A Valószínűségszámítás I. előadássorozat tizenharmadik előadása.
2007. május 15.

A nagy számok törvényét fogom tárgyalni. Felidézem a nagy számok már tárgyalt gyenge törvényét, és megfogalmazom a nagy számok erős törvényét is. Annak érdekében, hogy az eredményeket jobban megértsük felidézek néhány korábban tárgyalt fogalmat és eredményt. A bizonyítások jelentős részét elhagyom, vagy a nem kötelező *Kiegészítés* részben ismertetem.

A nagy számok erős törvénye.

Tárgyaltuk a nagy számok gyenge törvényét. Ez az eredmény azt mondja ki, hogy ha $\xi_1, \xi_2, \dots,$ független, egyforma eloszlású valószínűségi változók sorozata, és tekintjük e valószínűségi változók $S_n = \sum\limits_{j=1}^{n} \xi_j$, részletösszegeit és $\frac{S_n}{n}$ átlagait minden $n = 1, 2, \dots$ számra, akkor ezek az $\frac{S_n}{n}$ átlagok sztochasztikusan konvergálnak az ξ_1 valószínűségi változó $E\xi_1$ várható értékéhez. Beláttuk, (a Csebisev egyenlőtlenség segítségével), hogy a nagy számok gyenge törvénye érvényes akkor, ha a ξ_1 valószínűségi változónak létezik szórásnégyzete. Valójában ez a feltétel lényegesen gyengíthető. Megmutatom, hogy független valószínűségi változók átlagai nagyon általános feltételek mellett nemcsak sztochasztikusan, hanem egy valószínűséggel is konvergálnak a ξ_1 valószínűségi változó $E\xi_1$ várható értékéhez. Ezt az eredményt nevezik a nagy számok erős törvényének. Megfogalmazom az említett eredményeket és fogalmakat pontosan, és tárgyalni fogom ezen eredmények egymással való kapcsolatát. Sok eredményt csak kimondok, de időhiány miatt a bizonyítását elhagyom.

Először felírom a (részben már korábban ismertetett) különböző konvergenciafogalmakat.

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 konvergenciát a valószínűségszámításban és mértékelméletben majdnem mindenütt való konvergenciának is hívják.

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*

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

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* $\lim\limits_{n\to\infty} P(\xi_n < u) = F(u)$ 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 továbbiakban tárgyalni fogom (részben kitűzött feladatok formájában) a különböző konvergenciafogalmak közötti kapcsolatot.

A felsorolt konvergencia fogalmak között a következő a kapcsolat:

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

Tárgyaljuk először 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 azt kapjuk, hogy az egymásba skatulyázott A_n halmazokra, (azaz $A_1(\omega) \subset A_2 \subset \cdots$), teljesül a $P\left(\bigcup\limits_{n=1}^{\infty} A_n\right) = 1$ reláció. Ezért $\lim\limits_{n\to\infty} P(A_n) = 1$. Mivel érvényes a $\{\omega\colon |\xi_n(\omega) - \xi(\omega)| < \varepsilon\} \supset A_n$ reláció, ezért $P(|\xi_n(\omega) - \xi(\omega)| < \varepsilon) \to 1$, 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.

Nem kötelező házi feladat:

Valószínűségi változók ξ_n, $n = 1, 2, \dots,$ sorozata, akkor és csak akkor konvergál egy valószínűséggel egy ξ valószínűségi változóhoz, ha az $\eta_n = \sup\limits_{k\ge n} |\xi_k - \xi|$ valószínűségi változók sorozata sztochasztikusan konvergál nullához, azaz minden $\varepsilon > 0$ számra $P\left(\lim\limits_{n\to\infty} \sup\limits_{k\ge n} |\xi_k - \xi| > \varepsilon\right) = 0.$

Lássunk példát arra, hogy lehetséges olyan ξ_n, $n = 1, 2, \dots,$ és ξ valószínűségi változókat konstruálni, amelyekre 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

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

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

A sztochasztikus konvergencia és eloszlásban való konvergencia közötti kapcsolatra érvényesek a következő nem kötelező házi feladatok formájában megfogalmazott állítások.

