A Valószínűségszámítás II. előadássorozat tízenegyedik előadása.
2002. november 26.

A nagy számok törvényéről. Első rész.

Ezen előadás témája a nagy számok (erős és gyenge) törvénye. Ez a matematikai eredmény olyan jelenségeket kíván megmagyarázni, mint például az a tapasztalati tény, hogy minden évben a születendő gyerekek körülbelül fele fiú és fele lány. Ezt és hasonló jelenségeket azzal magyaráznak, hogy a sok független hatás (az egyes születendő gyerek neme) kiegyenlíti egymást, és azok átlaga lényegében nem függ az egyes véletlen hatásoktól, hanem egy (determinisztikus) számhoz konvergál. Célunk ennek az állításnak pontos megfogalmazása és bebizonyítása. Azt fogjuk megmutatni, hogy független és egyforma eloszlású valószínűségi változók átlaga nagyon általános feltételek esetén egy (determinisztikus) konstanshoz konvergál. A ,,nagyon általános feltételek esetén" kitétel jelen esetben azt jelenti, hogy a valószínűségi változók eloszlásfüggvénye nagyon általános családból származhat. Általában az a szám, amelyikhez a független valószínűségi változók átlagai konvergálnak az átlagban szereplő valószínűségi változók várható értékei. De mivel olyan esetet is tárgyalni kívánunk, amikor ez a várható érték nem létezik, ezért nem mondjuk azt, hogy a tekintett valószínűségi változók átlagának a limesze e valószínűségi változók várható értéke. (Mint a vizsgálatokból ki fog derülni a tipikus esetekben, amikor a várható érték létezik, akkor a határérték várakozásainknak megfelelően ez a szám.)

Ahhoz, hogy vizsgálatainkat elkezdjük, először tisztázni kell, mit értünk azon, hogy a valószínűségi változók átlaga konvergál. Ez a fogalom magyarázatra szorul, mivel a valószínűségi változók egy valószínűségi mezőn értelmezett függvények, és függvények konvergenciájának több különböző lehetséges definiciója van.

Jelen kontextusban két olyan kitüntetett konvergencia fogalom van, amelyiket érdemes tárgyalni. Ezek egyike az egy valószínűséggel való konvergencia, a másik pedig a sztochasztikus konvergencia. Mind a két konvergenciafogalom szerinti konvergenciát vizsgálni fogjuk. A két konvergenciafogalom nem esik egybe, és ez az oka annak, hogy beszélünk a nagy számok gyenge és erős törvényéről is. Mint látni fogjuk, a két különböző törvény között az a különbség, hogy az átlag egy valószínűségi vagy sztochasztikus konvergenciáját vizsgáljuk.

Ahhoz, hogy ezt a programot végrehajthassuk, először meg kell adnunk a szükséges konvergencia definiciókat és meg kell beszélnünk kapcsolatukat egymással. Felidézem az eloszlásban való konvergencia fogalmát is, annak érdekében, hogy ezt is összehasonlíthassuk a fenti két konvergenciafogalommal.

Az egy valószínűségű konvergencia definiciója: *Valószínűségi változók ξ_n, $n = 1, 2, \dots$, sorozata akkor konvergál egy valószínűséggel egy ξ valószínűségi változóhoz, ha (egyrészt ezek a valószínűségi változók ugyanazon az $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn vannak definiálva, másrészt)*

$$P\left(\omega\colon \lim_{n\to\infty} \xi_n(\omega) = \xi(\omega)\right) = 1.$$

Megjegyzés: Az egy valószínűségi konvergencia fogalmát a mértékelméletben is használják, de ott azt majdnem mindenütt való konvergenciának is hívják. (Az angol nyelvű irodalomban az almost sure convegence, almost everywhere convergence vagy convergence with probability one kifejezések használatosak.)

A sztochasztikus konvergencia definiciója: *Valószínűségi változók* ξ_n*,* $n = 1, 2, \dots,$ *sorozata akkor konvergál sztochasztikusan egy* ξ *valószínűségi változóhoz, ha (egyrészt ezek a valószínűségi változók ugyanazon az* $(\Omega, \mathcal{A}, P)$ *valószínűségi mezőn vannak definiálva, másrészt) minden* $\varepsilon > 0$ *számra*

$$\lim_{n\to\infty} P\left(|\xi_n(\omega) - \xi(\omega)| > \varepsilon\right) = 0.$$

Megjegyzés: A mértékelméletben előforduló kifejezések közül a mértékben való konvergencia felel meg ennek a fogalomnak. Az egyetlen apró különbség a mértékelmélet és valószínűségszámítás szóhasználata között abban vam, hogy a mértékelméletben nem csak valószínűségi (azaz egyre normált mértékeket tekintenek).

Az eloszlásban való konvergencia definiciója: *Valószínűségi változók* ξ_n*,* $n = 1, 2, \dots,$ *sorozata akkor konvergál eloszlásban egy* $F(u)$ *eloszlásfüggvényhez vagy az ezen eloszlásfüggvény által meghatározott eloszláshoz, ha* $\lim_{n\to\infty} P(\xi_n < u) = F(u)$ minden olyan u számra, ahol az $F(\cdot)$ eloszlásfüggvény függvény folytonos. *(Azt mondjuk, hogy a* ξ_n*,* $n = 1, 2, \dots,$ *valószínűségi változók sorozata eloszlásban konvergál egy* ξ *valószínűségi változóhoz, ha ez a sorozat eloszlásban konvergál az* $F(u) = P(\xi < u)$ *eloszlásfüggvényhez.)*

A következő kapcsolat érvényes a fenti konvergenciafogalmak között.

Egy valószínűségi konvergencia $\Rightarrow$ Sztochasztikus konvergencia $\Rightarrow$ Eloszlásban való konvergencia.

Először megtárgyalom az egy valószínűségi és sztochasztikus konvergencia kapcsolatát.

