.* **36** (1965), 322—326.
- [36] DANIELS, H. E.: "The statistical theory of the strengths of bundles of threads, I." *Proc. Roy. Soc. A* **183** (1945), 405—435.
- [37] DARLING, D. A.: "The Kolmogorov—Smirnov, Cràmèr-von Mises test", *Ann. Math. Stat.* **28** (1957), 823—838.
- [38] DARLING, D. A.—ROBBINS, H.: "Some further remarks on inequalities for sample sums", *Proc. Nat. Acad. Sci.* **60** (1968) 1175—1182.

- [39] DARLING, D. A.—ROBBINS, H.: "Some nonparametric sequential tests with power one", *Proc. Nat. Acad. Sci.* **61** (1968), 804—809.
- [40] DEMPSTER, A.: "Generalized D_n^+ statistics", *Ann. Math. Stat.* **30** (1959), 593—597.
- [41] DONSKEP, M. O.: "Justification and extension of Doob's heuristic approach to the Kolmogorov—Smirnov theorems", *Ann. Math. Stat.* **23** (1952), 277—281.
- [42] DOOB, J. L.: "*Stochastic processes*", Wiley, New York, 1953.
- [43] DOOB, J. L.: "Heuristic approach to the Kolmogorov—Smirnov theorems", *Ann. Math. Stat.* **20** (1949), 393—403.
- [44] DURBIN, J.: "The probability that the sample distribution function lies between two parallel straight lines", *Ann. Math. Stat.* **39** (1968), 398—411.
- [45] DURBIN, J.: "*Asymptotic distributions of some statistics based on the bivariate sample distribution function*", *Nonparametric Techniques in Statistical Inference*, ed. by M. L. Puri, Cambridge, 1970, 435—449.
- [46] DURBIN, J.: "Boundary crossing probabilities for the Brownian motion and Poisson processes and techniques for computing the power of the Kolmogorov—Smirnov test", *Journ. Appl. Prob.* **8** (1971), 431—453.
- [47] DVORETZKY, A.—KIEFER, J.—WOLFOWITZ, J.: "Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator", *Ann. Math. Stat.* **27** (1956), 642—669.
- [48] DWASS, M.: "The distribution of a generalized D_n^+ statistic", *Ann. Math. Stat.* **30** (1959), 1024—1028.
- [49] EICKER, F.: "A loglog-law for double sequences of random variables", *Zeitschr. Wahrsch.* **16** (1970), 107—133.
- [50] EICKER, F.: "On the probability that a sample distribution function lies below a line segment", *Ann. Math. Stat.* **41** (1970), 2075—2092.
- [51] FELLER, W.: "On the Kolmogorov—Smirnov limit theorems for empirical distributions", *Ann. Math. Stat.* **19** (1948) 177—189.
- [52] FINKELSTEIN, H.: "The law of the iterated logarithm for empirical distributions", *Ann. Math. Stat.* **42** (1971), 607—615.
- [53] FRASER, D. A. S.: "*Nonparametric Methods in Statistics*", Wiley, New York, 1957.
- [54] GLIVENKO, N.: "Sulla determinazione empirica delle leggi di probabilit ", *Giorn. dell' Istit. degli att.* **4** (1933), 92—99.
- [55] GUPTA, S. S.—PANCHAPAKESAN, S.: "On order statistics and some applications of combinatorial methods in statistics", *A Survey of Combinatorial Theory*, edited by J. N. Srivastava et al., North Holland Publishing Company, 1973, 217—250.
- [56] HAJEK, J.—SIDAK, Z.: "*Theory of Rank Tests*" Academic Press, New York, 1967.
- [57] HARTMAN, P.—WINTNER, A.: "On the law of the iterated logarithm", *Amer. J. Math.* **63** (1941), 169—176.
- [58] ISHII, G.: "Kolmogorov—Smirnov test in life test", *Ann. Inst. Stat. Math.* **10** (1958), 37—46.
- [59] ISHII, G.: "On the exact probabilities of R nyi's tests", *Ann. Inst. Stat. Math. Tokyo*, **11** (1959), 17—24.
- [60] KAC, M.: "On deviation between theoretical and empirical distribution", *Proc. Nat. Acad. Sci.* **35** (1949), 252—257.
