

A VÉLETLEN MINTAMÉRETŰ EMPIRIKUS FOLYAMAT

GYENGE KONVERGENCIÁJÁRÓL

Doktori értekezés

Készítette:

C s ö r g ő S á n d o r

ösztöndijas gyakornok

József Attila Tudományegyetem

Bolyai Intézet

Szeged

1971



Diss. B 800



Köszönetemet fejezem ki dr. Tandori Károly professzor urnak a dolgozat megírásában nyújtott segítségéért és bátyámnak, dr. Csörgő Miklós professzornak, aki a kutatás e területére bevezetett és akivel való közös munka szellemében a jelen dolgozat született.

1 §. Az empirikus felvétel. Bevezetés

Jelentsen $(\Omega, \mathcal{A}, P)$, a dolgozatban mindvégig, egy rögzített Kolmogorov-féle valószínűségi mezőt. $\xi_1, \xi_2, \dots, \xi_n$ legyenek független, azonos eloszlású valószínűségi változók, melyek közös eloszlásfüggvénye $F(x) = P\{\xi_1 \leq x\}$, $(i=1, \dots, n; n \text{ tetszőleges természetes szám})$. Tehát statisztikai szóhasználatnál, tekintettünk egy $F(x)$ eloszlású statisztikai sokaságból vett n elemű mintát. E minta tapasztalati vagy empirikus eloszlásfüggvényén az

$$F_n(x) = \sum_{i: \xi_i \leq x} \frac{1}{n}$$

(véletlentől függő) függvényt szokás érteni, ami a

$$\psi(x) = \begin{cases} 1, & \text{ha } x \geq 0 \\ 0, & \text{ha } x < 0 \end{cases}$$

függvény bevezetésével az

$$(1) \quad F_n(x) = \frac{1}{n} \sum_{i=1}^n \psi(x - \xi_i)$$

alakban is írható. Itt a $\psi_i^x = \psi(x - \xi_i)$ valószínűségi változók függetlenek, $M(\psi_i^x) = x$, $D^2(\psi_i^x) = x(1-x)$

és így egyszerű számolással $M(F_n(x)) = F(x)$,

$$\text{Cov}(F_n(x), F_n(y)) = \frac{1}{n} (\min(F(x), F(y)) - F(x)F(y)) \quad \cdot$$

• $F_n(x)$ (1)-beli egyszerű alakjából egyből látható, hogy rá igazak a valószínűségszámítás klasszikus tételei. Így a nagy számok erős törvénye, ami szerint minden x -re

$$P \left\{ \lim_{n \rightarrow \infty} F_n(x) = F(x) \right\} = 1,$$

vagy ennek még erősebb alakja, a Glivenko-Cantelli tétel:

$$P \left\{ \lim_{n \rightarrow \infty} \sup_{-\infty < x < \infty} |F_n(x) - F(x)| = 0 \right\} = 1,$$

továbbá az iterált logaritmus tétel,

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

valamint a többdimenziós centrális határeloszlástétel, miszerint

$$\left\{ \sqrt{n} (F_n(x_1) - F(x_1)), \dots, \sqrt{n} (F_n(x_k) - F(x_k)) \right\}$$

aszimptotikusan ($n \rightarrow \infty$, k rögzített) k -dimenziós normális eloszlást követ, melynek várható értékei nullák, kovarianciái pedig: $\min(F(x_i), F(x_j)) - F(x_i) F(x_j)$.

E tételek mindegyike - a formulákból leolvasható - nyilvánvaló statisztikai jelentőséggel bír, azonban éppen a matematikai statisztika egy problémája, a tiszta illeszke-

dés problémája miatt a tapasztalati eloszlásfüggvényre vonatkozó vizsgálatok velük nem merültek ki. Olyan statisztikai próbákat kerestek, melyekkel ellenőrizhető azon nullhipotézis, hogy egy statisztikai minta megadott, folytonos $F(x)$ eloszlásfüggvénnyel bíró statisztikai sokaságból származik-e. Ilyen próbákra nyújtanak lehetőséget az alábbi tételek. Bevezetve ugyanis a

$$D_n = \sqrt{n} \sup_{-\infty < x < \infty} |F_n(x) - F(x)|,$$

$$D_n^+ = \sqrt{n} \sup_{-\infty < x < \infty} (F_n(x) - F(x)),$$

ugynevezett Kolmogorov-Szmirnov statisztikákat és a

$$W_n^2 = n \int_{-\infty}^{\infty} (F_n(x) - F(x))^2 \varphi(F(x)) dF(x)$$

ugynevezett Cramér-von Mises-Szmirnov statisztikát - ahol az utóbbinál $\varphi(t)$ ($0 \leq t \leq 1$) egy a konkrét problémának megfelelő nemnegatív folytonos sulyfüggvény -, igazak a következő állítások:

1. TÉTEL (KOLMOGOROV, 1933) :

$$\lim_{n \rightarrow \infty} P \{ D_n < x \} = \sum_{k=-\infty}^{\infty} (-1)^k e^{-2k^2 x^2},$$

2. TÉTEL (SZMIRNOV, 1939) :

$$\lim_{n \rightarrow \infty} P \{ D_n^+ < x \} = 1 - e^{-2x^2},$$

3. TÉTEL (SZMIRNOV, 1937) :

$$\lim_{n \rightarrow \infty} P \{ W_n^2 < x \} = 1 - \frac{1}{\pi} \sum_{k=1}^{\infty} (-1)^{k-1} \int_{\lambda}^{\lambda^{2k}} \frac{e^{-\frac{xy}{2}}}{\sqrt{-y^2 D(y)}} dy ,$$

ahol $D(y)$ a $k(s,t) = \varphi(s)\varphi(t) [\min(\varphi(s), \varphi(t)) - \varphi(s)\varphi(t)]$ magfüggvényhez tartozó Fredholm determináns, $\lambda_i (i=1,2,\dots)$ pedig (egyszeres) gyökei $D(y)$ -nak.

Felmerülhet továbbá a Kolmogorov-Szmirnov-féle D_n és D_n^+ statisztikák "súlyozásának" a kérdése is. Az erre vonatkozó eredmények rövid összefoglalása céljából vezessük be a következő jelöléseket:

$$R_n = \sqrt{n} \sup_{\alpha < x < \beta} \frac{|F_n(x) - F(x)|}{\mathcal{J}(F(x))} ,$$

$$R_n^+ = \sqrt{n} \sup_{\alpha < x < \beta} \frac{F_n(x) - F(x)}{\mathcal{J}(F(x))} ,$$

ahol $\mathcal{J}(t)$ ($0 \leq t \leq 1$) nemnegatív folytonos függvény, (α, β) pedig olyan intervallum (az egész egyenes, félegyenes, zárt intervallum) amelyen a $\mathcal{J}(F(x))$ -szel való osztásnak értelme van.

4. TÉTEL (ANDERSON és DARLING, 1952; RÉNYI, 1953):

$$\lim_{n \rightarrow \infty} P \{ R_n < x \} = G_{\mathcal{J}}(x) ,$$

$$\lim_{n \rightarrow \infty} P \{ R_n^+ < x \} = H_{\mathcal{J}}(x) .$$

RÉNYI [21] a $\mathcal{J}(t)=t$ és $\mathcal{J}(t)=1-t$ esetre vonatkozóan megadta $G_{\mathcal{J}}$ és $H_{\mathcal{J}}$ explicit alakját, ANDERSON és DARLING [2] pedig bebizonyították a tételbeli limeszek létezését minden $\mathcal{J}(t)$ nemnegatív folytonos függvényre és a statisztikai szempontból fontos $\mathcal{J}(t)=\sqrt{t(1-t)}$ súlyfüggvény esetén kiszámították $G_{\mathcal{J}}(x)$ Laplace transzformáltját, amelyet invertálni azóta sem sikerült.

A Kolmogorov-Szmirnov, Cramér-von Mises próbákkal kapcsolatban az említetteken kívül még igen sok publikáció jelent meg, DARLING [12] 1957-ben 89-et sorolt fel a teljesség igénye nélkül, és az utóbbi tizenöt év termése is jelentős. Általában a rendezett minták elméletében jelentős lépés volt az, amikor GNYEGYENKO-nak és munkatársainak sikerült a D_n és D_n^+ valószínűségi változók pontos eloszlását kiszámolniuk az ötvenes évek elején. Ugy lehetne itélni, hogy ezekkel az eredményekkel az előbb felsorolt határeloszlástételek statisztikai alkalmazásának köre leszűkült, ennek ellenére ezek a tételek elvi fontosságuk miatt továbbra is a valószínűségelméleti kutatások középpontjában maradtak és számos új bizonyítás született rájuk. Azért, hogy a felsorolt négy tétel bizonyításának egy, a dolgozat további részeiben (3-6 §) tárgyalandó újabb elmélet kialakulásában is döntő szerepet vivő módszerét körvonalazni tudjuk, a következő ismert megjegyzést tesszük.

Amint az 1., 2. és 3. tételek esetén látható, a határeloszlásfüggvény független a kiindulási $F(x)$ -től, ugyanez érvényes a 4. tétel esetén is. Emel azonban még

több is igaz, nevezetesen az, hogy már a D_n , D_n^+ , W_n^2 , R_n és R_n^+ statisztikák maguk is eloszlásmentesek, azaz eloszlásuk nem függ $F(x)$ -től. Például W_n^2 esetén, ha a nullhipotézis igaz, akkor 1 valószínűséggel (alkalmazva $F(x)$ folytonosságát)

$$\begin{aligned} W_n^2 &= n \int_{-\infty}^{\infty} \left(\frac{1}{n} \sum_{i=1}^n \psi(x - \xi_i) - (F(x))^2 \varphi(F(x)) \right) dF(x) = \\ &= n \int_0^1 \left(\frac{1}{n} \sum_{i=1}^n \psi(t - F(\xi_i)) - t^2 \varphi(t) \right) dt, \end{aligned}$$

ahol az $F(\xi_i)$ ($i=1, \dots, n$) valószínűségi változók függetlenek és a $[0,1]$ intervallumon egyenletes eloszlásúak.

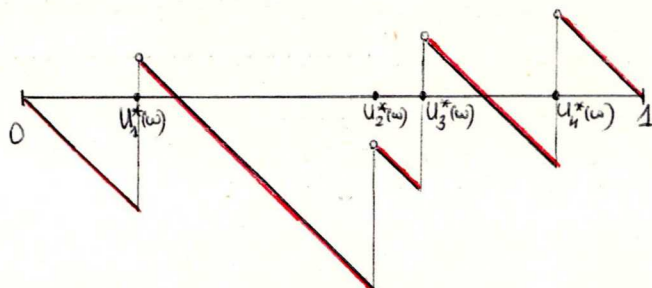
Legyenek tehát $U_1, \dots, U_n$ a $[0,1]$ intervallumon egyenletes eloszlású független valószínűségi változók és legyen $F_n(t)$ ($0 \leq t \leq 1$) tapasztalati eloszlásfüggvényük. Az

$$(2) \quad Y_n(t) = \sqrt{n}(F_n(t) - t)$$

sztochasztikus folyamatot empirikus folyamatnak szokás nevezni. Ez tehát minden rögzített t -re valószínűségi változó, az Ω eseménytér minden rögzített ω elemi eseményére pedig egy a $[0,1]$ intervallumon értelmezett függvény, az empirikus folyamat egy ugynevezett mintafüggvénye.

Illusztrációképp felrajzoljuk (rögzített $\omega \in \Omega$ mellett)

a $\sqrt{4}$ szorzó nélküli $\hat{Y}_n(t) = F_n(t) - t$ függvényt:



ahol $U_1^*, \dots, U_n^*$ a mintaelemek nagyság szerinti sorbarendezése. Az eloszlásmentességre tett fenti megjegyzés alapján

$$D_n = \sup_{0 \leq t \leq 1} |Y_n(t)|, \quad D_n^+ = \sup_{0 \leq t \leq 1} Y_n(t),$$

$$R_n = \sup_{\alpha \leq t \leq \beta} \frac{|Y_n(t)|}{J(t)}, \quad R_n^+ = \sup_{\alpha \leq t \leq \beta} \frac{Y_n(t)}{J(t)}$$

és

$$W_n^2 = \int_0^1 Y_n^2(t) J(t) dt.$$

Valamint egyszerű számolással

$$M(Y_n(t)) = 0, \quad 0 \leq t \leq 1;$$

$$M((Y_n(t) - Y_n(s))^2) = (t-s)(1-(t-s)), \quad 0 \leq s \leq t \leq 1.$$

DOOB 1949-ben írt alapvető [15] dolgozatában a következőképpen járt el. Tekintsünk egy $\{W^0(t), 0 \leq t \leq 1\}$ sztochasztikus folyamatot, amely a következő kikötéseknek tesz eleget:

- a) minden j természetes számra, ha $0 \leq t_1 < \dots < t_j \leq 1$, a $W^0(t_1), \dots, W^0(t_j)$ valószínűségi változók együttes eloszlása Gauss-féle (azaz normális);
- b) $M(W^0(t)) = 0, \quad 0 \leq t \leq 1;$
 $M(W^0(t) - W^0(s))^2 = (t-s)(1-(t-s)), \quad 0 \leq s \leq t \leq 1;$
- c) $P\{W^0(0) = 0\} = 1.$

(Ilyen folyamat van, mert $(t+1) W^0(\frac{t}{t+1}) = W(t)$ az ismert Wiener-féle folyamat, vagy más néven Brown-mozgás).