Ha $\xi_n(\omega) \to \xi(\omega)$ egy valószínűséggel, akkor definiálva az

$$A_n = A_n(\varepsilon) = \left\{\omega\colon \sup_{k\ge n} |\xi_k(\omega) - \xi(\omega)| < \varepsilon\right\}$$

halmazokat kapjuk, hogy az egymásba skatulyázott A_n halmazokra, (azaz $A_1(\omega) \subset A_2 \subset \cdots$), $P\left(\bigcup\limits_{n=1}^{\infty} A_n\right) = 1$. Ezért $\lim_{n\to\infty} P(A_n) = 1$. Mivel $\{\omega\colon |\xi_n(\omega) - \xi(\omega)| < \varepsilon\} \supset A_n$, $P(|\xi_n(\omega) - \xi(\omega)| < \varepsilon) \to 1$, azaz $P(|\xi_n(\omega) - \xi(\omega)| \ge \varepsilon) \to 0$, ha $n \to \infty$. Ez azt jelenti, hogy az egy valószínűségű konvergenciából következik a sztochasztikus konvergencia.

Megfogalmazom az alábbi állítást, melyet nem nehéz bebizonyítani. De mivel nem lesz rá később szükségünk, azért elhagyom a bizonyítást.

Állítás: *Valószínűségi változók ξ_n, $n = 1, 2, \dots$, sorozata, akkor és csak akkor konvergál egy valószínűséggel egy ξ valószínűségi változóhoz, ha az $\eta_n = \sup_{k \geq n} |\xi_k - \xi|$ valószínűségi változók sorozata sztochasztikusan konvergál nullához, azaz minden $\varepsilon > 0$ számra*

$$\lim_{n\to\infty} P\left(\sup_{k\geq n} |\xi_k - \xi| > \varepsilon\right) = 0.$$

Lássunk példát arra, hogy lehetséges olyan ξ_n, $n = 1, 2, \dots$, és ξ valószínűségi változókat konstruálni, melyekre a ξ_n, $n = 1, 2, \dots$, sorozat sztochasztikusan tart ξ-hez, de a ξ_n sorozat nem konvergál egy valószínűséggel a ξ valószínűségi változóhoz.

Tekintsük a következő $(\Omega, \mathcal{A}, P)$ valószínűségi mezőt: Ω a $[0, 1]$ intervallum, $\mathcal{A}$ a Borel mérhető halmazok σ-algebrája a $[0, 1]$ intervallumon, a P valószínűségi mérték a Lebesgue mérték. Legyen

$$\xi_n(x) = \begin{cases} 1 & \text{ha } x \in \left[(n - 2^k)2^{-k}, (n + 1 - 2^k)2^{-k}\right] \\ 0 & \text{ha } x \notin \left[(n - 2^k)2^{-k}, (n + 1 - 2^k)2^{-k}\right] \end{cases}$$
$$\text{akkor ha } 2^k \leq n < 2^{k+1}, \quad k = 1, 2, \dots,$$

és $\xi(x) = 0$ minden $0 \leq x \leq 1$ számra. Ekkor $P(|\xi_n - \xi| > \varepsilon) = 2^{-k}$ minden $1 > \varepsilon > 0$ számra, ha $2^k \leq n < 2^{k+1}$. Tehát a ξ_n, $n = 1, 2, \dots$, sorozat sztochasztikusan konvergál a ξ valószínűségi változóhoz. Viszont mivel $\limsup_{n\to\infty} \xi_n(x) = 1$ minden $0 \leq x \leq 1$ számra, ezért a ξ_n sorozat nem konvergál egy valószínűséggel a ξ valószínűségi változóhoz.

Másrészt a Borel-Cantelli lemmából következik, hogy amennyiben minden $\varepsilon > 0$ számra $\sum_{n=1}^{\infty} P(|\xi_n - \xi| > \varepsilon) < \infty$, akkor a ξ_n sorozat egy valószínűséggel konvergál a ξ valószínűségi változóhoz. Miért?

A sztochasztikus konvergencia és eloszlában való konvergencia közötti kapcsolatra érvényesek a következő állítások.

1. Álllítás: *Ha ξ_n, $n = 1, 2, \dots$, valószínűségi változók sztochasztikusan konvergálnak egy ξ valószínűségi változóhoz, akkor a ξ_n valószínűségi változók eloszlásban is konvergálnak ehhez a ξ valószínűségi változóhoz.*

Indoklás: Legyen x folytonossági pontja a ξ valószínűségi változó $F(\cdot)$ eloszlásfüggvényének, és rögzítve egy $\varepsilon > 0$ számot válasszunk olyan $\delta = \delta(\varepsilon) > 0$ számot, melyre $F(x) - \frac{\varepsilon}{2} \leq F(x - \delta) \leq F(x + \delta) < F(x) + \frac{\varepsilon}{2}$. Ezután válasszunk olyan $n_0 = n_0(\varepsilon, \delta)$ számot, melyre $P(|\xi_n - \xi| \geq \delta) < \frac{\varepsilon}{2}$. Ekkor $P(\xi_n < x) < P(\xi < x + \delta) + P(|\xi_n - \xi| \geq \delta) \leq F(x + \delta) + \frac{\varepsilon}{2} \leq F(x) + \varepsilon$. Másrészt $P(\xi_n > x) < P(\xi > x - \delta) + P(|\xi_n - \xi| \geq \delta) \leq 1 - F(x - \delta) + \frac{\varepsilon}{2} \leq 1 - F(x) + \varepsilon$, ha $n \geq n_0$. Innen $F(x) - \varepsilon \leq P(\xi_n < x) \leq F(x) + \varepsilon$ $n \geq n_0$ esetén. Mivel minden $\varepsilon > 0$ esetén érvényes egy ilyen becslés, innen következik a megfogalmazott állítás.