Nem kötelező házi feladat:

a.) 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.

b.) Ennek az állításnak a megfordítása nem igaz. Például, ha ξ_n, $n = 1, 2, \dots$, független, egyforma eloszlású valószínűségi változók, amelyek nem elfajultak, azaz nincs olyan konstans amelyeket ezek a valószínűségi változók egy valószínűséggel vesznek fel, akkor a ξ_n valószínűségi változók eloszlásban konvergálnak a ξ_1 valószínűségi változóhoz, viszont nem konvergálnak sztochasztikusan.

c.) Viszont igaz a következő állítás: Ha ξ_n, $n = 1, 2, \dots$, valószínűségi változók eloszlásban konvergálnak egy a konstanshoz, (azaz egy olyan valószínűségi változóhoz, amely egy valószínűséggel az a konstanst veszi fel, akkor a ξ_n, $n = 1, 2, \dots$, sorozat sztochasztikusan is konvergál ehhez az a konstanshoz.

Most rátérek annak tárgyalására, hogy hogyan lehet a nagy számok gyenge törvényének a Csebisev egyenlőtlenségen alapuló bizonyítását finomítva a nagy számok erős törvényét is bebizonyítani alkalmas feltételek mellett. Az alább ismertetendő érvelésben valójában a tétel bizonyítása érdekében a szükségesnél sokkal erősebb kikötéseket teszek. Idézzük fel a Csebisev egyenlőtlenséget, illetve a nagy számok gyenge törvényének a bizonyítását a Csebisev egyenlőtlenség segítségével.

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, amely szerint $P(|\xi| > x) \le \frac{E|\xi|}{x}$. 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 $P(|\xi| > x) \le \frac{E\xi^{2k}}{x^{2k}}$

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, amelyet 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.

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

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

$$
\begin{aligned}
&+4 \sum_{\substack{1\le j,k,l\le n \\ j,k,l \text{ különböző számok}}} \underbrace{E\xi_j^2 E\xi_k E\xi_l}_{=0} \\
&+ \sum_{\substack{1\le j,k,l,m\le n \\ j,k,l,m \text{ különböző számok}}} \underbrace{E\xi_j E\xi_k E\xi_l E\xi_m}_{=0} \\
&= nE\xi_1^4 + 6n(n-1)(E\xi_1^2)^2,
\end{aligned}
$$

ezért

$$
P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \le \frac{\frac1n E\xi_1^4 + 6\left(1-\frac1n\right)(E\xi_1)^2\big)}{n^2\varepsilon^4} \le \frac{\text{const.}}{n^2\varepsilon^4}.
$$

Innen 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, amelyet az alábbi tételben fogalmazok meg.

A nagy számok erős törvényéről szóló tétel egy gyenge formája. *Legyen $\xi_1, \xi_2, \dots,$ független egyforma eloszlású valószínűségi változók sorozata egy $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyekre 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.*

Megfogalmazom a nagy számok gyenge törvényét kimondó tételt eredeti, éles formájában is.

A nagy számok erős törvényét kimondó tétel. *Legyen ξ_1, ξ_2, $\dots$, 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*

$$
\lim_{n\to\infty} \frac{S_n(\omega)}{n} = E\xi_1 \quad \textit{majdnem minden } \omega \in \Omega\textit{-ra.}
$$

Láttuk, hogy a nagy számok különböző törvényeit kimondó tételekben az összeadandók momentumairól teszünk fel bizonyos feltételeket. Sőt, az utoljára kimondott tétel a nagy számok erős törvényéről, egyben jelzi azt, hogy ez a feltétel nemcsak elégséges, hanem szükséges feltétel is. Felmerülhet a kérdés, miért játszik olyan fontos szerepet a momentumok létezése ezekben az eredményekben.

A nagy számok törvényei olyan állítást fogalmaznak meg, hogy független valószínűségi változók átlagai egyfajta tipikus viselkedést mutatnak, amelyikben az egyes összeadandók hatása önmagában nem jelentős, azaz nagy valószínűséggel nem fordulhat elő, hogy van az összegben olyan kiugróan nagy értéket felvevő tag, amelynek hatása összemérhető az összes többi tag hatásával. Márpedig az, hogy egy valószínűségi változó valamely momentuma kisebb egy adott számnál olyan tulajdonságot fogalmaz meg, amely szerint a valószínűségi változó kis valószínűséggel vesz fel nagy értékeket. Ahhoz, hogy finomabb vizsgálatokat el tudjunk végezni, ki kell tudnunk fejezni a momentumok végességét az eloszlásfüggvények nyelvén. A következő lemma állítása érdekes lesz a nagy számok erős törvényének a bizonyításában.