Ily módon, az előbb említett többdimenziós centrális határeloszlás tétel miatt $(Y_n(t_1), \dots, Y_n(t_j))$ j -dimenziós eloszlása aszimptotikusan $(n \rightarrow \infty)$ $(W^0(t_1), \dots, W^0(t_j))$ eloszlásával egyezik meg, a 3. §-ban bevezetendő terminológiával élve, az $Y_n(t)$ folyamat végesdimenziós eloszlásai a $W^0(t)$ folyamat végesdimenziós eloszlásaihoz konvergálnak. És itt tette DOOB, a 3. §-tól tárgyalandó elmélet kialakítását döntő mód befolyásoló heurisztikus feltevését:

az $Y_n(t)$ folyamat bizonyos "jó" funkcionáljainak (ilyen "jó funkcionáloknak" tartotta a $h_1(x) = \sup_{0 \leq t \leq 1} |x(t)|$ és a $h_2(x) = \sup_{0 \leq t \leq 1} x(t)$ funkcionálokat) határeloszlásának kiszámításánál az $Y_n(t)$ folyamat egyszerűen $W^0(t)$ -vel helyettesíthető. E feltevését természetesen az is motíválta, hogy $\sup_{0 \leq t \leq 1} |W^0(t)|$ és $\sup_{0 \leq t \leq 1} W^0(t)$ eloszlásait kiszámítva pontosan az 1. illetve a 2. tételben szereplő eloszlásfüggvényeket kapta.

DOOB heurisztikáját először DOWSKER [14] bizonyította be szabatosan, 1952-ben, meghatározva a "jó funkcionálok" egy elég tág osztályát. Eredményére azóta több új bizonyítás látott napvilágot, és a jelen disszertáció 4. §-a egy újabb, megítélésünk szerint egyszerű bizonyítást adva, ezek számát egyből gyarapítja.

2. §. Véletlen elemszámu statisztikai minták, véletlen

indexek. A probléma

Igen sok alkalmazott valószínűségi modellben független, azonos eloszlásu valószínűségi változók valamely x_n sorozatával kell bánni. Ha egy ilyen modellen belül valamilyen statisztikai következtetést akarunk levonni, ezt nagyon gyakran megfigyeléseknek egy olyan $x_1, \dots, x_N$ sorozatára támaszkodva kell megtennünk, ahol N , a minta mérete, maga is valószínűségi változó. Például N lehet egy adott időperiódusban lehetséges megfigyeléseknek a száma vagy rögzített pénzügyi keret mellett végezhető kísérletek száma. Ahogy PYKE [20] is írja, ilyen esetekben egyáltalán nem szokatlan, hogy a statisztikusok az 1. §-ban ismertett tételeken nyugvó hagyományos eljárásokat alkalmazzák, tekintet nélkül a mintaméret véletlen voltára, még akkor is, ha ez a véletlen mintaméret nem független magától a mintától. Éppen ezért fontos teoretikus igényként merülhet fel a használt eljárások vizsgálata abban az esetben, amikor a minta mérete is véletlen. A jelenlegi dolgozat célja az, hogy az 1. §-ban felsorolt tételek érvényességét ilyen körülmények között is bebizonyítsa.

ANSCOMBE [3] dolgozatának 1952-ben való megjelenése óta igen sok vizsgálat folyt ezen a téren, főleg az $x_1 + \dots + x_N$ összeg és $\max(x_1, \dots, x_N)$ maximumokra vonatko-

zó eljárásokkal kapcsolatban (például [3], [23], [18], [19], [10] illetve [4], [24], [11]). Az egész kérdés általában a következő módon vethető fel. Legyen η_n valószínűségi változók egy sorozata, $H(x)$ egy eloszlásfüggvény. Ismert, hogy $H(x)$ minden folytonossági pontjában

$$(3) \quad \lim_{n \rightarrow \infty} P\{\eta_n < x\} = H(x),$$

vagyis az $\{\eta_n\}$ sorozat egy határeloszlástételnek tesz eleget. ν_n legyen minden n -re pozitív egész értékű valószínűségi változó. Milyen körülmények között következik (3)-ból, hogy

$$\lim_{n \rightarrow \infty} P\{\eta_{\nu_n} < x\} = H(x) ?$$

Természetes követelmény, hogy ν_n legalábbis sztochasztikusan (az ilyen konvergenciát a következőkben $\xrightarrow{P}$ -vel jelöljük) ∞ -hez tartson. Ez, abban a speciális esetben, amikor ν_n (minden n -re) független az összes η_i -től (az $\{\eta_i\}$ -re tett minden külön feltevés nélkül) elegendő is, hiszen ekkor

$$P\{\eta_{\nu_n} < x\} = \sum_{i=1}^{\infty} P\{\eta_i < x\} P\{\nu_n = i\},$$

ahol $p_{n_i} = P\{\nu_n = i\}$ egy Toeplitz típusú mátrix. Általában, vagyis ha a függetlenséget feltenni módunkban nincs, ν_n

sztochasztikus végtelenbe-tartásának mikéntjére megszorításokat kell tennünk. Erre utal BILLINGSLEY [5] következő példája is. Legyen ξ_n független valószínűségi változók olyan sorozata, hogy mindegyik ξ_1 $a+1$ és -1 értékeket veszi fel $1/2$, $1/2$ valószínűséggel. A centrális határeloszlás tétele miatt ekkor

$$\lim_{n \rightarrow \infty} P \left\{ \sum_{i=1}^n \xi_i / \sqrt{n} < x \right\} = \phi(x),$$

ahol $\phi(x)$ a standard normális eloszlásfüggvény. Ha most ν_n az n -edik olyan k index, amelyre $\sum_{i=1}^k \xi_i = 0$, akkor,

minthogy $\nu_n \geq \nu_{n-1} + 1$, $\nu_n \xrightarrow{P} \infty$ és ugyanekkor $\sum_{i=1}^{\nu_n} \xi_i / \sqrt{n} = 0$. Ezért aztán a ν_n viselkedésére

vonatkozó kikötések olyanok szoktak lenni, amelyek a ν_n/n , vagy általánosabban a $\nu_n/\omega(n)$ hányadosok sztochasztikusan konvergens voltát tételezik fel, ahol $\omega(n)$ valamilyen, $n \rightarrow \infty$ esetén végtelenbe divergáló számelméleti függvény.

Az első eredményt, mint említettük ANSCOMBE [3] érte el a $\xi_1 + \dots + \xi_n$ összegekkel kapcsolatban, nevezetesen a centrális határeloszlástétel Lindeberg-Lévy-féle alakját általánosította a most tárgyalt irányba, feltételezve, hogy $\nu_n/n \xrightarrow{P} 1$. Az ő eredményét RÉNYI [23] általánosította arra az esetre, amikor $\nu_n/n \xrightarrow{P} \nu$, ahol ν pozitív, diszkrét valószínűségi változó, majd a BLUM,

HANSON, ROSENBLATT [6] hármas és MOGYORÓDI [18], egymástól függetlenül, elejtették a $\mathcal{D}$ -re tett diszkréttségi kikötést. Végül MOGYORÓDI [19] bebizonyított egy, az összes most említett eredményeket mind tartalmazó általános tételt, amelynek a jelen dolgozat szerzője által adott enyhe általánosítását meg is fogalmazzuk, mivel a későbbiek során szükségünk lesz rá. Ehhez fel kell, hogy idézzük a következő definíciókat.

1. DEFINÍCIÓ (RÉNYI, 1958; [22]): Legyen η_n valószínűségi változóknak egy olyan sorozata, amelynek van határeloszlása, azaz van olyan $H(x)$ eloszlásfüggvény, hogy annak minden folytonossági pontjában

$$\lim_{n \rightarrow \infty} P\{\eta_n < x\} = H(x) .$$

Az η_n sorozatot keverőnek nevezzük, ha tetszőleges A pozitív valószínűségű eseményre $H(x)$ minden folytonossági pontjában fennáll

$$(4) \quad \lim_{n \rightarrow \infty} P\{\eta_n < x \mid A\} = H(x) .$$

5. TÉTEL (RÉNYI, 1958; [22]): Ahhoz, hogy $\{\eta_n\}$ keverő legyen szükséges és elegendő, hogy (4) az $A = \{\eta_k < x\}$ ($k=1, 2, \dots$ rögzített) eseményekkel fennálljon.

2. DEFINÍCIÓ (ANScombe, 1952; [3] - elnevezés [19]) :

Az η_n sorozat Anscombe feltételnek tesz eleget, ha bármely $\varepsilon > 0$ és $\delta > 0$ számokhoz van olyan $c > 0$ és n_0 természetes szám, hogy $n \geq n_0$ esetén

$$P \left\{ \max_{n(1-c) \leq m \leq n(1+c)} |\eta_m - \eta_n| \geq \varepsilon \right\} \leq \delta$$

Igaz a következő

A. TÉTEL ([10]). Tegyük fel, hogy teljesülnek a következő feltételek:

1) Az η_n valószínűségi változók, a d_n pozitív és m_n tetszőleges konstansok sorozatai olyanok, hogy az $(\eta_n - m_n)/d_n$ sorozat keverő egy $H(x)$ eloszlásfüggvénnyel,

2) η_n és d_n teljesítik a következő Anscombe típusú feltételt: Ha $\varepsilon > 0$ és $\delta > 0$, akkor van olyan $c > 0$ és n_0 , hogy $n \geq n_0$ esetén

$$P \left\{ \max_{n(1-c) \leq m \leq n(1+c)} |\eta_m - \eta_n| \geq d_n \varepsilon \right\} \leq \delta,$$

3) ν_n pozitív egész értékű valószínűségi változók olyan sorozata, hogy egy $\omega(n)$ ($n \rightarrow \infty$ esetén) ∞ -be tartó függvényre

$$\frac{\nu_n}{\omega(n)} \xrightarrow{P} \nu,$$

ahol ν pozitív valószínűségi változó.

Ekkor

$$(5) \quad \lim_{n \rightarrow \infty} P \left\{ \eta_{\nu_n} < d_{\nu_n} \frac{x + m_{\nu_n}}{\nu_n} \right\} = H(x)$$

$H(x)$ minden folytonossági pontjában teljesül.

1. MEGJEGYZÉS. (5)-nél még több is igaz, fennáll ugyanis, hogy $(\eta_{\nu_n} - m_{\nu_n})/d_{\nu_n}$ is keverő $H(x)$ -szel, sőt a stabilitás fogalmának bevezetésével még ennél is általánosabb állítás fogalmazható meg, ezekre azonban a jelen dolgozatban nem lesz szükség.

2. MEGJEGYZÉS. MOGYORÓDI [19]-beli tétele az A. tételből $m_n \equiv 0$, $d_n \equiv 1$ vételével adódik, ekkor az A. tétel 2) kikötése az Anscombe feltétel 2. definícióbeli alakjára redukálódik. Itt jegyezzük még meg, hogy ebben az esetben $(m_n \equiv 0, d_n \equiv 1)$ a 2) kikötés helyett elég azt feltenni, hogy bármely $\varepsilon > 0$ és $\vartheta > 0$ számokhoz legyen olyan $c > 0$ és n_0 , hogy $n \geq n_0$ esetén

$$P \left\{ \max_{n \leq m \leq n(1+c)} |\eta_m - \eta_n| \geq \varepsilon \right\} \leq \vartheta.$$

Kézenfekvőnek látszana az A. tételt közvetlenül alkalmazni az 1. §-beli négy tétel esetében is, hogy eljussunk ezeknek véletlen mintaméretes alakjához, vagyis, hogy az 1.-4. tételekben a szereplő $D_n, D_n^+, W_n^2, R_n, R_n^+$ statisztikák helyébe $D_{\nu_n}, D_{\nu_n}^+, W_{\nu_n}^2, R_{\nu_n}, R_{\nu_n}^+$ statisztikákat

írassunk. Amíg azonban valószínűségi változók összege illetve maximuma esetén az összeg linearitása illetve a maximum monotonitása miatt az A. tétel feltételei könnyen beláthatók, ezek, a tapasztalati eloszlásfüggvényre épült statisztikák sokkal bonyolultabb felépítésűek és vizsgálójuk természetesen jut el az empirikus folyamat gyenge konvergenciájának mélyebb kérdéséhez, amint ezt DOOB hozzáállásánál már láttuk, és a dolgozat további részeinél látni fogjuk.

A véletlen mintaméretű Kolmogorov-Szmirnov-Cramér-von Mises statisztikákkal foglalkozó eddigi munkák a következők. KAC [17] valamint ALLEN és BEEKMAN [1] a Kolmogorov-Szmirnov-féle D_λ és D_λ^+ statisztikákra bizonyítottak be határeloszlástételeket, abban az esetben, amikor λ -nak, a mintától független Poisson eloszlású véletlen mintaméretnek a várható értéke végtelenhez tart. PYKE [20], a kérdés felvetésének iménti szellemében az 1. §-beli 1., 2. és 3. tételeket általánosította úgy, hogy bennük D_n , D_n^+ és W_n^2 helyett D_{ν_n} , $D_{\nu_n}^+$ és $W_{\nu_n}^2$ szerepel, ahol, mint

ANSCOMBE-nál először, $\nu_n/n \xrightarrow{P} 1$. [8]-ban az 1. és 2. tételeket "randomizáltuk" úgy, mint PYKE (az övétől igen különböző módszerrel), de ahol már $\nu_n/n \xrightarrow{P} \nu$ (ν tetszőleges pozitív valószínűségi változó) és végül, a jelen dolgozattal egyidőben született [9]-ben, ugyanilyen ν_n -ekkel már mind a négy tétel véletlen mintaméretes változata kiadódott.

A jelen dolgozat célja az, hogy egy a [9]-ben használttól különböző általános módszerrel bebizonyítsa, hogy az 1. § mind a négy tétele igaz akkor is, ha a minta elemeinek a száma minden megfigyeléskor véletlen, vagyis, hogy e tételekben szereplő statisztikák n indexe mindenütt a véletlen ν_n -re cserélhető, azon általános kikötés mellett, hogy $\nu_n/\omega(n) \xrightarrow{P} \nu$, ahol ν pozitív valószínűségi változó, $\omega(n)$, $n \rightarrow \infty$ esetén, ∞ -be divergáló számsorozat. A módszer általánosságából és további kisebb következményekből le lehet majd vonni azt a kvalitatív jellegű végkövetkeztetést, hogy a megfigyelések számának minden határon túl való növelésekor a véletlen mintaméretű empirikus folyamat ugyanugy viselkedik, mint az eredeti empirikus folyamat.