Természetesen lehetséges, hogy ξ_n valószínűségi változók sorozata eloszlásban konvergáljon egy ξ valószínűségi változóhoz, de sztochasztikusan ne konvergáljon. Erre példa az az eset, ha a ξ_n valószínűségi változók függetlenek és azonos eloszlásúak. Ekkor az eloazlásban való konvergencia nyílván teljesül, de ha a ξ_n valószínűségi változók eloszlása nem elfajult, azaz a ξ_n valószínűségi változók nem egyenlőek egy konstanssal egy valószínűséggel, akkor e valószínűségi változók nem konvergálnak sztochasztikusan. Viszont abban az esetben, ha a limesz konstans akkor igaz a következő állítás:

2. Állítás: *Ha ξ_n, $n = 1, 2, \dots$, valószínűségi változók eloszlásban konvergálnak egy a konstanshoz, (azaz egy olyan valószínűségi változóhoz, mely egy valószínűséggel az a konstanst veszi fel, akkor a ξ_n, $n = 1, 2, \dots$, sorozat sztochasztikusan is konvergál ehhez az a konstanshoz.*

Indoklás: A határ valószínűségi változó eloszlásfüggvénye az az $F(x)$ eloszlás, melyre $F(x) = 0$, ha $x \le a$, és $F(x) = 1$ ha $x > a$. Ennek a függvénynek az $x = a$ pontot kivéve minden pontja folytonossági pontja. Ezért minden $\varepsilon > 0$ számra $\lim\limits_{n\to\infty} P(\xi_n < x+\varepsilon) = 1$, $\lim\limits_{n\to\infty} P(\xi_n < x-\varepsilon) = 0$. Innen következik, hogy $\lim\limits_{n\to\infty} P(|\xi_n-\xi| < \varepsilon) = \lim\limits_{n\to\infty} P(|\xi_n-a| < \varepsilon) = \lim\limits_{n\to\infty} (P(\xi_n < x+\varepsilon) - P(\xi_n < x-\varepsilon)) = 1$. Innen következik a 2. Állítás.

Ezután megfogalmazom a nagy számok gyenge és erős törvényét.

Nagy számok gyenge törvényének a definiciója. *Legyen ξ_n, $n = 1, 2, \dots$, független, egyforma eloszlású valószínűségi változók sorozata egy valószínűségi mezőn, $S_n = \sum\limits_{k=1}^{n} \xi_k$, $k = 1, 2, \dots$. Azt mondjuk, hogy ezek a ξ_n, $n = 1, 2, \dots$, valószínűségi változók teljesítik a nagy számok gyenge törvényét, ha létezik olyan E szám, melyre teljesül, hogy az $\dfrac{S_n}{n}$, $n = 1, 2, \dots$, valószínűségi változók sztochasztikusan konvergálnak az E számhoz, azaz ahhoz a valószínűségi változóhoz, mely egy valószínűséggel az E konstanssal egyenlő.*

Nagy számok erős törvényének a definiciója. *Legyen ξ_n, $n = 1, 2, \dots$, független, egyforma eloszlású valószínűségi változók sorozata egy valószínűségi mezőn, $S_n = \sum\limits_{k=1}^{n} \xi_k$, $k = 1, 2, \dots$. Azt mondjuk, hogy ezek a ξ_n, $n = 1, 2, \dots$, valószínűségi változók teljesítik a nagy számok erős törvényét, ha létezik olyan E szám, melyre teljesül, hogy az $\dfrac{S_n}{n}$, $n = 1, 2, \dots$, valószínűségi változók egy valószínűséggel konvergálnak az E számhoz, azaz ahhoz a valószínűségi változóhoz, mely egy valószínűséggel az E konstanssal egyenlő.*

Megjegyzés: Általában a nagy számok gyenge és erős törvénye az $E = E\xi_1$ számmal teljesül. Sőt, mint látni fogjuk a nagy számok erős törvénye mindig ezzel a konstanssal teljesül. De a nagy számok gyenge törvénye teljesülhet olyan esetben is, amikor a ξ_1 valószínűségi változónak nincs várható értéke. A fenti definiciót azért fogalmaztuk meg a fenti módon, hogy a lehető legáltalánosabb feltételek mellett fogalmazhassuk meg a nagy számok törvényét.

Először bebizonyítjuk a nagy számok gyenge és erős törvényét bizonyítását alkalmas feltételek mellett. Ezeknek az eredményeknek a bizonyítása egyszerű, viszont ezek megfogalmazásában még nem törekszünk a lehető legáltalánosabb eredmény bizonyítására.

A nagy számok gyenge törvénye egyszerűen bebizonyítható a Csebisev egyenlőtlenség segítségével, melyet felidézek.

Csebisev egyenlőtlenség: *Ha a ξ valószínűségi változó második momentuma $E\xi^2 = m_2$, akkor tetszőleges $x > 0$ számra*

$$P(|\xi| > x) \le \frac{m_2}{x^2}.$$

Innen következik, ha ezt az egyenlőtlenséget a $\bar{\xi} = \xi - E\xi$ valószínűségi változóra alkalmazzuk, hogy

$$P(|\xi - E\xi| > x) \le \frac{\operatorname{Var}\xi}{x^2},$$

Megjegyzés: A Csebisev egyenlőtlenség a Markov egyenlőtlenség következménye, mely szerint $P(|\xi| > x) \le \dfrac{E|\xi|}{x}$. Alkalmazva ezt az állítást a ξ^2 valószínűségi változóra kapjuk a Csebisev egyenlőtlenséget. Hasonló módon, alkalmazva a Markov egyenlőtlenséget a ξ^{2k} valószínűségi változóra, ahol k tetszőleges pozitív egész szám, $x > 0$, kapjuk, hogy $P(|\xi| > x) \le \dfrac{E\xi^{2k}}{x^{2k}}$

Tétel a nagy számok gyenge törvényéről. *Legyen ξ_n, $n = 1, 2, \dots$, független, egyforma eloszlású valószínűségi változók sorozata, melyekre teljesül az $E\xi_1^2 < \infty$ tulajdonság. Ezek a valószínűségi változók teljesítik a nagy számok gyenge törvényét az $E = E\xi_1$ konstanssal.*

Bizonyítás: Az adott feltételek mellett

$$P\left(\left|\frac{1}{n}\sum_{j=1}^{n}\xi_j - E\xi_1\right| > \varepsilon\right) \le \frac{\operatorname{Var}\left(\sum\limits_{j=1}^{n}\xi_j\right)}{n^2\varepsilon^2} = \frac{\operatorname{Var}\xi_1}{n\varepsilon^2}$$

minden $\varepsilon > 0$ számra, mert $\operatorname{Var}\left(\sum\limits_{j=1}^{n}\xi_j\right) = \sum\limits_{j=1}^{n}\operatorname{Var}\xi_j = n\operatorname{Var}\xi_1$. Mivel $\lim\limits_{n\to\infty}\dfrac{\operatorname{Var}\xi_1}{n\varepsilon^2} = 0$, innen következik a Tétel állítása.