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

Megjegyzés. Az alább ismertetendő bizonyítás finomításával be lehet látni, hogy egy $|\xi|$ valószínűségi változónak akkor és csak akkor van véges második momentuma, ha $\sum\limits_{n=1}^{\infty} nP(|\xi|>n)<\infty$. Általánosabban $E|\xi|^\alpha<\infty$ valamilyen $\alpha>0$, számra akkor és csak akkor, ha $\sum\limits_{n=1}^{\infty} n^{\alpha-1}P(|\xi|>n)<\infty$. De ezen állítások bizonyítását, amelyekre nem lesz szükségünk, elhagyom.

A Lemma bizonyítása: Tekintsük először azt a speciális esetet, amikor a ξ valószínűségi változó csak egész értékeket vesz fel. Ekkor az $E|\xi|<\infty$ reláció definició szerint akkor és csak akkor teljesül, ha $\sum\limits_{n=0}^{\infty} nP(|\xi|=n)<\infty$. Ezt úgy is írhatjuk, hogy $\sum\limits_{n=0}^{\infty} n\left(P(|\xi|>n-1)-P(\xi>n)\right)<\infty$, és bármely rögzített N egész számra

$$\sum_{n=0}^{N} nP(|\xi|=n)=\sum_{n=0}^{N} n\left(P(|\xi|>n-1)-P(|\xi|>n)\right)=\sum_{n=0}^{N-1} P(|\xi|>n)-NP(|\xi|>N). \tag{$*$}$$

Ha $E|\xi|=\infty$, akkor $\sum\limits_{n=1}^{\infty} P(|\xi|>n)=\lim\limits_{N\to\infty}\sum\limits_{n=0}^{N-1} P(|\xi|>n)\ge \lim\limits_{N\to\infty}\sum\limits_{n=0}^{N} nP(|\xi|=n)=\infty$ a $(*)$ reláció alapján. Ha $E|\xi|<\infty$, akkor $NP(|\xi|>N)\le E|\xi|$ minden N-re, azaz az előző kifejezésnek van egy N-től független felső becslése. Ezért $\sum\limits_{n=1}^{\infty} P(|\xi|>n)=\lim\limits_{N\to\infty}\sum\limits_{n=0}^{N-1} P(|\xi|>n)\le \lim\limits_{N\to\infty}\sum\limits_{n=0}^{N} nP(|\xi|=n)+\text{const.}<\infty$, szintén a $(*)$ reláció alapján.

Egy általános ξ valószínűségi változó esetén definiáljuk a ξ_1 és ξ_2 valószínűségi változókat úgy, hogy $k\le|\xi|<k+1$ esetén $\xi_1=k$, $\xi_2=k+1$, $k=0,1,2,\dots$. Ekkor $\xi_1\le|\xi|\le\xi_2$, és $\xi_2-\xi_1\equiv 1$, ezért az $E|\xi_1|$, $E|\xi|$ és $E|\xi_2|$ várható értékek egyszerre

végesek vagy végtelenek. Továbbá $\sum\limits_{n=0}^{\infty} P(|\xi|>n)=\sum\limits_{n=0}^{\infty} P(\xi_2>n)$, ezért a Lemma állítása az általános esetben következik annak már bebizonyított speciális esetéből.

Megmutatom az előző lemma eredményének és a Borel–Cantelli lemmának a segítségével, hogy amennyiben $E|\xi_1|=\infty$, akkor a ξ_1 valószínűségi változóval azonos eloszlású ξ_j, $j=1,2,\dots$, független valószínűségi változók nem teljesítik a nagy számok erős törvényét. Pontosabban azt állítom, hogy ebben az esetben az $S_n(\omega)=\sum\limits_{j=1}^{n}\xi_j(\omega)$, $n=1,2,\dots$, összegekre az $S_n(\omega)$ majdnem minden $\omega\in\Omega$ pontban nem konvergens.

Valóban, az előbb megfogalmazott eredmény, és a ξ_j valószínűségi változók azonos eloszlása miatt az $E|\xi_1|=\infty$ feltételből következik, hogy $\infty=\sum\limits_{n=1}^{\infty}P(|\xi_1|>n)=\sum\limits_{n=1}^{\infty}P(|\xi_n|>n)$. Ezért, és a ξ_n valószínűségi változók függetlensége miatt 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.