3. §. Valószínűségi mértékek metrikus tereken.

a C és D terek *)

Legyen S egy metrikus tér és $\mathcal{P}$ pedig az S -beli Borel halmazok összessége, azaz $\mathcal{P}$ az S nyitott halmazai által generált σ -algebra. Ha $\mathbb{P}_n$ ($n=1,2,\dots$) és $\mathbb{P}$ olyan az $\mathcal{P}$ -en értelmezett valószínűségi mértékek, hogy minden az S -en értelmezett korlátos és folytonos f függvényre

$$\int_S f d\mathbb{P}_n \xrightarrow{(n \rightarrow \infty)} \int_S f d\mathbb{P},$$

akkor azt mondjuk, hogy $\mathbb{P}_n$ gyengén konvergál $\mathbb{P}$ -hez ($\mathbb{P}_n \Rightarrow \mathbb{P}$). Ez különben akkor és csak akkor teljesül, ha minden olyan $\mathcal{P}$ -beli A -ra, amelynek ∂A határára $P(\partial A)=0$, igaz az, hogy $\mathbb{P}_n(A) \longrightarrow \mathbb{P}(A)$, ha $n \rightarrow \infty$.

$(S, \mathcal{P})$ mellett tekintsünk egy másik ilyen párt is, $(S', \mathcal{P}')$ -t és legyen h az S -nek S' -be való mérhető leképezése, vagyis, ha $A \in \mathcal{P}'$, akkor $h^{-1}A \in \mathcal{P}$. Ekkor bármely az $(S, \mathcal{P})$ -en értelmezett $\mathbb{P}$ valószínűségi mértékre a h leképezés egy egyértel

*) A valószínűségszámítás e modern fejezetét, DOOB, DONSKER, ERDŐS és KAC valamint KOLMOGOROV első cikkei után lényegében PROHOROV és SZKOROHOD dolgozták ki az ötvenes évek közepétől kezdve. A dolgozat e §-a BILLINGSLEY 1968-ban megjelent [5] könyvének első százötven oldala alapján foglalja össze (bizonyítások nélkül) a leglényegesebb és használatra kerülő ismereteket.

műen meghatározott $\mathcal{P}h^{-1}$ mértéket értelmez $(S, \mathcal{F})$ -en a következő definícióval: $\mathcal{P}h^{-1}(A) := \mathcal{P}(h^{-1}A)$, ha $A \in \mathcal{F}$.

A következő fontos tételt, amelyet euklidesi terek esetére először MANN és WALD bizonyítottak 1943-ban, igen gyakran fogjuk használni.

6. TÉTEL. Ha $\mathcal{P}_n \Rightarrow \mathcal{P}$ és h szakadási pontjai halmazát D_h -vel jelölve, $\mathcal{P}(D_h) = 0$, akkor $\mathcal{P}_n h^{-1} \Rightarrow \mathcal{P}h^{-1}$. (Speciálisan minden folytonos h -ra igaz az állítás)

Egy $(S, \mathcal{F})$ -en adott $\mathcal{P}$ valószínűségi mértéket feszésnek nevezünk, ha bármely pozitív ε -hoz van olyan $K \in \mathcal{F}$ kompakt halmaz, hogy $\mathcal{P}(K) > 1 - \varepsilon$. Igaz a következő ULAM-tól származó

7. TÉTEL. Ha S szeparábilis és teljes, akkor minden, az $(S, \mathcal{F})$ -en értelmezett valószínűségi mérték feszés.

Legyen Π az $(S, \mathcal{F})$ -en értelmezett valószínűségi mértékeknek valamely halmaza. Π -t relative kompaktnak nevezük, ha Π elemeinek bármely sorozata tartalmaz gyengén konvergens részsorozatot. Nyilván relative kompakt például valószínűségi mértékeknek egy gyengén konvergens sorozata. Π -t feszésnek nevezük, ha bármely pozitív ε -hoz van olyan $K \in \mathcal{F}$ kompakt halmaz, hogy bármely Π -beli $\mathcal{P}$ mértékre $\mathcal{P}(K)$

Az utóbbi két fogalom közötti kapcsolatra vonatkozik PROHOROV következő két tétele, melyek jelentősége alapvető.

8. TÉTEL. Ha Π fesszes, akkor relative kompakt.

9. TÉTEL. Tegyük fel, hogy az S alaptér teljes és szeparábilis. Ha Π relative kompakt, akkor fesszes.

Tekintsük most $(\Omega, \mathcal{A}, P)$ valószínűségi mezőnek az S metrikus térbe való X leképezését. Ha X mérhető $(X^{-1}\mathcal{G} \subset \mathcal{A})$, akkor ez $X(\omega)$ függvényt az S véletlen elemé-nek nevezzük (Speciálisan, ha $S = \mathbb{R}^1 - \mathbb{R}^k$ -val a k dimenziós euklidesi teret jelölve-, akkor X a hagyományos valószínűségi változó, ha $S = \mathbb{R}^k$, akkor k -dimenziós valószínűségi vektortváltozó).

X eloszlásán azt az $(S, \mathcal{F})$ -en egyértelműen meghatározott $\mathcal{P} = PX^{-1}$ valószínűségi mértéket értjük, amelyet a $\mathcal{P}(A) = PX^{-1}(A) = P(X^{-1}A) = P\{\omega \in \Omega : X(\omega) \in A\}$, $(A \in \mathcal{F})$ egyenlőséggel értelmezzük. Azt mondjuk, hogy véletlen elemeknek egy X_n sorozata az X véletlen elemhez eloszlásban konver-gál, $X_n \xrightarrow{\mathcal{D}} X$, ha az X_n és az X $\mathcal{P}_n$ és $\mathcal{P}$ eloszlásaira $\mathcal{P}_n \Rightarrow \mathcal{P}$. (Euklidesi térben a $\xrightarrow{\mathcal{D}}$ konvergencia egybeesik az eloszlásfüggvények hagyományos értelemben vett konvergenciájával.)

Rátérünk most a címbeli C tér ismertetésére. Legyen $C = C[0,1]$ a $[0,1]$ intervallumon folytonos függvények metrikus tere, ahol két $x, y \in C$ elem közötti távolságot a $\sup_{0 \leq t \leq 1} |x(t) - y(t)|$ mennyiséggel értelmezzük, amivel C teljes és szeparábilis. Jelölje $\mathcal{C}$ a C Borel halmazainak összességét. A $[0,1]$ valamely $t_1, t_2, \dots, t_k$ pontjaira je-

lentsse $h_{t_1, \dots, t_k}$ azt a leképezést, amely a C tér valamely x elemének az R^k euklidesi tér $(x(t_1), x(t_2), \dots, x(t_k))$ vektorát felelteti meg. Ekkor a $(C, \mathcal{C})$ -n értelmezett $\mathcal{P}$ valószínűségi mérték k -dimenziós eloszlásán az R^k Borel halmazainak σ -algebráján egyértelműen meghatározott $\mathcal{P}_{h_{t_1, \dots, t_k}}^{-1}$ mértéket és általában végesdimenziós eloszlásain a megfelelő mértékeket értjük. Ha mármint $\mathcal{P}_n$ és $\mathcal{P}$ a $(C, \mathcal{C})$ -n értelmezett olyan valószínűségi mértékek, hogy $\mathcal{P}_n \Rightarrow \mathcal{P}$, akkor bármely $t_1, t_2, \dots, t_k \in [0, 1]$ esetén, lévén $h_{t_1, \dots, t_k}$ folytonos C -n, a 6. tétel miatt $\mathcal{P}_{n, h_{t_1, \dots, t_k}}^{-1} \Rightarrow \mathcal{P}_{h_{t_1, \dots, t_k}}^{-1}$.

Könnyű példát adni arra, hogy ez fordítva nem áll, vagyis abból, hogy $\mathcal{P}_n$ végesdimenziós eloszlásai a $\mathcal{P}$ végesdimenziós eloszlásaihoz tartanak, még nem következik, hogy $\mathcal{P}_n \Rightarrow \mathcal{P}$. Ha azonban még azt is feltesszük, hogy $\{\mathcal{P}_n\}$ relative kompakt, akkor már igen. A relativ kompaktság viszont, a C tér teljes és szeparábilis volta miatt, PROKHOROV 8. és 9. tételei szerint ekvivalens a fessséssel, úgyhogy a következő tétel áll:

10. TÉTEL. Ha $\mathcal{P}_n$ végesdimenziós eloszlásai a $\mathcal{P}$ megfelelő végesdimenziós eloszlásaihoz konvergálnak és $\{\mathcal{P}_n\}$ fesszes, akkor $\mathcal{P}_n \Rightarrow \mathcal{P}$.

Hogy ez a tétel tényleg használható legyen, a fessséget, azaz C kompakt halmazait kellene tudni megfelelően jellemezni, amihez viszont rendelkezésre áll a klasszikus Arzela-Ascoli-féle tétel, melynek segítségével tehát belátható a következő

11. TÉTEL. $\{P_n\}$ akkor és csak akkor fesszes, ha

a) minden pozitív η -hoz van olyan a , hogy $n \geq 1$ esetén

$$P_n \{x : |x(0)| > a\} < \eta,$$

b) minden $\varepsilon > 0$ és $\eta > 0$ számokhoz van olyan δ , $0 < \delta < 1$, és n_0 , hogy ha $n \geq n_0$, akkor

$$P_n \{x : w_x(\delta) > \varepsilon\} < \eta,$$

ahol $w_x(\delta) = w(x, \delta) = \sup_{|s-t| < \delta} |x(s) - x(t)|$, az $x(t)$

függvény ugynevezett folytonossági modulusa.

Mindkét tétel természetesen fogalmazható át a $(C, \mathcal{C})$ -véletlen elemeire és azok $\xrightarrow{\mathcal{D}}$ konvergenciájára, ami miatt a következőkben a C (és később a D) véletlen elemeit ugyanazzal a betűvel jelöljük mint C -n (vagy D -n) való eloszlásait. Egyik leginkább vizsgált mérték $(C, \mathcal{C})$ -na W Wiener-féle, amely olyan valószínűségi mérték, hogy egyrészt

$$W \{x : x(t) < \alpha\} = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\alpha} e^{-\frac{u^2}{2t}} du, \text{ ha } 0 < t \leq 1$$

és

$$W \{x : x(0) = 0\} = 1,$$

másrészt ezzel az $\{x_t, 0 \leq t \leq 1\}$ "sztochasztikus folyamat" független növekményű, azaz, ha $0 \leq t_0 \leq t_1 \leq \dots \leq t_k \leq 1$, akkor

$$W \left\{ x : x(t_i) - x(t_{i-1}) \leq \alpha_i; i=1, \dots, k \right\} = \prod_{i=1}^k W \left\{ x : x(t_i) - x(t_{i-1}) \leq \alpha_i \right\}.$$

Megállapodásunknak megfelelően W -vel jelöljük ugyancsak a C azon véletlen elemét, melynek eloszlása a Wiener mérték, azaz tehát W az $(\Omega, \mathcal{A}, P)$ -nek olyan C -be való leképezése, hogy

$$P \left\{ \omega : W(\omega) \in A \right\} = W(A), \quad A \in \mathcal{C},$$

ahol a jobboldali W az előbb értelmezett Wiener mérték.

$W_t(\omega)$ -val a véletlen függvény t -beli értékét jelöljük, $\{W_t : 0 \leq t \leq 1\}$ tehát olyan sztochasztikus folyamat, melynek mintafüggvényei bármely $\omega \in \Omega$ esetén folytonosak.

Ezt a folyamatot szokás Wiener-féle folyamatnak vagy Brown-mozgásnak nevezni. $M(W_t) = 0$, minden t -re és $M(W_s W_t) = s$, ha $s \leq t$.

Általában C egy X véletlen elemét Gauss-félének szokás nevezni, ha összes végesdimenziós eloszlásai normálisak.

Például W is Gauss-féle, vagy más, szintén szokásos szóhasználatnál W_t Gauss-folyamat. Egy X Gauss-féle véletlen elem C -n való eloszlását egyértelműen meghatározzák az $M(X_t)$ és $M(X_s X_t)$ várható értékek, hiszen ezek meghatározzák a véges-

dimenziós eloszlásokat. Az empirikus folyamat vizsgálatánál szükségünk lesz, mint az 1. §-ban már láttuk, egy olyan W^0 Gauss-féle véletlen elemre, melynek C -n való eloszlásai az $M(W_t^0) = 0$, $M(W_s^0 W_t^0) = \min(s, t)$ -st kikötésekkel vannak meghatározva. A C ilyen véletlen eleme W -ből megszerkeszthető, legyen ugyanis W^0 az a véletlen elem, mely minden $0 \leq t \leq 1$ -re

$$W_t^0 = W_t - tW_1.$$

Ez nyilván Gauss-féle és az is igaz, hogy $M(W_t^0) = 0$, valamint $M(W_s^0 W_t^0) = s(1-t)$, ha $s \leq t$. Ezenkívül $P\{W_0^0 = W_1^0 = 0\} = 1$. Ezt a $\{W_t^0 : 0 \leq t \leq 1\}$ sztochasztikus folyamatot Brown-féle hídnak vagy lekötött Brown-mozgásnak szokás nevezni.

Sok szép eredményt nyertek a C téren bizonyított gyenge konvergenciával olyan folyamatokra, amelyek mintafüggvénye folytonos. Céljainknak e tér nem fog megfelelni, mivel mint láttuk az empirikus folyamat mintafüggvényei (vagy ahogy még nevezni szokták, trajektóriái) nem folytonosak, elsőfajú szakadással bírnak. Bár meg lehet adni olyan folyamatot, amelynek folytonos mintafüggvényei szuprémumban, $1/\sqrt{n}$ hibától eltekintve ugyanazok mint az $Y_n(t)$ empirikus folyamatéi és ez új folyamat gyenge konvergenciáját C -n bizonyítva egyuttal az $Y_n(t)$ konvergenciáját is bizonyítjuk, mégis természetesebb $Y_n(t)$ -t saját terében, az elsőfajú szakadással bíró függvények terében vizsgálni.