Megjegyzés: A fenti állítás bizonyítását ki lehetett volna olvasni a centrális határeloszlástételből is. Valóban a tétel feltételeinek teljesülése esetén teljesülnek a centrális határeloszlástétel feltételei, ezért tetszőleges kis $\varepsilon > 0$ és nagy $K > 0$ számra felírhatjuk $n \ge n_0 = n_0(\varepsilon, K)$ esetében, hogy

$$P\left(\left|\frac{1}{n}\sum_{j=1}^{n}\xi_j - E\xi_1\right| > \varepsilon\right) \le P\left(\left|\frac{1}{\sqrt{n}}\sum_{j=1}^{n}\xi_j - E\xi_1\right| > K\right) \le 3(1-\Phi(K)),$$

ahol $\Phi(\cdot)$ a standard normális eloszlásfüggvényt jelöli. A jobboldal viszont tetszőlegesen kicsivé tehető a K konstans alkalmas megválasztásával. Érdemes észrevenni, hogy a nagy számok gyenge és mint látni fogjuk a nagy számok erős törvényének a bizonyítása is azon múlik, hogy olyan jó egyenlőtlenségeket kell bizonyítunk, melyek független valószínűségi változók összegeinek a viselkedését írják le.

Megmutatom, hogyan lehet a nagy számok gyenge törvényének fenti, a Csebisev egyenlőtlenségen alapuló bizonyítását finomítva a nagy számok erős törvényét is bebizonyítani alkalmas feltételek mellett. Az alábbi érvelésben valójában a tétel bizonyítása érdekében a szükségesnél sokkal erősebb kikötéseket teszek.

A nagy számok gyenge törvényét úgy bizonyítottuk be, hogy megvizsgáltuk milyen becslést ad a Csebisev egyenlőtlenség annak valószínűségére, hogy független, egyforma eloszlású valószínűségi változók átlaga legalább ε-nal eltér e változók várható értékétől. Idézzük fel ezt a számolást, illetve ezzel párhuzamosan vizsgáljuk meg azt is, milyen becslést kapunk, ha továbbra is azt a valószínűségi változót vizsgáljuk, melyet úgy kapunk, hogy független, egyforma eloszlású valószínűségi változók átlaga minusz azok várható értékét vesszük, de most ennek a valószínűségi változónak nemcsak a második momentumát számoljuk ki, mint tettük a Csebisev egyenlőtlenség alkalmazásában, hanem a negyediket is. Látni fogjuk, hogy ekkor érdekes új eredményeket kapunk.

Vezessük be a következő jelöléseket. Legyenek $\xi_1, \xi_2, \dots,$ független, egyforma eloszlású valószínűségi változók, $S_n = \xi_1 + \cdots + \xi_n$. Tegyük fel, hogy $E\xi_1 = 0$, és vizsgáljuk az

$$E\left(\frac{S_n}{n}\right)^2, \quad E\left(\frac{S_n}{n}\right)^4 \quad \text{és} \quad P\left(\frac{|S_n|}{n} > \varepsilon\right) = P\left(\left|\frac{S_n}{n} - E\xi_1\right| > \varepsilon\right), \; \varepsilon > 0$$

kifejezéseket. Az utolsó valószínűség vizsgálatában nem jelent megszorítást az $E\xi_1 = 0$ feltétel, mert a ξ_j valószínűségi változókat a $\xi_j - E\xi_j$ valószínűségi változókkal helyettesítve az általános esetet erre a speciális esetre redukálhatjuk.

$$\operatorname{Var}\frac{S_n}{n} = \frac{1}{n^2}\operatorname{Var}S_n = \frac{n}{n^2}\operatorname{Var}\xi_1 = \frac{1}{n}\operatorname{Var}\xi_1, \quad P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \le \frac{\operatorname{Var}\xi_1}{n\varepsilon^2}.$$

Innen következik, hogy $\lim\limits_{n\to\infty} P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) = 0$, tehát érvényes a nagy számok gyenge törvénye.

$$E\left(\frac{S_n}{n}\right)^4 = \frac{1}{n^4}E(\xi_1 + \cdots + \xi_n)^4,$$

$$E(\xi_1 + \cdots + \xi_n)^4 = \sum_{k=1}^{n} E\xi_k^4 + 6\sum_{1\le j<k\le n} E\xi_j^2 E\xi_k^2 + 4\sum_{\substack{1\le j,k\le n\\ j\ne k}} \underbrace{E\xi_j^3 E\xi_k}_{=0}$$
$$+\, 12\sum_{\substack{1\le k<l\le n,\, 1\le j\le n\\ j\ne k,\, j\ne l}} \underbrace{E\xi_j^2 E\xi_k E\xi_l}_{=0}$$

$$+ 24 \sum_{1\le j,k,l,m\le n} \underbrace{E\xi_j E\xi_k E\xi_l E\xi_m}_{=0}$$
$$= nE\xi_1^4 + 3n(n-1)(E\xi_1^2)^2,$$

ezért
$$P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \le \frac{\frac1n E\xi_1^4 + 3\left(1-\frac1n\right)(E\xi_1)^2)}{n^2\varepsilon^4} \le \frac{\text{const.}}{n^2\varepsilon^4}.$$

De a fenti becslés csak akkor érvényes, ha $E\xi_1^4 < \infty$. Ebből a becslésbő következik, hogy $\sum\limits_{n=1}^{\infty} P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) < \infty$ minden $\varepsilon > 0$ számra, és a Borel–Cantelli lemma (könnyebbik feléből) következik, hogy minden $\varepsilon > 0$ számra és majdnem minden $\omega \in \Omega$ pontban teljesül, hogy $\left|\frac{S_n(\omega)}{n}\right| \le \varepsilon$, ha $n \ge n_0(\omega)$. Alkalmazva ezt a relációt minden $\varepsilon = \frac1k$ számra $k = 1, 2, \dots$, kapjuk a nagy számok erős törvényét, melyet az alábbi tételben fogalmazunk meg.