- [61] KARLIN, S.: "*A first course in Stochastic Processes*", Academic Press, New York, 1966.
- [62] KIEFER, J.: "Iterated logarithm analogues for sample quantiles when $p_n \downarrow 0$ ", *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, 1970, Vol. I, 227—244.
- [63] KNOTT, M.: "The small sample power of one-sided Kolmogorov tests for a shift in location of the normal distribution", *Journ. Amer. Stat. Ass.* **65** (1970), 1384—1391.
- [64] KOLMOGOROV, A. N.: "Sulla determinazione empirica di una legge di distribuzione", *Giorn. dell'Istit. degli att.* **4** (1933), 83—91.
- [65] L SZL , Z.: "Egy teljesen v letlen megbizhat s gi jelleg  készlet-modell", *MTA III. Oszt. K zl.* **21** (1972), 77—117.
- [66] LAUWERIER, H. A.: "The asymptotic expansion of the statistical distribution of N. V. Smirnov", *Zeitschr. Wahrsch.* **2** (1963), 61—68.
- [67] MALMQUIST, S.: "On certain confidence contours for distribution functions", *Ann. Math. Stat.* **25** (1954), 523—533.
- [68] МАНУЯ, Г. М.: «Обобщение критерия А. Н. Колмогорова для оценки закона распределения по эмпирическим данным», *Докл. Акад. Наук*, **69** (1949), 495—497.

- [69] MASSEY, F. J.: "A note on the power of a non-parametric test", *Ann. Math. Stat.* **21** (1950), 440—443. Correction: *Ann. Math. Stat.* **23** (1952), 637—638.
- [70] NEF, W.: »Über die Differenz zwischen theoretischer und empirischer Verteilungsfunktion«, *Zeitschr. Wahrsch.* **3** (1964), 154—162.
- [71] NOÉ, M.—VANDEWIELE, G.: "The calculation of distributions of Kolmogorov—Smirnov type statistics including a table of significance points for a particular case", *Ann. Math. Stat.* **39** (1968), 233—241.
- [72] NOETHER, G. E.: "Note on the Kolmogorov-statistic in the discrete case", *Metrika* **7** (1963), 115.
- [73] PURI, M. L.—SEN P. K.: "*Nonparametric methods in multivariate analysis*", Wiley, New York, 1971.
- [74] PYKE, R.: "The supremum and infimum of the Poisson process", *Ann. Math. Stat.* **30** (1959), 568—576.
- [75] QUADE, D.: "On the asymptotic power of the one sample Kolmogorov—Smirnov tests", *Ann. Math. Stat.* **36** (1965), 1000—1018.
- [76] RAGHAVACHARI, M.: "Limiting distributions of Kolmogorov—Smirnov type statistics under the alternative", *Ann. Stat.* **1** (1973), 67—73.
- [77] RÉNYI, A.: "On the theory of order statistics", *Acta Math. Acad. Sci. Hung.* **4** (1953), 191—231.
- [78] RÉNYI, A.: „Valószínűségsszámítás”, 2. kiadás, Tankönyvkiadó, Budapest, 1968.
- [79] RÉNYI, A.: „A rendezett minták elméletének egy problémaköréről”, *MTA III. Oszt. Közl.* **18** (1968), 23—30.
- [80] ROBBINS, H.: "A one-sided confidence interval for an unknown distribution function", (Abstract), *Ann. Math. Stat.* **25** (1954), 409.
- [81] ROSENBLATT, J.: "Some modified Kolmogorov—Smirnov tests of approximate hypotheses and their properties", *Ann. Math. Stat.* **33** (1962), 513—524.
- [82] SAHLER, W.: "A survey on distribution-free statistics based on distances between distribution functions", *Metrika* **13** (1968) 149—169.
- [83] SARKADI, K.: "Combinatorial proof for a theorem of H. E. Daniels", *Coll. Math. Soc. János Bolyai*, **4** (1969), *Combinatorial theory and its applications*, 975—979.
- [84] SARKADI, K.: "On the exact distributions of statistics of Kolmogorov—Smirnov type", *Periodica Math. Hung.* **3** (1973), 9—12.
- [85] SCHMID, P.: "On the Kolmogorov and Smirnov limit theorems for discontinuous distribution functions", *Ann. Math. Stat.* **29** (1958), 1011—1027.