Legyen tehát D a $[0,1]$ intervallumon értelmezett balról folytonos és jobboldali határértékkel is mindenütt rendelkező függvények tere. Hogy ez tényleg "tér" legyen, a C -hez hasonlóan, valamilyen metrikát kellene bevezetni, mindenesetre olyat, hogy a látható célok érdekében a Prohorov tételek kényelme rendelkezésre állhasson, tehát, hogy D ezzel a metrikával teljes és szeparábilis legyen, sőt ezen túlmenően, hogy a fessséget is használni tudjuk, még a kompakt halmazokat is jellemezni lehessen valahogy Arzela-Ascoli-féle módon. Ezt a problémát SZKOROHOD oldotta meg 1956-ban, a következő módon.

Jelentse Λ a $[0,1]$ intervallumnak önmagára való, szigorúan monoton növekvő folytonos leképezései összességét. Két $x, y \in D$ elem $d(x, y)$ távolsága ezek után legyen az olyan pozitív ε -ok infimuma, amelyekhez van olyan $\lambda \in \Lambda$, hogy $\sup_{0 \leq t \leq 1} |\lambda(t) - t| \leq \varepsilon$ és $\sup_{0 \leq t \leq 1} |x(t) - y(\lambda(t))| \leq \varepsilon$. Ez a d tényleg távolság és vele a D tényleg teljes és szeparábilis metrikus tér. Ha $x, x_n \in D$ esetén $\sup_t |x_n(t) - x(t)| \longrightarrow 0$, (sup itt és ezután mindig $\sup_{0 \leq t \leq 1}$ -et jelent) akkor $d(x_n, x) \longrightarrow 0$, fordítva azonban ez nem áll. Ha azonban $d(x_n, x) \longrightarrow 0$, és x (egyenletesen) folytonos, akkor már $\sup_t |x_n(t) - x(t)| \longrightarrow 0$, speciálisan, a C -re relativizált Szkorohod topológia megegyezik a C térnek a szuprénum-távolság által származtatott saját (ugynevezett uniform) topológiájával. Jelölje a továbbiak-

ban $\mathcal{D}$ a D Borel-halmazait. A kompakt halmazok Arzela-Ascoli típusu karakterizálásával itt is mód nyílt a feszség következő jellemzésére:

12. TÉTEL. A $(D, \mathcal{D})$ -n értelmezett valószínűségi mértékek egy P_n sorozata akkor és csak akkor feszes, ha

a) minden pozitív η -hoz van olyan $\underline{a}$, hogy minden n -re

$$P_n \{ x : \sup_t |x(t)| > \underline{a} \} < \eta .$$

b) bármely $\varepsilon > 0$ és $\eta > 0$ számokhoz van olyan δ ($0 < \delta < 1$) és egy n_0 egész, hogy $n \geq n_0$ esetén

$$P_n \{ x : w_x^*(\delta) > \varepsilon \} < \eta .$$

ahol a $w_x^*(\delta) = w^*(x, \delta)$ "modulus" ($0 < \delta < 1$) a következő:

$$w_x^*(\delta) = \inf_{\{t_i\}} \max_{0 < i \leq r} \sup_{s, t \in [t_{i-1}, t_i]} |x(s) - x(t)| ,$$

ahol az infimum $[0, 1]$ -beli pontoknak minden olyan véges $\{t_i\}$ rendszerére értendő, melyre

$$0 = t_0 < t_1 < \dots < t_r = 1 \quad \text{és} \quad t_i - t_{i-1} > \delta , \quad i=1, 2, \dots, r.$$

Egy $(D, \mathcal{D})$ -n értelmezett $\mathcal{P}$ valószínűségi mértékekre jelentse $T_{\mathcal{P}}$ mindazon $[0,1]$ -beli t pontok halmazát, melyre a (már a C térnél értelmezett) h_t , az egydimenziós euklidesi térre való projekció folytonos, $\mathcal{P}$ -nullamértékű halmazt alkotó D -beli függvényektől eltekintve, azaz

$$T_{\mathcal{P}} = \left\{ t \in [0,1] : \mathcal{P} \{ x : x(t) \neq x(t-) \} = 0 \right\}.$$

Ha most $\mathcal{P}_n$ és $\mathcal{P}$ a $(D, \mathcal{D})$ -n értelmezett valószínűségi mértékek és $\mathcal{P}_n \Rightarrow \mathcal{P}$, akkor a 6. tétel miatt

$$\mathcal{P}_n^{h_{t_1, \dots, t_k}^{-1}} \Rightarrow \mathcal{P}^{h_{t_1, \dots, t_k}^{-1}}, \text{ feltéve, hogy } t_1, t_2, \dots, t_k \in T_{\mathcal{P}}$$

Fordítva itt is csak a feszség kikötésével lesz érvényben az állítás, vagyis igaz a következő

13. TÉTEL. Ha $\mathcal{P}_n$ a $(D, \mathcal{D})$ -n értelmezett valószínűségi mértékek feszes sorozata és

$$\mathcal{P}_n^{h_{t_1, \dots, t_k}^{-1}} \Rightarrow \mathcal{P}^{h_{t_1, \dots, t_k}^{-1}},$$

hacsak $t_1, \dots, t_k \in T_{\mathcal{P}}$, akkor

$$\mathcal{P}_n \Rightarrow \mathcal{P}.$$

Még egy, a C és D terek kapcsolatát ugyancsak jól jellemző, a későbbiekben használandó tételt említünk.

14. TÉTEL. Ha a $(D, \mathcal{D})$ -n értelmezett valószínűségi mértékekre

a) minden pozitív η -hoz van olyan a , hogy

$$P_n \{x : |x(0)| > a\} < \eta, \quad n \geq 1,$$

b) bármely $\varepsilon > 0$ és $\eta > 0$ esetén van olyan δ , $(0 < \delta < 1)$ és n_0 , hogy $n \geq n_0$ esetén

$$P_n \{x : w_x(\delta) > \varepsilon\} < \eta,$$

akkor $\{P_n\}$ fesszes, sőt ha P olyan valószínűségi mérték, hogy P_n egy P_{n_k} részsorozata hozzá gyengén konvergál, akkor $P(C) = 1$.

E tétel alkalmazását sok esetben megkönnyíti a következő

15. TÉTEL. Ahhoz, hogy valószínűségi mértékek valamely P_n sorozata kielégítse a 14. tétel b) feltételét elegendő, hogy bármely $\varepsilon > 0$ és $\eta > 0$ esetén legyen olyan δ $(0 < \delta < 1)$ és n_0 , hogy ha $n \geq n_0$, akkor

$$P_n \left\{ x : \sup_{t \leq s \leq t+\delta} |x(s) - x(t)| > \varepsilon \right\} < \delta \eta$$

fennálljon minden $t \in [0,1]$ esetén (ami $t > 1 - \delta$ esetén
egy értendő, hogy a szuprémumot $0 \leq s \leq 1$ esetre megszo-
rítjuk).

Mivel $C \in \mathcal{D}$ és, mint említettük, a C -re megszo-
ritott Szkerohod-topológia C saját uniform-topológiájával
esik egybe, ezért ha $A \in \mathcal{D}$, akkor $A \cap C \in \mathcal{C}$. Így a W
Wiener-féle mérték a következő definícióval kiterjeszthető
 $(D, \mathcal{D})$ -re: ha $A \in \mathcal{D}$, akkor $W(A) := W(A \cap C)$. Pontosan ugyan-
így értelmezzük a $(D, \mathcal{D})$ téren a W^0 "Brown-féle hid" mérté-
ket is.

4. § Az empirikus folyamat gyenge konvergenciája

Tekintsük tehát az 1. §-ban értelmezett $Y_n(t)$ empirikus folyamatot. Y_n (minden n -re), mint $(\Omega, \mathcal{A})$ -nak $(D, \mathcal{D})$ -be való leképezése, a D véletlen eleme. (Ennek általában egy X leképezést tekintve az a szükséges és elegendő feltétele, hogy rögzített $\omega \in \Omega$ esetén $X(t, \omega)$ D -beli függvény legyen és rögzített $t \in [0, 1]$ esetén $X(t, \omega)$ valószínűségi változó legyen $(\Omega, \mathcal{A})$ -n — Y_n esetén e feltételek teljesülnek.) Tekintsük szintén W^0 -t, a lekötött Brown-mozgást, mint D véletlen elemét. Bebizonyítjuk, hogy

$$(5) \quad Y_n \xrightarrow{\mathcal{D}} W^0.$$

A bizonyítást a 13. tétel segítségével végezzük el. Eszerint, mivel, mint az 1. §-ban beláttuk, az Y_n végesdimenziós eloszlásai a W^0 megfelelő végesdimenziós eloszlásaihoz tartanak, már csak azt kell belátnunk, hogy Y_n feszes sorozatot alkot (véletlen elemek valamely halmazának feszsége alatt értelemszerűen az eloszlásukul szolgáló mértékek halmazának feszségét értjük). A jelen bizonyítás annyiban új, amennyiben ezt a feszséget új módon bizonyítja, az eddigi két monográfiában (GIHMAN-SZKOROHOD [16], BILLINGSLEY [5]) látottaknál egyszerűbben, amit BREIMAN [7] következő eredménye tesz lehetővé:

1. SEGÉDTÉTEL. Legyenek $\xi_1, \dots, \xi_n, \dots$ független exponenciális eloszlású valószínűségi változók, amelyek mindegyikének várható értéke 1 és legyen $Z_k = \xi_1 + \dots + \xi_k$. Továbbra is $(U_1^*, \dots, U_n^*)$ -al jelölve a $[0,1]$ intervallumon való egyenletes eloszlású sokaságból vett rendezett mintát igaz az, hogy $(U_1^*, \dots, U_n^*)$ együttes eloszlása megegyezik $(Z_1/Z_{n+1}, \dots, Z_n/Z_{n+1})$ együttes eloszlásával.

Szükség lesz a következő, PAUL LÉVY-től származó egyenlőtlenségre is ([5]):

2. SEGÉDTÉTEL. Ha $\eta_1, \dots, \eta_m$ független 0 várható értékű valószínűségi változók, $D^2(\eta_i) = D_i^2$ végesek $(i = 1, \dots, m)$, akkor

$$P \left\{ \max_{1 \leq i \leq m} \frac{\sum_{j=1}^i \eta_j}{\sqrt{D_1^2 + \dots + D_m^2}} \geq \lambda \right\} \leq 2 P \left\{ \frac{\sum_{j=1}^m \eta_j}{\sqrt{D_1^2 + \dots + D_m^2}} \geq \lambda - \sqrt{2} \right\}$$

Mivel $P \{ Y_n(0)=0 \} = 1$, Y_n feszességének bizonyításához, a 14. tétel szerint, elég annyit igazolni, hogy bármely $\varepsilon > 0$ és $\eta > 0$ esetén van olyan $\delta, 0 < \delta < 1$, és n_0 , hogy ha $n \geq n_0$ akkor

$$(6) \quad P \left\{ \sup_{|s-t| < \delta} |Y_n(s) - Y_n(t)| \geq \varepsilon \right\} \leq \eta,$$

amihez viszont, alkalmazva a 15. tételt, elegendő, hogy tetszőleges $0 \leq t \leq 1$ esetén adott ε -hoz és η -hoz meg tudjuk δ -t és n_0 -t úgy választani, hogy $n \geq n_0$ esetén

$$(7) \quad P \left\{ \sup_{t \leq s \leq t+\delta} |Y_n(s) - Y_n(t)| \geq \varepsilon \right\} \leq \eta \delta$$

is teljesüljön. Sőt, minthogy $Y_n(t)$ növekményei stationárius eloszlásúak, (7)-ben $t=0$ vehető. Azaz, amit elég bizonyítanunk az az, hogy adott ε -hoz és η -hoz van olyan δ ($0 < \delta < 1$) és n_0 , hogy

$$(8) \quad P \left\{ \sup_{s \leq \delta} |Y_n(s)| \geq \varepsilon \right\} \leq \delta \eta,$$

ha csak $n \geq n_0$. Bizonyításunk eddig megegyezik BILLINGSLEY bizonyításával, innét kezdve viszont eltér. (GIHMAN és SZKOROHOD (6)-ot látják be direktén, eléggé bonyolult módon.) (8) jobboldalát tekintve, az 1. segédétel felhasználásával kapjuk, hogy

$$P \left\{ \sup_{s \leq \delta} |Y_n(s)| \geq \varepsilon \right\} = P \left\{ \sqrt{n} \max_{\frac{k}{n} \leq \delta} |U_k^* - \frac{k}{n}| \geq \varepsilon \right\} =$$

$$= P \left\{ \sqrt{n} \max_{k \leq n\delta} \left| \frac{Z_k}{Z_{n+1}} - \frac{k}{n} \right| \geq \varepsilon \right\} =$$

$$= P \left\{ \frac{n}{Z_{n+1}} \max_{k \leq n\delta} \left| \frac{Z_k - k}{\sqrt{n}} - \frac{k}{n} \frac{Z_{n+1} - n}{\sqrt{n}} \right| \geq \varepsilon \right\}.$$

Bevezetve a $\xi_1 = S_1 - 1$ valószínűségi változókat, melyek függetlenek, azonos eloszlásúak 0 várható értékkel és 1 szórással, valamint az $S_k = \xi_1 + \dots + \xi_k$ jelölést, az utóbbi valószínűség a következő módon becsülhető:

$$\begin{aligned}
 & P \left\{ \frac{n}{Z_{n+1}} \max_{k \leq \delta n} \left| \frac{S_k}{\sqrt{n}} - \frac{k}{n} \left(\frac{S_n}{\sqrt{n}} + \frac{f_{n+1}}{\sqrt{n}} \right) \right| \geq \varepsilon \right\} = \\
 & = P \left\{ \frac{n}{Z_{n+1}} \sup_{s \leq \delta} \left| \left(\frac{S_{[ns]}}{\sqrt{n}} - s \frac{S_n}{\sqrt{n}} \right) + s \frac{f_{n+1}}{\sqrt{n}} \right| \geq \varepsilon \right\} \leq \\
 & \leq P \left\{ \frac{n}{Z_{n+1}} \sup_{s \leq \delta} \left| \frac{S_{[ns]}}{\sqrt{n}} - s \frac{S_n}{\sqrt{n}} \right| + \delta \frac{f_{n+1}}{\sqrt{n}} \geq \varepsilon \right\}.
 \end{aligned}$$

Mint hogy $M(f_n) = D(f_n) = 1$, így $f_{n+1}/n \xrightarrow{P} 0$ és a nagyszámok erős törvénye szerint $n/Z_{n+1} \longrightarrow 1$, 1 valószínűséggel, ezért a legutolsó valószínűség nem nagyobb mint

$$(9) \quad P \left\{ \sup_{s \leq \delta} \left| \frac{S_{[ns]}}{\sqrt{n}} - s \frac{S_n}{\sqrt{n}} \right| \geq \frac{\varepsilon}{2} \right\} + \Theta_n,$$

ahol $\Theta_n \longrightarrow 0$, ha $n \longrightarrow \infty$.