A nagy számok erős törvényéről szóló tétel *gyenge formája***:** *Legyen $\xi_1, \xi_2, \dots$, független egyforma eloszláaú valószínűségi változók sorozata egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, melyekre teljesül az $E\xi_1^4 < \infty$ feltétel, és vezessük be az $S_n = \sum\limits_{j=1}^{n} \xi_j$ valószínűségi változókat. Ekkor az $\frac{S_n(\omega)}{n}$ valószínűségi változók majdnem minden $\omega \in \Omega$-ra konvergálnak ez $E\xi_1$, számhoz, azaz ezek a valószínűségi változók teljesítik a nagy számok gyenge törvényét.*

Most megfogalmazom a nagy számok erős és gyenge törvényét kimondó tételt eredeti, éles formájában. A bizonyításokat csak később ismertetem.

A nagy számok erős törvényét kimondó tétel éles alakja. *Legyen ξ_1, ξ_2, ... , független egyforma eloszlású valószínűségi változók sorozata, és definiáljuk az $S_n = \sum\limits_{k=1}^{n} \xi_k$, $n = 1, 2, \dots$, részletösszegeket. Az $\frac{S_n(\omega)}{n}$, $n = 1, 2, \dots$, sorozat akkor és csak akkor konvergál pozitív valószínűséggel, ha $E|\xi_1| < \infty$. Ha $E|\xi_1| < \infty$, akkor ez a sorozat teljesíti a nagy számok erős törvényét, azaz*
$$\lim_{n\to\infty} \frac{S_n(\omega)}{n} = E\xi_1 \quad \textit{majdnem minden } \omega \in \Omega\textit{-ra.}$$

A nagy számok gyenge törvényét kimondó tétel éles alakja. *Legyen ξ_k, $k = 1, 2, \dots$, független, egyforma eloszlású valószínűségi változók sorozata F eloszlásfüggvénnyel. Ezek a valószínűségi változók akkor és csak akkor teljesítik a nagy számok gyenge törvényét, azaz a $\frac1n \sum\limits_{k=1}^{n} \xi_k$ átlag akkor és csak akkor konvergál sztochasztikusan $n \to \infty$ esetében egy a számhoz, $-\infty < a < \infty$, ha*
$$\lim_{x\to\infty} x\left[F(-x) + (1 - F(x))\right] = 0, \quad \textit{és} \quad \lim_{u\to\infty} \int_{-u}^{u} xF(\,dx) = a.$$

Annak érdekében, hogy jobban megértsük a nagy számok erős és gyenge törvényét lássunk néhány példát. Először példát mutatok olyan független egyforma eloszlású valószínűségi változókra, melyek teljesítik a nagy számok gyenge törvényét, de nem teljesítik a nagy számok erős törvényét. Az indoklásban hivatkozhatnánk az előbb kimondott tételekre, de tanulságosabb lehet közvetlen bizonyítást adni. Utána példát mutatok olyan független egyforma eloszlású valószínűségi változókra, melyek a nagy számok gyenge törvényét sem teljesítik.

1. példa: *Legyen ξ_k, $k=1,2,\dots$, független, egyforma eloszlású valószínűségi változók sorozata, az $f(u)=\dfrac{C}{u^2\log|u|}$, ha $|x|>3$, $f(u)=0$, ha $|u|\le 3$, képlettel megadott sűrűségfüggvénnyel. (A C konstanst a $\displaystyle\int_{|u|>3}\frac{C\,du}{u^2\log|u|}=1$ határozza meg.) Definiáljuk az $S_n=\sum\limits_{k=1}^{m}\xi_k$, $n=1,2,\dots$, részletösszegeket. Ekkor $E|\xi_1|=\infty$, ezért ezek a valószínűségi változók nem teljesítik a nagy számok erős törvényét. Másrészt az $\dfrac{S_n}{n}$ átlagok sztochasztikusan tartanak nullához, azaz minden $\varepsilon>0$ számra $P\left(\left|\dfrac{S_n}{n}\right|>\varepsilon\right)\to 0$, ha $n\to\infty$. Ez azt jelenti, hogy ez a sorozat teljesíti a nagy számok gyenge törvényét.*