- [86] STANLEY, R. M.: "Boundary crossing probabilities for the Kolmogorov—Smirnov statistics", *Ann. Math. Stat.* **43** (1972), 664—668.
- [87] STECK, G. P.: "Rectangle probability that the uniform order statistics and the probability that the empirical distribution function lies between two distribution functions", *Ann. Math. Stat.* **42** (1971), 1—11.
- [88] SUZUKI, G.: "On the Glivenko—Cantelli theorem", *Ann. Inst. Stat. Math.* **18** (1966), 29—37.
- [89] SUZUKI, G.: "On exact probabilities of some generalised Kolmogorov's D-Statistics", *Ann. Inst. Stat. Math.* **19** (1967), 373—388.
- [90] Санов, И. Н.: «О вероятности больших отклонений случайных величин», *Мат. Сборник* **42** (81), (1957), 11—44.
- [91] Смирнов, Н. В.: «Об уклонениях эмпирической функции распределения», *Мат. Сборник* **6** (48), (1939), 3—26.
- [92] Смирнов, Н. В.: «Приближение законов распределения по эмпирическим данным», *Успехи Мат. Наук*, **10** (1944), 179—206.
- [93] Смирнов, Н. В.: «Вероятности больших значений непараметрических односторонних критериев согласия», *Труды Мат. Инст. им. Стеклова*, **LXIV**, (1961), 185—210.
- [94] TAKÁCS, L.: "The use of a ballot theorem in order statistics", *Journal Appl. Prob.* **1** (1964), 389—392.
- [95] TAKÁCS, L.: "Applications of a ballot theorem in physics and in order statistics", *Journ. Royal Stat. Soc., Series B*, **27** (1965), 130—137.
- [96] TAKÁCS, L.: "The distributions of some statistics depending on the deviations between empirical and theoretical distribution functions", *Sankhya, Series A*, **27** (1965) 93—100.
- [97] TAKÁCS, L.: "*Combinatorial methods in the theory of stochastic processes*", Wiley, New York, 1967.
- [98] TAKÁCS, L.: "On the comparison of a theoretical and an empirical distribution function", *Journ. Appl. Prob.* **8** (1971), 321—330.

- [99] TANG, S. C.: "Some theorems on the ratio of empirical distribution to the theoretical distribution to the theoretical distribution", *Pac. Journ.* **12** (1962), 1107—1114.
- [100] TOPSOE, F.: "On the Glivenko—Cantelli theorem", *Zeitschr. Wahrsch.* **14** (1970), 239—250.
- [101] VAN DER WAERDEN, B. L.: "Testing a distribution function", *Indag. Math.* **15** (1953), 201—207.
- [102] VINCZE, I.: „Kétváltozós empirikus eloszlásfüggvények eltéréséről”, *MTA III. Oszt. Közl.* **10** (1960), 361—372.
- [103] VINCZE, I.: „A Kolmogorov—Szmirnov és más nem-paraméteres próbák erőfüggvényéről”, *MTA III. Oszt. Közl.* **15** (1965), 97—105.
- [104] VINCZE, I.: "On Kolmogorov—Smirnov type distribution theorems", *Nonparametric Techniques in Statistical Inference*, ed. by M. L. Puri, Cambridge, 1970, 385—401.
- [105] VINCZE, I.: "On some results and problems in connection with statistics of the Kolmogorov—Smirnov type", *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability*, 1970, Vol. I. 459—470.
- [106] WALD, A.—WOLFOWITZ, J.: "Confidence limits for continuous distribution functions", *Ann. Math. Stat.* **10** (1939), 105—118.
- [107] WALSH, J. E.: "Bounded probability properties of Kolmogorov—Smirnov and similar statistics for discrete data", *Ann. Inst. Stat. Math.* **15** (1963), 153—158.
- [108] WHITTLE, P.: "Some exact results for one sided distribution tests of the Kolmogorov—Smirnov type", *Ann. Math. Stat.* **32** (1961), 499—505.
- [109] WICHURA, M. J.: "Some Strassen-type laws of the iterated logarithm for multiparameter stochastic processes with independent increments", *Ann. Prob.* **1** (1973), 272—296.

(Beérkezett: 1973. XII. 21.)