Viszont a Csebisev egyenlőtlenség alkalmazásával

$$\begin{aligned}
 & P \left\{ \sup_{s \leq \delta} \left| \frac{S_{[ns]}}{\sqrt{n}} - s \frac{S_n}{\sqrt{n}} \right| \geq \frac{\varepsilon}{2} \right\} \leq \\
 & \leq P \left\{ \sup_{s \leq \delta} \left| \frac{S_{[ns]}}{\sqrt{n}} \right| \geq \frac{\varepsilon}{4} \right\} + P \left\{ \sup_{s \leq \delta} \left| s \frac{S_n}{\sqrt{n}} \right| \geq \frac{\varepsilon}{4} \right\} \leq \\
 & \leq P \left\{ \max_{k \leq \delta n} \frac{|S_k|}{\sqrt{n}} \geq \frac{\varepsilon}{4} \right\} + P \left\{ \frac{|S_n|}{\sqrt{n}} \geq \frac{\varepsilon}{4\delta} \right\} \leq \\
 & \leq P \left\{ \max_{k \leq \delta n} \frac{|S_k|}{\sqrt{[\delta n]}} \geq \frac{\varepsilon}{4\sqrt{\delta}} \right\} + \frac{16\delta^2}{\varepsilon^2}.
 \end{aligned}$$

Ha most $\delta < \frac{\varepsilon^2}{128}$, akkor $\frac{\varepsilon}{4\sqrt{\delta}} > 2\sqrt{2}$ és így

$\frac{\varepsilon}{8\sqrt{\delta}} < \frac{\varepsilon}{4\sqrt{\delta}} - \sqrt{2}$, tehát a P. Lévy egyenlőtlenség (2. segédttétel) alkalmazásával kapjuk, hogy az utóbbi kifejezés nem nagyobb, mint

$$2 P \left\{ \frac{|S[\delta n]|}{\sqrt{[\delta n]}} \geq \frac{\varepsilon}{8\sqrt{\delta}} \right\} + \frac{16\delta^2}{\varepsilon^2}.$$

Viszont a centrális határeloszlástétel szerint ez utóbbi $n \rightarrow \infty$ esetén a következő kifejezéshez tart:

$$2 P \left\{ |N| \geq \frac{\varepsilon}{8\sqrt{\delta}} \right\} + 16 \frac{\delta^2}{\varepsilon^2},$$

ahol N standard normális eloszlású valószínűségi változó.

Mivel a köbös Csebisev egyenlőtlenség szerint

$$P \left\{ |N| \geq \frac{\varepsilon}{8\sqrt{\delta}} \right\} \leq \frac{8^3 \delta^{3/2}}{\varepsilon^3} M(|N|^3),$$

így végül azt kaptuk, hogy $\delta < \varepsilon^2/128$ esetén

$$\lim_{n \rightarrow \infty} P \left\{ \sup_{s \leq \delta} |Y_n(s)| \geq \varepsilon \right\} \leq \frac{1024 \sqrt{\frac{2}{\pi}}}{\varepsilon^3} \delta^{3/2} + \frac{16}{\varepsilon^2} \delta^2,$$

és amennyiben $\delta < \min \left(\frac{\eta^2 \varepsilon^6 \pi}{8 \cdot 388 \cdot 608}, \frac{\eta \varepsilon^2}{32} \right)$, akkor

ez a jobboldal határozottan kisebb mint $\delta \eta$, ami ~~állíté~~ sünkát bizonyítja.

Mivel a

$$h_1(x) = \sup_{0 \leq t \leq 1} |x(t)|, \quad h_2(x) = \sup_{0 \leq t \leq 1} x(t),$$

$$h_3(x) = \int_0^1 x^2(t) dt, \quad h_4(x) = \inf_{0 \leq t \leq 1} x(t)$$

(ahogy BILLINGSLEY [5] és DARLING [12] is bizonyítják) a D -n mind folytonos funkcionálok, (5)-ből és a 6. tételből következik, hogy

$$\begin{aligned} \sup_t |Y_n(t)| &\xrightarrow{\mathcal{D}} \sup_t |W^0(t)|, \\ \sup_t Y_n(t) &\xrightarrow{\mathcal{D}} \sup_t W^0(t), \\ \inf_t Y_n(t) &\xrightarrow{\mathcal{D}} \inf_t W^0(t), \\ \int_0^1 Y_n^2(t) dt &\xrightarrow{\mathcal{D}} \int_0^1 (W^0(t))^2 dt, \end{aligned} \quad (10)$$

vagyis többek között az 1. § első három tétele is (az előző relációk jobboldalainak eloszlásait mind kiszámolták már külön-külön (pl. [5]), természetesen a már régebből ismert eredményekre jutva). Ha $h_5(x)$ azon $[0,1]$ -beli t pontok halmazának Lebesgue mértéke, melyekre $x(t) > 0$, akkor h_5 szintén folytonos a Szkorohod topológiában és így

$$h_5(Y_n) \xrightarrow{\mathcal{D}} h_5(W^0)$$

és minthogy BILLINGSLEY azt is belátta, hogy $h_5(w^0)$ a $[0,1]$ -en egyenletes eloszlású, így $0 \leq \alpha \leq 1$ esetén

$$\lim_{n \rightarrow \infty} P \left\{ h_5(Y_n) \leq \alpha \right\} = \alpha.$$

Az 1. §-nak azonban nemcsak az első három tétele, hanem a 4. tétel is következik (5)-ből. Ugyanis a

$$h_6(x) = \sup_{a \leq t \leq 1} \frac{|x(t)|}{t}, \quad h_7(x) = \sup_{a \leq t \leq 1} \frac{x(t)}{t},$$

$$h_8(x) = \sup_{0 \leq t \leq b} \frac{|x(t)|}{1-t}, \quad h_9(x) = \sup_{0 \leq t \leq b} \frac{x(t)}{1-t},$$

$$h_{10}(x) = \sup_{a \leq t \leq b} \frac{|x(t)|}{\sqrt{t(1-t)}}, \quad h_{11}(x) = \sup_{a \leq t \leq b} \frac{x(t)}{\sqrt{t(1-t)}}$$

funkcionálok mindegyike ($0 < a < b < 1$) szubadditív és pozitív homogén, úgyhogy akkor és csak akkor folytonosak (lásd [13], 10. oldal) ^{ha folytonosok} az $x(t) \equiv 0$ ($\in D$) pontban. Ezen funkcionálok mindegyike viszont tényleg folytonos $\in 0$ pontban, mivel ha x_n a Szkorohod topológiában 0 -hoz tart (ami egyenletesen folytonos függvény), oda tart egyenletesen, ^{is} amiből következik, hogy $h_i(x_n) \rightarrow 0 = h_i(0)$ ($n \rightarrow \infty$), $i=6, \dots, 11$. Úgyhogy a (10) alatti relációkhoz hasonlóak a $h_6, \dots, h_{11}$ funkcionálokra is felírhatók.

5. §. A véletlen mintaméretű empirikus folyamat gyenge konvergenciája

Rátérünk most a dolgozat fő eredményének ismertetésére. Fel fogjuk használni a következő segédtételeket.

3. SEGÉDTÉTEL (CRAMÉR és WOLD, 1936; [5]). Az R^k euklideszi tér $\eta_n = (\eta_{n1}, \dots, \eta_{nk})$ $n=1, 2, \dots$ és $\eta = (\eta_1, \dots, \eta_k)$ valószínűségi vektorváltozókra az

$$\eta_n \xrightarrow{\mathcal{D}} \eta$$

összefüggés akkor és csak akkor áll fenn, ha bármely $(c_1, \dots, c_k)$ valós vektorra

$$\sum_{i=1}^k c_i \eta_{ni} \xrightarrow{\mathcal{D}} \sum_{i=1}^k c_i \eta_i$$

4. SEGÉDTÉTEL (PYKE, 1968; [20]). Továbbra is az $(n$ -dik) empirikus folyamatot jelölve $Y_n(t)$ -vel, igaz, hogy

$$\lim_{y \rightarrow \infty} \lim_{r \rightarrow \infty} P \left\{ \max_{1 \leq j \leq r} \left(\frac{1}{r} \right)^{1/2} \sup_{0 \leq t \leq 1} |Y_j(t)| > y \right\} = 0.$$

Ez az eredmény igen fontos a jelen dolgozat szempontjából. Bizonyítása igen bonyolult és hosszú, bizonyos kétdimenziós paraméterhalmazzal rendelkező sztochasztikus folyamatok gyenge konvergenciájára vonatkozó eredményeket használ fel. (Kétdimenziós folyamatra azért van szükség, hogy az igen nehezen kezelhető $\max_j \sup_t$ kifejezés egyetlen - két indexes - $\sup_{s,t}$ kifejezéssel legyen helyettesíthető.)

Erre a 4. segédtételekre épül a következő

5. SEGÉDTÉTEL. Tetszőleges pozitív ε -hoz és η -hoz van olyan $c > 0$ és n_0 , hogy ha $n \geq n_0$, akkor

$$(11) \quad P \left\{ \max_{n \leq m \leq n(1+c)} \sup_{0 \leq t \leq 1} |Y_m(t) - Y_n(t)| > \varepsilon \right\} < \eta.$$

(azaz

$$\lim_{c \rightarrow 0} \lim_{n \rightarrow \infty} P \left\{ \max_{n \leq m \leq n(1+c)} \sup_{0 \leq t \leq 1} |Y_m(t) - Y_n(t)| > \varepsilon \right\} = 0).$$

BIZONYÍTÁS. Jelölje $F_{m-n}^i(t)$ az $(U_{n+1}, \dots, U_m)$ minta eloszlásfüggvényét, $Y_{m-n}^i(t)$ pedig az ehhez tartozó $\sqrt{m-n}(F_{m-n}^i(t) - t)$ empirikus folyamat. Ha $m > n$, akkor

$$Y_n(t) - Y_m(t) = n^{1/2} F_n(t) - m^{1/2} F_m(t) - t(n^{1/2} - m^{1/2}) =$$

$$= n^{-1/2} n F_n(t) - m^{-1/2} m F_m(t) - t(n^{1/2} - m^{1/2}) =$$

$$= n^{-1/2} n F_n(t) - m^{-1/2} (n F_n(t) + (m-n) F'_{m-n}(t)) - t(n^{1/2} - m^{1/2}) =$$

$$= n^{-1/2} n F_n(t) - m^{-1/2} n F_n(t) - t n^{1/2} + m^{-1/2} t n - m^{-1/2} t n =$$

$$- \frac{(m-n)^{1/2}}{(m-n)^{1/2}} m^{-1/2} (m-n) F'_{m-n}(t) + t m^{1/2} =$$

$$= n F_n(t) n^{-1/2} (1 - (\frac{n}{m})^{1/2}) - t n^{1/2} (1 - (\frac{n}{m})^{1/2}) - m^{-1/2} t n =$$

$$- (m-n)^{1/2} F'_{m-n}(t) (\frac{m-n}{m})^{1/2} + t m^{1/2} =$$

$$= (1 - (\frac{n}{m})^{1/2}) (n^{1/2} F_n(t) - n^{1/2} t) - (m-n)^{1/2} F'_{m-n}(t) (1 - \frac{n}{m})^{1/2} +$$

$$+ t \frac{(m-n)^{1/2}}{m^{1/2}} (m-n)^{1/2} =$$

$$= (1 - (\frac{n}{m})^{1/2}) Y_n(t) - (m-n)^{1/2} F'_{m-n}(t) (1 - \frac{n}{m})^{1/2} + (m-n)^{1/2} t (1 - \frac{n}{m})^{1/2} =$$

$$= (1 - (\frac{n}{m})^{1/2}) Y_n(t) - (1 - \frac{n}{m})^{1/2} Y'_n(t).$$

Ezért

$$\max_{n \leq m \leq n(1+c)} \sup_{0 \leq t \leq 1} |Y_n(t) - Y_m(t)| \leq$$

$$\leq \max_{n \leq m \leq n(1+c)} (1 - (\frac{n}{m})^{1/2}) \sup_{0 \leq t \leq 1} |Y_n(t)| + \max_{n \leq m \leq n(1+c)} (\frac{m-n}{m})^{1/2} \times$$

$$\times \sup_{0 \leq t \leq 1} |Y'_{m-n}(t)| \leq (1 - (\frac{n}{n(1+c)})^{1/2}) \sup_{0 \leq t \leq 1} |Y_n(t)| +$$

$$+ \max_{n \leq m \leq n(1+c)} (\frac{m-n}{n})^{1/2} \sup_{0 \leq t \leq 1} |Y'_{m-n}(t)| =$$

$$= (1 - (\frac{1}{1+c})^{1/2}) \sup_t |Y_n(t)| + \max_{1 \leq j \leq n(1+c)-n} (\frac{1}{n})^{1/2} \sup_t |Y'_j(t)|,$$

és így (11) baloldala nem nagyobb, mint

$$P \left\{ \sup_{0 \leq t \leq 1} |Y_n(t)| > \frac{\varepsilon}{2(1 - (\frac{1}{1+c})^{1/2})} \right\} +$$

$$+ P \left\{ \max_{1 \leq j \leq nc} (\frac{1}{n})^{1/2} \sup_{0 \leq t \leq 1} |Y'_j(t)| > \frac{\varepsilon}{2} \right\}.$$

Mivel $\sup_t |Y_n(t)|$ -nek van határeloszlása (1. tétel), így

(mivel $c \rightarrow 0$ esetén $\varepsilon/2(1 - (\frac{1}{1+c})^{1/2}) \rightarrow \infty$)

$$\lim_{c \rightarrow 0} \lim_{n \rightarrow \infty} P \left\{ \sup_{0 \leq t \leq 1} |Y_n(t)| > \frac{\varepsilon}{2(1 - (\frac{1}{1+c})^{1/2})} \right\} = 0,$$

és így elég azt bizonyítani, hogy

$$(12) \quad \lim_{c \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} P \left\{ \max_{1 \leq j \leq nc} \left(\frac{1}{n}\right)^{1/2} \sup_{0 \leq t \leq 1} |Y'_j(t)| > \frac{\varepsilon}{2} \right\} = 0.$$

Mint hogy az U_1 változók azonos eloszlásúak és így $Y'_j(t)$ -ről a vessző leghagyható, azaz a (12)-ben szereplő valószínűség egyenlő a következővel:

$$P \left\{ \max_{1 \leq j \leq nc} \left(\frac{1}{nc}\right)^{1/2} \sup_{0 \leq t \leq 1} |Y_j(t)| > \frac{\varepsilon}{2\sqrt{c}} \right\},$$

ezért (12) következik a 4. segédtételből $r=nc$ és $y = \frac{\varepsilon}{2\sqrt{c}}$ vételével.