Indoklás: Ezekre a valószínűségi változókra $E|\xi_1|=\displaystyle\int uf(u)\,du=\int_3^\infty\frac{2C}{u\log u}\,du=\infty$, mert $\displaystyle\int_3^x\frac{1}{u\log u}\,du=[\log\log u]_3^x$, és $\lim\limits_{x\to\infty}\log\log x=\infty$. Látni fogjuk, hogy az $E|\xi_1|=\infty$ relációból és a Borel–Cantelli lemmából következik, hogy nem teljesül a nagy számok gyenge törvénye. Ennek indoklását később ismertetem. Annak érdekében, hogy bebizonyítsuk a nagy számok gyenge törvényét erre a sorozatra, definiáljuk a következő valószínűségi változókat: Legyen $\bar\xi_k=\bar\xi_k^{(n)}=\xi_kI\left(|\xi_k|\le n\right)$, $\bar{\bar\xi}_k=\bar{\bar\xi}_k^{(n)}=\xi_kI\left(|\xi_k|>n\right)$, és $\bar S_n=\sum\limits_{k=1}^n\bar\xi_k$, $\bar{\bar S}_n=\sum\limits_{k=1}^n\bar{\bar\xi}_k$. Ekkor $P(|S_n|>\varepsilon n)\le P\left(|\bar S_n|>\dfrac{\varepsilon}{2}n\right)+P\left(|\bar{\bar S}_n|>\dfrac{\varepsilon}{2}n\right)\le \dfrac{\operatorname{Var}\bar S_n}{\frac{\varepsilon^2}{4}n^2}+nP\left(\bar{\bar\xi}_1\ne 0\right)$. Adjunk jó becslést a $\operatorname{Var}\bar\xi_1$ kifejezésre. $\operatorname{Var}\bar\xi_1=\displaystyle\int_3^n\frac{2C\,du}{\log u}\le\frac{\text{const.}\,n}{\log n}$, mert $\displaystyle\int_{\sqrt n}^n\frac{2C\,du}{\log u}\le\frac{\text{const.}\,n}{\log n}$ és $\displaystyle\int_3^{\sqrt n}\frac{2C\,du}{\log u}\le\text{const.}\sqrt n$. Innen a fenti egyenlőtlenség jobboldalán szereplő első tagot úgy becsülhetjük felülről, hogy $\dfrac{\operatorname{Var}\bar S_n}{\frac{\varepsilon^2}{4}n^2}\le\dfrac{\operatorname{Var}\bar\xi_1}{\frac{\varepsilon^2}{4}n}\le\dfrac{\text{const.}}{\varepsilon^2\log n}$. A második tag becslése érdekében vegyük észre, hogy $P\left(\bar{\bar\xi}_1\ne 0\right)=\displaystyle\int_n^\infty\frac{2C}{u^2\log u}\le\frac{2C}{\log n}\int_n^\infty\frac{du}{u^2}\le\frac{\text{const.}}{n\log n}$. Innen $nP\left(\bar{\bar\xi}_1\ne 0\right)\le\dfrac{\text{const.}}{\log n}$. A fenti egyenlőtlenségekből következik, hogy $\lim\limits_{n\to\infty}P(|S_n|>\varepsilon n)=0$, és ezt akartuk megmutatni.

A következő példa megtárgyalása érdekében bevezetjük a következő úgynevezett Cauchy eloszlásokat.

A Cauchy eloszlás definiciója. *Egy eloszlást a paraméterű Cauchy eloszlásnak nevezünk, ha sűrőségfüggvénye* $f(u) = \dfrac{1}{\pi}\dfrac{a}{1+a^2u^2}$ *alakú,* $-\infty < u < \infty$.

Megjegyzés: A Cauchy eloszlással találkozhattak a gyakorlaton is. Ott szerepelt az egyik példában, hogy két független standard normális eloszlású valószínűségi változó hányadosa ($a = 1$ paraméterű) Cauchy eloszlású valószínűségi változó.

Nem nehéz belátni, hogy független Cauchy eloszlású valószínűségi sorozata változók (ugyanazzal az a paraméterrel) nem teljesíti a nagy számok gyenge törvényét, mert (például $a = 1$ választással a $G(x)$ Cauchy eloszlásra a

$$\lim_{x\to\infty} x(1-G(x)) = \lim_{x\to\infty} x\int_x^\infty \frac{1}{\pi}\frac{du}{1+2u^2} = \frac{1}{\pi}$$

reláció teljesül. Valóban $x\displaystyle\int_x^\infty \frac{du}{1+u^2} \ge x\int_x^\infty \frac{du}{u^2} = 1$ minden $x \ge 0$-ra, és

$$x\int_x^\infty \frac{du}{1+u^2} \le x\int_x^\infty \frac{(1+\varepsilon)\,du}{u^2} = 1+\varepsilon, \quad \text{ha } x \ge x(\varepsilon).$$

Független Cauchy eloszlású valószínűségi változók összegének viselkedéséről többet is mondhatunk. Azok teljesítik a következő rendkívül figyelemre méltó tulajdonságot.

Tétel. *Legyen* ξ *és* η *két független Cauchy eloszlású valószínűségi változó,* $a > 0$ *és* $b > 0$ *paraméterrel. Ekkor* $\xi + \eta$ *is Cauchy eloszlású* $a + b$ *paraméterrel. Speciálisan* n *független* a *paraméterű Cauchy eloszlású valószínűségi változó összege* na *paraméterű Cauchy eloszlású valószínűségi változó, átlaga pedig* a *paraméterű Cauchy eloszlású valószínűségi változó. Tehát független azonos paraméterű Cauchy eloszlású valószínűségi változók átlagának az eloszlása megegyezik az összeadandók eloszlásával.*

Magyarázat: Nem végzem el az állítás bizonyítását. Ki lehetne számolni független Cauchy eloszlású valószínűségi változók összegének a sűrűségfüggvényét konvolució segítségével, és így be lehet bizonyítani a Tételt. Ez meglehetősen kellemetlen, de (parciális törtekre bontás segítségével) kiszámolható integrálokhoz vezetne. Lényegesen egyszerűbben célhoz érünk, ha tudjuk a Cauchy eloszlás karakterisztikus függvényét. Az $a = 1$ paraméterű Cauchy eloszlás karakterisztikus függvénye $\varphi(t) = e^{-|t|}$, ahonnan az a paraméterű Cauchy eloszlás karakterisztikus függvénye $\varphi_a(t) = e^{-a|t|}$. Innen látszik, hogy $\varphi_a(t)\varphi_b(t) = \varphi_{a+b}(t)$, ahonnan a tételben megfogalmazott állítások azonnal láthatóak. Természetesen ez az indoklás csak úgy teljes, ha ki tudjuk számítani a Cauchy eloszlás karakterisztikus függvényét. Ezt a kiegészítésben megteszem. Azok számára, akik tanultak komplex függvénytant és megértették a reziduumszámítást, ez egyszerű feladat.

2. példa. *Legyen* $\xi_1, \xi_2, \dots,$ *független egyforma eloszlású valószínűségi változók sorozata* $f(u) = \dfrac{1}{\pi}\dfrac{1}{1+u^2}$, $-\infty < u < \infty$ *sűrűségfüggvénnyel. Ezek a valószínűségi változók nem teljesítik a nagy számok gyenge törvényét, azaz az* $\dfrac{1}{n}\sum\limits_{k=1}^{n} \xi_k$ *valószínűségi változók nem konvergálnak sztochasztikusan egy konstanshoz.*

3. példa. *Legyen* $\xi_1, \xi_2, \dots,$ *független egyforma eloszlású valószínűségi változók sorozata* $f(u) = \dfrac{C}{u^2 \log^2 u}$, *ha* $|u| \ge 3$, *és* $f(u) = 0$, *ha* $|u| < 3$ *sűrűségfüggvénnyel. (A* C *konstanst az* $\displaystyle\int_{|u|>3} \frac{C\,du}{u^2 \log^2 u} = 1$, *képlet határozza meg.) Ekkor* $E|\xi_1| < \infty$, *ezért a fent kimondott tétel alapján ezek a valószínűségi változók teljesítik a nagy számok erős törvényét, azaz az* $\dfrac{1}{n}\sum\limits_{k=1}^{n} \xi_k$ *valószínűségi változók egy valószínűséggel konvergálnak nullához.*

Indoklás: Elegendő ellenőrízni, hogy $E|xi_1 = \displaystyle\int_3^\infty \frac{2C\,du}{u \log^2 u} = \left[-\frac{2C}{\log u}\right]_3^\infty < \infty$. Ezért alkalmazható a nagy számok erős törvénye, és szimmetria meggondolások alapján $E\xi_1 = 0$.

A fenti pélák hasonlítottak egymáshoz abban, hogy mindegyikben a sűrűségfüggvény a plusz-minusz végtelen környezetében aszimptotikusan úgy viselkedett, mint $\frac{1}{u^2}$ szorozva egy logaritmus hatvány rendű korrekciós taggal. Ez a korrekciós tag befolyásolta, hogy bízonyos integrálok konvergensek vagy divergensek, és ettől függött, hogy teljesül-e a nagy számok erős vagy gyenge törvénye. Láttuk, hogy a nagy számok különböző törvényeit kimondó tételekben és példákban az játszott fontos szerepet, hogy az összeadandók eloszlásfüggvényei hogyan viselkednek a $\pm\infty$-ben, milyen gyorsan tartanak ottan az eloszlásfüggvények egyhez illetve nullához. Ettől függ ugyanis, hogy az ott szereplő integrálok konvergensek vagy divergensek. Hasonló képpel találkoztunk a centrális határeloszástétel vizsgálatában. Ott a Lindeberg feltétel jelent meg, mint a centrális határeloszlástétel elégséges (és bizonyos értelemben) szükséges feltétele, és ez is hasonló jellegű feltételt fejezett ki, nevezetesen azt, hogy az egyes összeadandók F_n eloszlásaiból képzett $\int_{|u|\ge D_n} u^2 F_n(\,du)$ integrálok viszonylag kicsik. Ez is olyan jellegű feltétel, amelyik azt fejezi ki, hogy az egyes összeadandók csak viszonlag kis valószínűséggel vesznek fel nagy értékeket. Kissé pongyolán, de talán a lényeget jól kifejezően azt mondhatjuk, hogy a valószínűségszámítás független valószínűségi változók összegeiről szóló tételei olyan feltételek mellett érvényesek, melyek azt biztosítják, hogy tipikus esetekben az egyes összeadandók rendkívül nagy értékei nem zavarják meg az összeg viselkedését.

A nagy számok törvényének vizsgálatában szükségünk van egy olyan eredményre, mely az összeadandók abszolut értékének véges várható értékét az itt szükséges vizsgálatokban jobban kezelhető módon fejezi ki. Ez a tartalma a következő lemmának.

Lemma. *Egy* ξ *valószínűségi változó abszolut értékének akkor és csak akkor van véges* *várható értéke, ha* $\sum\limits_{n=1}^{\infty} P(|\xi|>n)<\infty.$

Megjegyzem, hogy ennek az állításnak igaz a következő általánosítása, amelyiket hasonlóan lehet bizonyítani. De mivel erre nem lesz szükségünk, ennek bizonyítását elhagyom.

A lemma általánosítása. *Egy ξ valószínűségi változó akkor és csak akkor teljesíti az $E|\xi|^r<\infty$ momentum feltételt valamely ≥ 1 számra, ha* $\sum\limits_{n=1}^{\infty} n^{r-1}P(|\xi|>n)<\infty.$

A lemma bizonyítása. Vezessük be a következő $\tilde{\xi}$ valószínűségi változót: $\tilde{\xi}=j$, ha $j-1<|\xi|\leq j$, $j=1,2,\dots$. Ekkor $P(0\leq|\xi|-\tilde{\xi}\leq 1)=1$, ezért a $|\xi|$ és $\tilde{\xi}$ valószínűségi változók várható értéke egyszerre véges vagy végtelen. Viszont $E\tilde{\xi}=\sum\limits_{j=1}^{\infty} jP(\tilde{\xi}=j)=\sum\limits_{j=1}^{\infty} jP(j-1<|\xi|\leq j)$, ezért ennek az összegnek a konvergenciáját vagy divergenciáját kell vizsgálnunk.