Ha most minden n -re ν_n egész értékű valószínűségi változó, akkor az Y_{ν_n} véletlen mintaméretű empirikus folyamat a D térnek véletlen eleme, ugyanazon okoknál fogva, mint Y_n . Bebizonyítjuk a következő tételt.

B. TÉTEL. Y_{ν_n} -el jelölve a véletlen mintaméretű empirikus folyamatot és W^0 -al a Brown-féle hidat, ha ν_n pozitív egész értékű valószínűségi változóknak olyan soro-

zata, hogy egy $\omega(n)$ végtelenhez tartó számsorozatra

$$\frac{\nu_n}{\omega(n)} \xrightarrow{P} \nu,$$

ahol ν pozitív valószínűségi változó, akkor

$$Y_{\nu_n} \xrightarrow{\mathcal{D}} W^0.$$

BIZONYÍTÁS. A 13. tétel feltételeit fogjuk belát-
ni, vagyis, hogy I) Y_{ν_n} végesdimenziós eloszlásai a W^0
megfelelő végesdimenziós eloszlásaihoz konvergálnak és hogy
II) Y_{ν_n} feszes.

I) A W^0 -nak a $(D, \mathcal{D})$ -n való értelmezése miatt
 $T_{W^0} = [0, 1]$, úgyhogy tetszőleges $[0, 1]$ -beli "időpilla-
natokkal" kell foglalkoznunk. Legyen s tetszőleges $[0, 1]$ -be-
li szám. Be kell látnunk, hogy

$$(13) \quad Y_{\nu_n}(s) \xrightarrow{\mathcal{D}} W_s^0.$$

Ezt az A. tétel felhasználásával fogjuk bebizonyítani. Az
előző §-ban ugyanis láttuk, hogy $Y_n \xrightarrow{\mathcal{D}} W^0$, amiből követ-
kezik (a $h_s(x) = x(s)$ projekcióval), hogy $Y_n(s) \xrightarrow{\mathcal{D}} W_s^0$.
Ha most belátjuk, hogy Ia) $Y_n(s)$ keverő az $N(\cdot; 0, s(1-s))$
normális eloszlásfüggvénnyel és Ib) teljesíti az Anscombe ^{feltétel}
A. tételbeli alakját, akkor az A. tételből következik (13).

Ia) Figyelembe véve $P_n(t)$ -nek (1) alatti alakját,

$$Y_n(s) = \sqrt{n} \left(\frac{1}{n} \sum_{i=1}^n \psi(s-U_i) - s \right) = \frac{1}{\sqrt{n}} \sum_{i=1}^n (\psi(s-U_i) - s).$$

Rövidség kedvéért ψ_1 -vel jelölve most a $\psi(s-U_1)$ valószínűségi változót ($M(\psi_1)=s$, $D^2(\psi_1)=s(1-s)$), ahogy az imént említettük, tetszőleges $-\infty < x < \infty$ esetén

$$P \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n (\psi_i - s) < x \right\} \longrightarrow N(x; 0, s(1-s)),$$

ha $n \longrightarrow \infty$. Így van olyan n_0 , hogy ha $k \geq n_0$, akkor

$$P \left\{ \frac{1}{\sqrt{k}} \sum_{i=1}^k (\psi_i - s) < x \right\} > 0.$$

Az 5. tétel szerint így elég belátni, hogy $n \longrightarrow \infty$ esetén

$$(14) \quad P \left\{ \frac{1}{\sqrt{n}} \sum_{i=1}^n (\psi_i - s) < x \mid \frac{1}{\sqrt{k}} \sum_{i=1}^k (\psi_i - s) < x \right\} \longrightarrow N(x; 0, s(1-s))$$

minden rögzített $k \geq n_0$ mellett. Legyen $Y_n^*(s) = \frac{1}{\sqrt{n}} \sum_{i=p_n}^n (\psi_i - s)$,

ahol p_n egész, $p_n \longrightarrow \infty$, de olyan lassan, hogy $p_n/n \longrightarrow 0$ (pl. $p_n = [\log n]$). Ekkor a Cszebisev egyenlőtlenség alapján

$$P \left\{ |Y_n(s) - Y_n^*(s)| > \varepsilon \right\} = P \left\{ \frac{1}{\sqrt{n}} \left| \sum_{i=1}^{p_n} (\psi_i - s) \right| > \varepsilon \right\} \leq$$

$$\leq \frac{s(1-s)p_n}{\varepsilon^2 n} \longrightarrow 0,$$

így (14) igaz lesz, ha $n \longrightarrow \infty$ esetén

$$P \left\{ \frac{1}{\sqrt{n}} \sum_{i=p_n}^n (\psi_i - s) < x \mid \frac{1}{\sqrt{k}} \sum_{i=1}^k (\psi_i - s) < x \right\} \longrightarrow N(x; 0, s(1-s))$$

minden $k \geq n_0$ mellett teljesül. Ez így is van, mert ha n már olyan nagy, hogy $p_n \geq k \geq n_0$, akkor $Y_n(s)$ független $Y_k(s)$ -től.

Ib) Az 5. segédtételből következik, hogy bármely pozitív ε -hoz és η -hoz van olyan c és n_0 , hogy $n \geq n_0$ esetén

$$P \left\{ \max_{n \leq m \leq n(1+c)} |Y_n(s) - Y_m(s)| > \varepsilon \right\} < \eta,$$

és figyelembe véve az A. tétel utáni 2. megjegyzést, az A. tétel 2) kikötése is bizonyított. Azonban, mivel sem a 2. megjegyzés sem az 5. segédtételnek alapul szolgáló 4. segédtétel bizonyítása e dolgozatban nem található, mutatóban itt most külön belátjuk, hogy az A. tétel 2) kikötése a jelen esetben teljesül. $Y_n(t) = \sqrt{n}(F_n(t) - t) = \sqrt{n}Y_n^{**}(t)$ jelöléssel (tehát $m_n \equiv 0$, $d_n = n^{-1/2}$)

$$\begin{aligned}
 & P \left\{ \max_{n(1-c) \leq m \leq n(1+c)} |Y_n^{**}(s) - Y_m^{**}(s)| > \frac{\varepsilon}{\sqrt{n}} \right\} \leq \\
 & \leq P \left\{ \max_{n(1-c) \leq m \leq n(1+c)} |Y_n^{**}(s) - Y_m^{**}(s)| > \frac{\varepsilon}{2\sqrt{n}} \right\} + \\
 & + P \left\{ \max_{n(1-c) \leq m \leq n(1+c)} |Y_n^{**}(s) - Y_m^{**}(s)| > \frac{\varepsilon}{2\sqrt{n}} \right\}
 \end{aligned}$$

miatt elegendő megmutatni, hogy minden pozitív ε -hoz és η -hoz van olyan $c > 0$ és n_0 , hogy $n \geq n_0$ esetén

$$P \left\{ \max_{n(1-c) \leq m \leq n} |Y_n^{**}(s) - Y_m^{**}(s)| > \frac{\varepsilon}{2\sqrt{n}} \right\} < \frac{\eta}{2}$$

és

$$P \left\{ \max_{n \leq m \leq n(1+c)} |Y_n^{**}(s) - Y_m^{**}(s)| > \frac{\varepsilon}{2\sqrt{n}} \right\} < \frac{\eta}{2}.$$

Csak a második egyenlőtlenséggel foglalkozunk, az első teljesen hasonlóan igazolható. Nyilván

$$\begin{aligned}
 & P \left\{ \max_{n \leq m \leq n(1+c)} |Y_n^{**}(s) - Y_m^{**}(s)| > \frac{\varepsilon}{2\sqrt{n}} \right\} = \\
 & = P \left\{ \max_{n \leq m \leq n(1+c)} \left| \frac{1}{n} \sum_{i=1}^n \psi_i - \frac{1}{m} \sum_{i=1}^m \psi_i \right| > \frac{\varepsilon}{2\sqrt{n}} \right\},
 \end{aligned}$$

Ez pedig a következő módon becsülhető:

$$\begin{aligned}
 & P \left\{ \max_{n \leq m \leq n(1+c)} \left| \frac{1}{m} \sum_{i=1}^{m-n} (\psi_{n+1} - s) - \right. \right. \\
 & \left. \left. - \frac{m-n}{mn} \sum_{i=1}^n (\psi_i - s) \right| > \frac{\varepsilon}{2\sqrt{n}} \right\} \leq \\
 & \leq P \left\{ \max_{n \leq m \leq n(1+c)} \frac{1}{m} \left| \sum_{i=1}^{m-n} (\psi_{n+1} - s) \right| > \frac{\varepsilon}{4\sqrt{n}} \right\} + \\
 & + P \left\{ \max_{n \leq m \leq n(1+c)} \left(\frac{1}{n} - \frac{1}{m} \right) \left| \sum_{i=1}^n (\psi_i - s) \right| > \frac{\varepsilon}{4\sqrt{n}} \right\} \leq \\
 & \leq P \left\{ \max_{0 \leq k \leq [nc]} \left| \sum_{i=1}^k (\psi_{n+1} - s) \right| > \frac{\varepsilon \sqrt{n}}{4} \right\} + \\
 & + P \left\{ \left(\frac{1}{n} - \frac{1}{n + [nc]} \right) \left| \sum_{i=1}^n (\psi_i - s) \right| > \frac{\varepsilon}{4\sqrt{n}} \right\} .
 \end{aligned}$$

A Kolmogorov és a Csobisev egyenlőtlenségek alkalmazásával adódik, hogy az utóbbi összeg nem nagyobb mint

$$\frac{16s(1-s)[nc]}{\varepsilon^2 n} + \frac{16s(1-s)}{\varepsilon^2} \left(\frac{n[nc]}{n^2 + n[nc]} \right)^2 < \eta/2 ,$$

ha c -t elég kicsire és n -et elég nagyra választjuk, hiszen az utolsó sor baloldalán lévő összeg $n \rightarrow \infty$ esetén

$$\frac{16 s(1-s)}{\varepsilon^2} \left(c + \frac{c^2}{(1+c)^2} \right) \text{-hez tart.}$$

Beláttuk tehát, hogy rögzített $s \in [0,1]$ esetén

$$Y_{\nu_n}(s) \xrightarrow{\mathcal{D}} W_s^0.$$

Tekintsünk most két pontot: $s, t \in [0,1]$. Igazolnunk kell, hogy

$$(Y_{\nu_n}(s), Y_{\nu_n}(t)) \xrightarrow{\mathcal{D}} (W_s^0, W_t^0).$$

A 3. segédttétel miatt ehhez elegendő az, hogy bármely (c_1, c_2) vektorra

$$(15) \quad c_1 Y_{\nu_n}(s) + c_2 Y_{\nu_n}(t) \xrightarrow{\mathcal{D}} c_1 W_s^0 + c_2 W_t^0.$$

Y_n -nek W^0 -hoz való gyenge konvergenciájából, a 6. tétel alkalmazásával, mindenesetre következik, hogy

$$(Y_n(s), Y_n(t)) \xrightarrow{\mathcal{D}} (W_s^0, W_t^0),$$

amiből viszont, újra csak a 3. segédttétel alapján kapjuk, hogy minden (c_1, c_2) valós vektorra

$$(16) \quad c_1 Y_n(s) + c_2 Y_n(t) \xrightarrow{\mathcal{D}} c_1 W_s^0 + c_2 W_t^0 .$$

Az A. tétel feltételeit, tehát a keverést és az Anscombe-féle kikötést most a (16) bal oldalán lévő valószínűségi változókra pontosan ugyanugy láthatjuk be, mint az imént egy időpillanat esetén. Ugyhogy (15) tényleg igaz. Kettőnél több rögzített pont esetét teljesen ugyanugy tudjuk elintézni, tehát Y_{ν_n} végesdimenziós eloszlásai tényleg W^0 megfelelő végesdimenziós eloszlásaihoz tartanak.