Felírhatjuk, hogy $\sum\limits_{j=1}^{\infty} jP(j-1<|\xi|\leq j)=\sum\limits_{j=1}^{\infty} j(P(|\xi|>j-1)-P(|\xi|>j))$. Rendezzük át a fenti összeget a következő módon. Tetszőleges N számra

$$\begin{aligned}\sum_{j=1}^{N} jP(j-1<|\xi|\leq j)&=\sum_{j=1}^{N} j(P(|\xi|>j-1)-P(|\xi|>j))\\ &=\sum_{j=1}^{N-1} P(|\xi|>j)((j+1)-j)-NP(|\xi|>N).=\sum_{j=1}^{N-1} P(|\xi|>j)-NP(|\xi|>N).\end{aligned}$$

Megmutatjuk, hogy $N\to\infty$ határátmenettel a fenti relációból következik a Lemma állítása. Valóban, ha $E|\xi|<\infty$, akkor létezik olyan $K<\infty$ szám, melyre $NP(|\xi|>N)\leq\sum\limits_{j=N+1}^{\infty} jP(j-1<|\xi|\leq j)\leq\sum\limits_{j=1}^{\infty} jP(j-1<|\xi|\leq j)\leq K$, ahol a K konstans független az N számtól. Így ebben az esetben az előző becslés alapján léteznek olyan univerzáslis L és K számok, melyekre

$$L\geq\sum_{j=1}^{N-1} P(|\xi|>j)-K \quad \text{minden } L=1,2,\dots \text{ számra,}$$

ahonnan $\sum\limits_{N=1}^{\infty} P(|\xi|>j)<\infty$).

Ha $E|\xi|=\infty$, akkor $\sum\limits_{N=1}^{\infty} P(|\xi|>j)=\infty$, mert ekkor

$$\sum_{j=1}^{N-1} P(|\xi|>j)\geq\sum_{j=1}^{N} jP(j-1<|\xi|\leq j),$$

és $\lim\limits_{N\to\infty}\sum\limits_{j=1}^{N} jP(j-1<|\xi|\le j)=\infty$.

Megjegyzés: Az összegnek a fenti számolásban történt átrendezését Abel féle átrendezésnek nevezik, és ez sokszor hasznos. Az Abel féle átrendezés egyébként az integrálszámításban alkalmazott parciális integrálás diszkrét megfelelője.

Megfogalmazom a fenti lemma egy számunkra fontos következményét.

A Lemma következménye. *Legyenek $\xi_1,\xi_2,\dots,$ független, egyforma eloszlású valószínűségi változók. Ha $E|\xi_1|=\infty$, akkor e valószínűségi változók nem teljesítik a nagy számok erős törvényét. Pontosabban azt állíthatjuk, hogy ebben az esetben az $S_n(\omega)=\sum\limits_{j=1}^{n}\xi_j(\omega)$, $n=1,2,\dots,$ összegekre az $\dfrac{S_n(\omega)}{n}$ kifejezés majdnem minden $\omega\in\Omega$ pontban nem konvergens.*

Bizonyítás. Ebben az esetben a Lemma alapján a ξ_j valószínűségi változók azonos eloszlása miatt az $E|\xi_1|=\infty$ feltételből következik, hogy $\sum\limits_{n=1}^{\infty}P(|\xi_1|>n)=\sum\limits_{n=1}^{\infty}P(|\xi_n|>n)=\infty$. A ξ_n valószínűségi változók függetlensége miatt az $\{\omega\colon |\xi_n(\omega)|>n\}$ események is függetlenek. Ezért a Borel–Cantelli lemmából következik, hogy majdnem minden $\omega\in\Omega$-ra $|\xi_n(\omega)|>n$ végtelen sok az (ω elemi eseménytől függő) n indexre. Tekintsünk egy olyan $\omega\in\Omega$ elemi eseményt, melyre $\lim\limits_{n\to\infty}\dfrac{S_n(\omega)}{n}=a$ valamilyen véges a számra, melynek értéke függhet ettől az ω elemi eseménytől. Ilyen ω pontokban a $\lim\limits_{n\to\infty}\dfrac{S_{n-1}(\omega)}{n}=a$ reláció is teljesül. Ezért $\lim\limits_{n\to\infty}\left(\dfrac{S_n(\omega)}{n}-\dfrac{S_{n-1}(\omega)}{n}\right)=\lim\limits_{n\to\infty}\dfrac{\xi_n(\omega)}{n}=0$. Ez a reláció viszont, mint láttuk csak egy nulla valószínűségi halmazon teljesülhet.

A Cauchy eloszlás karakterisztikus függvényének a kiszámítása.

Legyen ξCauchy eloszlású valószínűségi változó $a=1$ paraméterrel . Ekkor

$$Ee^{it\xi}=\int_{-\infty}^{\infty}\frac{1}{\pi}\frac{e^{itu}}{1+u^2}\,du.$$

Ezt az integrált a komplex függvénytan reziduum tétele segítségével ki tudjuk számítani.

Vegyük észre, hogy a $g(z)=g(z,t)=\dfrac{e^{itz}}{\pi(1+z^2)}$ függvény analitikus a komplex számsíkon, két pólusa van a $z=\pm i$ pontokban. A $g(z)$ függvény reziduuma az i pontban e^{-t} a $-i$ pontban e^{t}. Tekintsük a következő körintegrált. A $g(z)=g(z,t)$ függvényt integráljuk a $[-R,R]$ szakaszon, majd a a $|z|=R$ $\mathrm{Im}\,z\ge 0$ félkörön, ha $t\ge 0$ és a $|z|=R$, $\mathrm{Im}\,z\le 0$ félkörön, ha $t\le 0$. Ekkor ennek a körintegrálnak az értéke a $g(z)$ függvény i pontbeli reziduumával egyenlő $t>0$ és a $-i$ pontbeli reziduumával a $t<0$ esetben. Másrészt az integrál megszorítása az R sugarú félkörre nullához tart, ha $R\to 0$. Innen következik, hogy $Ee^{it\xi}=\int_{-\infty}^{\infty}g(t,u)\,du=e^{-t}$, ha $t>0$, és $Ee^{it\xi}=\int_{-\infty}^{\infty}g(t,u)\,du=e^{t}$, ha $t<0$. Egységes jelöléssel azt írhatjuk, hogy $Ee^{it\xi}=e^{|t|}$.