II) Mivel $Y_{\nu_n}(0)$ 1 valószínűséggel 0, a 14. tétel szerint Y_{ν_n} feszességéhez elegendő belátni, hogy bármely $\varepsilon > 0$ és $\eta > 0$ számokhoz található olyan $(0,1)$ -beli δ és n_0 egész, hogy

$$P \left\{ \sup_{|s-t| < \delta} |Y_{\nu_n}(s) - Y_{\nu_n}(t)| > \varepsilon \right\} < \eta .$$

ha $n \geq n_0$, azaz, hogy (felidézve $w(Y_{\nu_n}, \delta) = \sup_{|s-t| < \delta} |Y_{\nu_n}(s) - Y_{\nu_n}(t)|$ jelölést)

$$\lim_{\delta \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} P \{ w(Y_{\nu_n}, \delta) > \varepsilon \} = 0 .$$

Legyenek $\varepsilon > 0$ és $\eta > 0$ adott számok és legyen $\ominus$ olyan pozitív szám, hogy $\frac{\ominus}{1-\ominus} < \frac{\eta}{4}$. Válasszunk olyan $0 < a < b$ számokat, amelyekre

$$(17) \quad P \{ a \leq \nu < b \} > 1 - \frac{\eta}{4} .$$

Tegyük fel, hogy $\varepsilon < a$. Ez a kikötés nem megy az általánosítás rovására. Az 5. segédtétel szerint θ -hoz és ε -hoz megadható olyan n_1 és c_1 ($i=1,2,\dots$), hogy $n \geq n_1$ esetén

$$P \left\{ \max_{n \leq m \leq n(1+c_1)} \sup_t |Y_m(t) - Y_n(t)| > \frac{\varepsilon}{4} \right\} < \theta^1.$$

Két lehetőség van:

$$\alpha) (a-\varepsilon) \prod_{i=1}^{\infty} (1+c_i) > b+\varepsilon; \quad \beta) (a-\varepsilon) \prod_{i=1}^{\infty} (1+c_i) \leq b+\varepsilon.$$

Először α -val foglalkozunk. Legyen k a legkisebb természetes szám, amelyre $(a-\varepsilon) \prod_{i=1}^k (1+c_i) > b+\varepsilon$. Osszuk be az $[a-\varepsilon, b+\varepsilon]$ intervallumot az

$$a_0 = a - \varepsilon; \quad a_i = a_0 \prod_{j=1}^i (1+c_j), \quad i=1, \dots, k-1; \quad a_k = b + \varepsilon.$$

osztópontokkal. x valós szám esetén jelölje $\{x\}$ a az x -nél nem kisebb egész számok közül a legkisebbet. Az

$$A_n^{(1)} = \left\{ a_{i-1} \leq \frac{\nu_n}{\omega(n)} < a_i \right\}$$

és

$$C(n, k, i, \varepsilon) = \left\{ \sup_t |Y_{\nu_n}(t) - Y_{\{ \omega(n) a_{i-1} \}}(t)| \leq \varepsilon \right\}$$

($i=1, \dots, k$) jelölésekkel mostmár, felhasználva a teljes valószínűség tételét, nyerjük

$$P \left\{ (Y_{\nu_n}, \delta) > \varepsilon \right\} \leq P \left\{ w(Y_{\nu_n}, \delta) > \varepsilon, \quad a \leq \nu < b, \right.$$

$$\left. \left| \frac{\nu_n}{\omega(n)} - \nu \right| < \varepsilon \right\} + P \left\{ \nu \notin [a, b] \right\} + P \left\{ \left| \frac{\nu_n}{\omega(n)} - \nu \right| \geq \varepsilon \right\}.$$

Ha n nagyobb, mint valamely megfelelő n^* , akkor a jobb oldal nem nagyobb mint



$$\begin{aligned}
 & P \left\{ w(Y_{\nu_n}, \delta) > \varepsilon, a - \varepsilon \leq \frac{\nu_n}{w(n)} < b + \varepsilon \right\} + \frac{\eta}{4} + \frac{\eta}{4} = \\
 & = \sum_{i=1}^k P \left\{ w(Y_{\nu_n}, \delta) > \varepsilon, A_n^{(i)} \right\} + \frac{\eta}{2} \leq \\
 & \leq \sum_{i=1}^k P \left\{ w(Y_{\nu_n}, \delta) > \varepsilon, A_n^{(i)}, C(n, k, i, \varepsilon/4) \right\} + \\
 & + \sum_{i=1}^k P \left\{ A_n^{(i)}, \overline{C(n, k, i, \varepsilon/4)} \right\} + \eta/2 \leq \\
 & \leq \sum_{i=1}^k P \left\{ \sup_{|s-t| < \delta} (|Y_{\nu_n}(s) - Y_{\{\omega(n)a_{i-1}\}}(s)| + |Y_{\{\omega(n)a_{i-1}\}}(s) - \right. \\
 & \left. - Y_{\{\omega(n)a_{i-1}\}}(t)| + |Y_{\{\omega(n)a_{i-1}\}}(t) - Y_{\nu_n}(t)|) > \varepsilon, A_n^{(i)}, C(n, k, i, \varepsilon/4) \right\} + \\
 & + \sum_{i=1}^k P \left\{ A_n^{(i)}, \overline{C(n, k, i, \varepsilon/4)} \right\} + \eta/2 \leq \\
 & \leq \sum_{i=1}^k P \left\{ \sup_{|s-t| < \delta} (2 \sup_{0 \leq v \leq 1} |Y_{\nu_n}(v) - Y_{\{\omega(n)a_{i-1}\}}(v)| + \right. \\
 & \left. + |Y_{\{\omega(n)a_{i-1}\}}(s) - Y_{\{\omega(n)a_{i-1}\}}(t)|) > \varepsilon, A_n^{(i)}, C(n, k, i, \varepsilon/4) \right\}
 \end{aligned}$$

$$\begin{aligned}
 & + \sum_{i=1}^k P \left\{ A_n^{(i)}, C(n, k, i, \varepsilon/4) \right\} + \eta/2 \leq \\
 & \leq \sum_{i=1}^k P \left\{ w(Y_{\{w(n)a_{i-1}\}}, \delta) > \frac{\varepsilon}{2} \right\} + \\
 & + \sum_{i=1}^k P \left\{ A_n^{(i)}, \overline{C(n, k, i, \varepsilon/4)} \right\} + \eta/2 .
 \end{aligned}$$

Az előző paragrafusban bebizonyítottuk, hogy Y_n a feszesség most fennforgó követelményének is eleget tesz, azaz, $\varepsilon/2$ -hez és θ^1 -hez is megadható olyan δ_1 és $n_{(1)}$, hogy $n \geq n_{(1)}$ esetén

$$P \left\{ w(Y_n, \delta_i) > \frac{\varepsilon}{2} \right\} < \theta^i (i = 1, \dots, k) .$$

Ha most $n^{**} (\geq n^*)$ olyan, hogy $n \geq n^{**}$ esetén $w(n)(a-\varepsilon) \geq \max(n_{(1)}, \dots, n_{(k)})$ és az eddig meg nem határozott δ -t $\min(\delta_1, \dots, \delta_k)$ -nak válasszuk, akkor

$$\sum_{i=1}^k P \left\{ w(Y_{\{w(n)a_{i-1}\}}, \delta) > \frac{\varepsilon}{2} \right\} < \sum_{i=1}^k \theta^i < \frac{\eta}{4} .$$

Ilyen δ -val, $n \geq n^{**}$ esetén tehát

$$P \left\{ w(Y_n, \delta) > \varepsilon \right\} < \sum_{i=1}^k P \left\{ A_n^{(i)}, \overline{C(n, k, i, \varepsilon/4)} \right\} + \frac{\eta}{4} .$$

Ugyanakkor

$$P \left\{ A_n^{(i)}, C(n, k, i, \varepsilon/4) \right\} \leq P \left\{ \max_{1 \leq j \leq k} \{ \omega(n) a_{i-1}^{(j)} \} < \sup_{1 \leq j \leq k} \{ \omega(n) a_{i-1}^{(j)} \} (1 + c_i) \right\}$$

$$\sup_t |Y_j(t) - Y_{\{\omega(n) a_{i-1}\}}(t)| > \frac{\varepsilon}{4} \} < \theta^i,$$

ha $\omega(n) a_{i-1} \geq n_i$. És ha mostmár n_0 ($\geq n^{**}$) olyan nagy, hogy $n \geq n_0$ esetén $\omega(n)(a - \varepsilon) \geq \max(n_1, \dots, n_k)$, akkor δ előbbi választásával, az n_0 -nál nagyobb n -ekre

$$P \left\{ \omega(Y_n, \delta) > \varepsilon \right\} < \sum_{i=1}^k \theta^i + \frac{3}{4} \eta < \eta.$$

Az α) esetben tehát készen vagyunk.

β) Amennyiben $(a - \varepsilon) \prod_{i=1}^{\infty} (1 + c_i) \leq b + \varepsilon$, akkor legyen K olyan (≥ 1) szám, hogy

$$(18) \sqrt[k]{a - \varepsilon} \prod_{i=1}^{\infty} (1 + c_i) > \sqrt[k]{b + \varepsilon}.$$

Ekkor $\mu = \sqrt[k]{a - \varepsilon}$, $\mu_n = [\sqrt[k]{\mu_n}]$, $\alpha(n) = [\sqrt[k]{\omega(n)}]$ jelölésekkel

$$\frac{\mu_n}{\alpha(n)} \xrightarrow{P} \mu$$

és

$$P \left\{ \sqrt[k]{a} \leq \mu < \sqrt[k]{b} \right\} > 1 - \frac{\eta}{4}.$$

Legyen ismét k az a legkisebb természetes szám, amelyre

$$\sqrt[k]{a - \varepsilon} \prod_{i=1}^k (1 + c_i) > \sqrt[k]{b + \varepsilon}.$$

Ílyen k (18) miatt létezik. Osszuk be a $[\sqrt[k]{a - \varepsilon}, \sqrt[k]{b + \varepsilon}]$

intervallumot az

$$a_0 = \sqrt[k]{a-\varepsilon} \quad ; \quad a_i = a_0 \prod_{j=1}^i (1+c_j), \quad i=1, \dots, k-1;$$

$a_k = \sqrt[k]{b+\varepsilon}$ osztópontokkal. Ezek után, μ , μ_n és $\alpha(n)$ -el ν , ν_n és $\omega(n)$ helyén, az egész α -beli bizonyítás változtatás nélkül megismételhető, és így arra jutunk, hogy $\{Y_{\mu_n}\}$ fészes. Minthogy az I)-beli bizonyítás nyilván Y_{μ_n} -re is érvényes, így Y_{μ_n} végesdimenziós eloszlásai W^0 megfelelő végesdimenziós eloszlásaihoz tartanak, ezért $Y_{\mu_n} \xrightarrow{\mathcal{D}} W^0$. E tényből valamint abból, hogy (mint az 5. segédtétel bizonyításánál láttuk)

$$\begin{aligned} Y_{[\sqrt[k]{\mu_n}]}(t) - Y_{\nu_n}(t) &= \left(1 - \left(\frac{\sqrt[k]{\mu_n}}{\nu_n} \right)^{1/2} \right) Y_{[\sqrt[k]{\mu_n}]}(t) - \\ &= \left(1 - \frac{\sqrt[k]{\mu_n}}{\nu_n} \right)^{1/2} Y'_{[\sqrt[k]{\mu_n}]}(t), \end{aligned}$$

a kívánt eredmény adódik, ha figyelembe vesszük a következő, egyszerű (az 51. oldalon látott) számolással adódó egyenlőtlenséget:

$$\begin{aligned} & \left| \lim_{\delta \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} P \left\{ w(Y_{\nu_n}, \delta) > 2\varepsilon \right\} - \lim_{\delta \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} P \left\{ w(Y_{\mu_n}, \delta) > \varepsilon \right\} \right| \leq \\ & \lim_{\delta \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} P \left\{ w(Y_{\mu_n}, \delta) > \varepsilon \right\} + \\ & + \lim_{n \rightarrow \infty} P \left\{ \sup_t |Y_{\nu_n}(t) - Y_{[\sqrt[k]{\mu_n}]}(t)| > \frac{\varepsilon}{2} \right\}. \end{aligned}$$

Ezzel a B. tételt bebizonyítottuk.

Tételünk, de ezzel tulajdonképpen általában a gyenge konvergencia erejét demonstrálandó felírjuk most a B. tétel már eddig is jelzett következményét: Mivel az eddig bevezetett $h_1, h_2, \dots, h_{11}$ funkcionálok mind folytonosak a D tér Szkorohod topológiájában, így $i=1, 2, \dots, 11$ esetén

$$h_i(Y_{\nu_n}) \xrightarrow{\mathcal{D}} h_i(W^0),$$

vagyis az empirikus folyamatra ismert minden határeloszlástétel érvényben van a véletlen mintaméretű empirikus folyamatra is. Ez az eredmény nyújt tehát lehetőséget arra a kijelentésre, hogy a megfigyelések számának minden határon túl való növelésekor a véletlen mintaméretű empirikus folyamat ugyanugy viselkedik, mint az empirikus folyamat.

6. §. Általánosítások

Mint ahogyan a 2. §-ban valószínűségi változókra tettük, a jelen dolgozatban tárgyalt kérdést meg lehet általánosan, legalábbis a C vagy D tér véletlen elemeit illetően fogalmazni a következőképpen. Ha tudjuk, hogy a C vagy D tér X_n és X véletlen elemeire $X_n \xrightarrow{\mathcal{D}} X$, akkor melyek azok a feltételek, amik biztosítják, hogy $X_n \xrightarrow{\mathcal{D}} X$, a $\mathcal{D}_n$ -re vonatkozó legáltalánosabb, B. tételbeli kikötés mellett. A [9] dolgozatban erre vonatkozóan már kaptunk egy eléggé használható eredményt, amit most bizonyítás nélkül vázolunk. Minden rögzített $\omega \in \Omega$ esetén legyen a $p(v) = p(v, \omega)$ függvény a $[0, \infty)$ intervallumon folytonos (illetve balról folytonos és jobboldali határértékekkel bíró). Legyen ezenkívül a_n pozitív, lassan oszcilláló számsorozat, amely ∞ -hez divergál (hogy a_n lassan oszcilláló, az azt jelenti, hogy $n^\alpha L(n)$ alakú, ahol α valós szám, $L(n)$ pedig olyan sorozat, hogy bármely $c > 0$ esetén

$L([cn])/L(n) \longrightarrow 1$.) Ha $p(v, \omega)$ mérhető az $(\Omega, \mathcal{A})$ -ra vonatkozóan, akkor az $X_n(t, \omega) = p(t, n, \omega) a_n^{-1}$ ($0 \leq t \leq 1$) egyenlőséggel értelmezett X_n leképezések a C (illetve D) véletlen elemei. Az ilyen alakú folyamatoknak (amelyek igen gyakran fordulnak elő az elméletben) az az igen hasznos tulajdonságuk, hogy ha X_n feszes, akkor X_n is (ugy a $(C, \mathcal{C})$ mint a $(D, \mathcal{D})$ terekre vonatkozóan). Ha tehát feltesszük, hogy

$X_n \xrightarrow{\mathcal{D}} X$, a C-n, vagy D-n, akkor mindkét tér esetén a Prohorov-féle 9. tételből következik, hogy $\{X_n\}$ fesszes és így $\{X_{\nu_n}\}$ is az. Így ha az X_{ν_n} végesdimenziós eloszlásainak megfelelő konvergenciáját az A. tétel alkalmazhatóságának megfelelően biztosítjuk, akkor a következő általános eredményre jutunk:

Ha a C (vagy D) tér X_n és X véletlen elemeire igaz az, hogy $X_n \xrightarrow{\mathcal{D}} X$, ahol is $X_n(t, \omega)$ a mondott $p(nt, \omega) a_n^{-1}$ alaku, ezenkívül tetszőleges $s \in T_X$ esetén $X_n(s)$ kielégíti az A. tétel 2) alatti Anscombe-féle feltételét ($d_n = a_n$, $m_n \equiv 0$), valamint minden $(C_1, \dots, C_k)$ vektorra és $t_1, \dots, t_k \in T_X$ pontokra $C_1 X_n(t_1) + \dots + C_k X_n(t_k)$ keverő $C_1 X(t_1) + \dots + C_k X(t_k)$ eloszlásfüggvényével, akkor $X_{\nu_n} \xrightarrow{\mathcal{D}} X$.

Itt jegyezzük meg, hogy [9]-ben a véletlen mintaméretű empirikus folyamat gyenge konvergenciáját éppen azzal sikerült bizonyítani, hogy $Y_{\nu_n}(t)$ fessességének ellenőrzése átjátszható egy $p(\nu_n t) a_{\nu_n}^{-1}$ alaku folyamat fessességének ellenőrzésére.

Az e dolgozatban látott módszer azonban szintén mutat az empirikus folyamaton. Vegyük ugyanis észre, hogy a B. tételbeli Y_{ν_n} fessességének bizonyításánál, azon túlmenően, hogy $Y_{\nu_n}(0)$ 1 valószínűséggel 0 és, hogy teljesíti az 5. segédtételbeli "általánosított" Anscombe-féle feltételt, csak azt használtuk ki, hogy bármely ε és η pozitív számokhoz van olyan δ és n_0 , hogy mihelyst $n \geq n_0$.

$$(19) \quad P \{ w(Y_n, \delta) > \varepsilon \} < \eta .$$

Ezen utóbbi feltétel viszont (a bal végpontban 0 folyamatoknál) a C-térben nemcsak elegendő, hanem szükséges is a fessességhez, ahogy a 11. tétel mutatja. Így jutunk tehát arra, hogy igaz a következő általános

C₁. TÉTEL: Ha a (C, C) tér X_n és X véletlen elemei olyanok, hogy

$$(i) \quad X_n \xrightarrow{\mathcal{D}} X,$$

(ii) minden (c₁, ..., c_k) vektorra és t₁, ..., t_k ∈ [0, 1] számokra C₁X_n(t₁) + ... + C_kX_n(t_k) keverő C₁X(t₁) + ... + C_kX(t_k) eloszlásfüggvényével,

(iii) bármely pozitív ε-ra

$$\lim_{c \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} P \left\{ \max_{n \leq m \leq n(1+c)} \sup_{0 \leq t \leq 1} |X_n(t) - X_m(t)| > \varepsilon \right\} = 0$$

(iv) {X_n} a B. tétel feltételének eleget tesz,

akkor

$$X_{\nu_n} \xrightarrow{\mathcal{D}} X .$$

BIZONYÍTÁS: Mivel (iii)-ből következik, hogy

$$(20) \quad \lim_{c \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} P \left\{ \max_{n \leq m \leq n(1+c)} |X_n(t) - X_m(t)| > \varepsilon \right\} = 0$$

minden $0 \leq t \leq 1$ esetén, így (ii) és (iii) a Mogyoródi-féle tételhez fűzött 2. megjegyzésen keresztül biztosítják, hogy X_n végesdimenziós eloszlásai X -éihez tartsanak. A 10. tétel szerint állításunk bizonyítva lesz, ha $\{X_n\}$ fesszeségét is kimutatjuk. Mivel a C tér teljes és szeparábilis, Prohorov 9. tételét használva (i)-ből következik, hogy $\{X_n\}$ fesszes. Ekkor viszont a 11. tételből következik (19) és (iii)-t felhasználva a 3. tétel bizonyításának II) része szó szerint megismételhető. Azt nem tettük fel, hogy $P\{X_n(0) = 0\} = 1$, így be kell még bizonyítanunk, hogy ha minden $\eta > 0$ számhoz található olyan L szám, amelyre

$$P\{|X_n(0)| > L\} < \eta$$

minden n -re fennáll, akkor olyan M is található, amelyre

$$(21) \quad P\{|X_n(0)| > M\} < \eta.$$

igaz minden n -re. Tudván azonban (20)-at, erre a kérdésre a B. tétel bizonyításának II) része megint csak megismételhető, és ennek révén oda jutunk, hogy (21) fennáll valamely n_0 -tól nagyobb vagy egyenlő összes n -re. Ilymódon egyedüli problémaként az marad, hogy ezt az n_0 -át 1-re csökkentsük. Ehhez pedig felidézzük ULAM 7. tételét, ami szerint $X_1, \dots, X_{n_0-1}$ önmagukban mind fesszesek, vagyis $i=1, 2, \dots, n_0-1$ esetén mindig van olyan M_i , hogy $P\{|X_i(0)| > M_i\} < \eta$, és most ha a (21)-beli M -et még megnöveljük úgy, hogy nagyobb

legyen mint $\max (M_1, \dots, M_{n_0-1})$, akkor (21) minden n -re fenn fog állni. Eszerint $\{X_{\nu_n}\}$ fesszes is, így $X_{\nu_n} \xrightarrow{\mathcal{D}} X$, és ezt akartuk bizonyítani.

A C és D terek közötti szoros analógiát látva már várható, hogy a C_1 tétel a D téren is igaz. Ha tehát X_n és X a D tér véletlen elemei, akkor a végesdimenziós eloszlások ugyanugy viselkednek mint előbb. Ekkor $X_n \xrightarrow{\mathcal{D}} X$ alapján (a 9. és 12. tételekkel) most az következik, hogy egyrészt, minden $\eta > 0$ számhoz van olyan L , hogy

$$(22) \quad P \left\{ \sup_{0 \leq t \leq 1} |X_n(t)| > L \right\} < \eta, \quad n \geq 1;$$

másrészt, hogy minden $\varepsilon > 0$ számra

$$(23) \quad \lim_{\delta \rightarrow 0} \lim_{n \rightarrow \infty} P \{ w(X_n, \delta) > \varepsilon \} = 0;$$

Nekünk azt kellene belátni, hogy mindkét relációban az n index helyébe ν_n -et írva is mindkettő fennáll. (22) esetében ez teljesen ugyanugy mutatható meg, mint az előbb (21)-nél; (23)-al kapcsolatban pedig vegyük észre, hogy a B. tétel bizonyításának II) részében még az sem játszott szerepet, hogy ott $w(Y_{\nu_n}, \delta)$ szerepelt, vagyis ha w helyett w' -ből indulunk ki, akkor a bizonyítás ugyanugy elvégezhető. A bizonyítás során ugyanis w -t mindössze egyszer használtuk fel, ez a lépés most, az ottaniakkal összhangban, így alakul: ha $\{t_j\}$ tetszőleges olyan véges rendszer,

amelyre $0=t_0 < t_1 < \dots < t_r=1$ és $t_j - t_{j-1} > \delta$ ($j=1, \dots, r$),
akkor

$$\begin{aligned} w(X_n, \delta) &\leq \inf_{\{t_j\}} \max_{0 \leq j \leq r} \sup_{s, t \in [t_{j-1}, t_j]} (|X_n(s) - X_{\{\omega(n)a_{i-1}\}}(s)| + \\ &+ |X_{\{\omega(n)a_{i-1}\}}(s) - X_{\{\omega(n)a_{i-1}\}}(t)| + |X_{\{\omega(n)a_{i-1}\}}(t) - X_n(t)|) \leq \\ &\leq 2 \sup_{0 \leq j \leq 1} |X_n(\omega) - X_{\{\omega(n)a_{i-1}\}}(\omega)| + w(X_{\{\omega(n)a_{i-1}\}}, \delta), \end{aligned}$$

és innét a bizonyítás úgy folytatható, mint előbb. Kaptuk tehát (ami a C tér esetében a C_1 tétel is volt) a valószínűségi változókra vonatkozó Mogyoródi-féle tétel, vagy az A. tétel direkt általánosítását a D tér véletlen elemeire:

C_2 TÉTEL. Ha a $(D, \mathcal{D})$ tér X_n és X véletlen elemei olyanok, hogy

$$X_n \xrightarrow{\mathcal{D}} X,$$

valamint ν_n pozitív egész értékű valószínűségi változók-
nak olyan sorozata, hogy egy ∞ -hez divergáló $\omega(n)$
számsorozatra

$$\frac{\nu_n}{\omega(n)} \xrightarrow{P} \nu(n \rightarrow \infty),$$

ahol ν pozitív valószínűségi változó, és teljesül a követ-
kező két feltétel:

(i) minden $(c_1, \dots, c_k)$ valós vektorra és tetsző-
leges $t_1, \dots, t_k \in T_X$ pontokra $c_1 X_n(t_1) + \dots + c_k X_n(t_k)$ keve-
rő $c_1 X(t_1) + \dots + c_k X(t_k)$ eloszlásfüggvényével,

(ii) bármely pozitív ε -ra

$$\lim_{c \rightarrow 0} \overline{\lim}_{n \rightarrow \infty} P \left\{ \max_{n \leq m \leq n(1+c)} \sup_{0 \leq t \leq 1} |X_m(t) - X_n(t)| > \varepsilon \right\} = 0,$$

akkor

$$X_{\nu_n} \xrightarrow{\otimes} X.$$

Hivatkozások

- [1] ALLEN, J. L. and BEEKMAN, J. A., A statistical test involving a random number of random variables, Ann. Math. Statist., 37(1966), 1305-1311.
- [2] ANDERSON, T. W. and DARLING, D. A., Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes, Ann. Math. Statist., 23(1952), 193-212.
- [3] ANSCOMBE, F. J., Large-sample theory of sequential estimation, Proc. Cambridge Philos. Soc., 48(1952), 600-607.
- [4] BARNDORFF-NIELSEN, O., On the limit distribution of the maximum of a random number of independent random variables, Acta Math. Acad. Sci. Hung., 15(1964), 399-403.
- [5] BILLINGSLEY, P., Convergence of probability measures (Wiley, New York, 1968).
- [6] BLUM, J., HANSON, D. and ROSENBLATT, J., On the central limit theorem for the sum of a random number of independent random variables, Zeitschrift für Wahrsch. und Verw. Gebiete, (1963), 389-393.
- [7] BREIMAN, L., Probability (Addison-Wesley, Reading, Mass; 1968).

- [8] CSÖRGÖ, M. and CSÖRGÖ, S., An invariance principle for the empirical process with random sample size, Bull. Amer. Math. Soc., 76(1970), 706-710.
- [9] CSÖRGÖ, M. and CSÖRGÖ, S., On weak convergence of randomly selected partial sums and the empirical process with random sample size. (Előkészületben)
- [10] CSÖRGÖ, S., A note on a limit theorem of J. Mogyoródi for sequences of random variables with random indices, Ann. Math. Statist. (Közlésre benyújtva)
- [11] CSÖRGÖ, S., On limit distributions of sequences of random variables with random indices, Acta Math. Acad. Sci. Hung. (Megjelenés alatt.)
- [12] DARLING, D. A., The Kolmogorov-Smirnov, Cramér- von Mises tests, Ann. Math. Stat., 28(1957), 823-838.
- [13] DAY, M. M., Normed linear spaces, Ergebnisse der Math. Heft 21, 1958.
- [14] DONSKER, M., Justification and extension of Doob's heuristic approach to the Kolmogorov-Smirnov theorems, Ann. Math. Statist., 23(1952), 277-281.
- [15] DOOB, J. L., Heuristic approach to the Kolmogorov-Smirnov theorems, Ann. Math. Statist., 20(1949), 393-403.

- [16] GIKHMAN, I. I. and SKOROKHOD, A. V., Introduction to the theory of random processes (W. B. Saunders Comp., Philadelphia, London, Toronto, 1969).
- [17] KAC, M., On deviations between theoretical and empirical distributions, Proc. Nat. Acad. Sci. U.S.A., 35(1949), 252-257.
- [18] MOGYORÓDI, J., A central limit theorem for the sum of a random number of random variables, Publ. Math. Inst. Hung. Acad. Sci., 7A(1962), 409-424.
- [19] MOGYORÓDI, J., Limit distributions for sequences of random variables with random indices, Trans. Fourth Prague Conf., Prague, 1965, 463-470.
- [20] PYKE, R., The weak convergence of the empirical process with random sample size, Proc. Cambridge Philos. Soc., 64(1968), 155-160.
- [21] RÉNYI, A., On the theory of order statistics, Acta Math. Acad. Sci. Hung., 4(1953), 191-227.
- [22] RÉNYI, A., On mixing sequences of sets, Acta. Math. Acad. Sci. Hung., 9(1958), 215-228.
- [23] RÉNYI, A., On the central limit theorem for the sum of a random number of random variables, Acta Math. Acad. Sci. Hung., 11(1960), 383-393.
- [24] SREEHARI, M., A limit theorem for the maximum of commutative sums, Acta Math. Acad. Sci. Hung., 19(1968), 117-120.