Tegyük fel, hogy valamely $\omega\in\Omega$ elemi eseményre $\lim\limits_{n\to\infty}\frac{S_n(\omega)}{n}=a$ valamilyen véges a számra, amelynek értéke függhet az ω elemi eseménytől. Ekkor a $\lim\limits_{n\to\infty}\frac{S_{n-1}(\omega)}{n}=a$ reláció is teljesül, ahonnan $\lim\limits_{n\to\infty}\left(\frac{S_n(\omega)}{n}-\frac{S_{n-1}(\omega)}{n}\right)=\lim\limits_{n\to\infty}\frac{\xi_n(\omega)}{n}=0$. Ez a reláció viszont, mint láttuk egy 1 valószínűségi halmazon nem teljesül.

Láttuk, hogy mind a nagy számok gyenge mind a nagy számok erős törvényének bizonyításához arra van szükségünk, hogy képesek legyünk független (és egyforma eloszlású), nulla várható értékű $\xi_1,\xi_2,\dots$, valószínűségi változók $S_n=\sum\limits_{j=1}^{n}\xi_j$, $n=1,2,\dots$, összegének eloszlásfüggvényére jó becslést tudjunk adni. Pontosabban arra van szükségünk, hogy jól meg tudjuk becsülni minden n indexre és rögzített $\varepsilon>0$ számra a $P(|S_n|>\varepsilon n)$ valószínűségeket. Valójában a nagy számok erős törvényének a bizonyításában hasznosabb, ha a $P\left(\sup\limits_{1\le m\le n}|S_m|>\varepsilon n\right)$ valószínűségeket tudjuk jól megbecsülni. Első látásra azt gondolhatnánk, hogy az ilyen valószínűségek becslésére alkalmazhatnánk a centrális határeloszlástételt. Valójában a helyzet bonyolultabb.

A centrális határeloszlástétel jó aszimptotikát ad a $P(S_n>x\sqrt{n})$ valószínűségekre nagy n indexre, és rögzített x számra. De szabad-e rögzített x helyett n-től függő $x_n=\varepsilon\sqrt{n}$ (azaz $\sqrt{n}x_n=\varepsilon n$) számot írni, és alkalmazni formálisan a centrális határeloszlásban szereplő képletet nem törődve az x_n számnak az n indextől való függésével? Azt, hogy e kérdés felvetésénél nem formális kötözködésről van szó mutatja a következő egyszerű példa: Legyenek $\xi_1,\xi_2\dots$, független egyforma eloszlású valószínűségi változók, amelyekre $P(\xi_1=1)=P(\xi_1=-1)=\frac{1}{2}$, $S_n=\sum\limits_{j=1}^{n}\xi_j$. Ekkor a $p_n(x)=P\left(S_n\ge nx\right)$ valószínűségre $p_n(x)=0$, ha $x>1$, és $p_n(x)=2^{-n}$, ha $x=1$. (Az $x<1$ esetben is jó aszimptotikus formulát lehet adni a $p_n(x)$ mennyiségre, de ahhoz külön vizsgálat lenne szükséges, ezért ezt most nem tesszük.) Ez a példa azt mutatja, hogy a $p_n(x)$ függvény nem úgy viselkedik, ahogy a centrális határeloszlástételből adódó formális analógia sugallná. A $P(S_n \ge nx)$ alakú valószínűségek vizsgálatával a valószínűségszámítás egyik fontos és érdekes fejezete a nagy eltérések elmélete foglalkozik. Ez az elmélet nem témája a jelen előadássorozatnak.

A nagy számok erős törvényének a bizonyítását (amikor annak éles változatát vizsgáljuk, és csak a tétel szükséges feltételeit használjuk a bizonyításban) az előadás fő részében nem dolgozom ki. Viszont megfogalmazom a valószínűségszámítás egy fontos egyenlőtlenségét az úgynevezett Kolmogorov egyenlőtlenséget. Ennek bizonyítását megadom egy *Kiegészítésben*, és ott ismertetem azt is, hogyan lehet ennek segítségével bebizonyítani a nagy számok nerős törvényét.

Annak szükséges és elégséges feltétele is ismert, hogy független egyforma eloszlású valószínűségi változók teljesítsék a nagy számok gyenge törvényét. Megfogalmazom ezt az eredményt, de az állítás bizonyítását elhagyom. Ennek az eredménynek az ismeretét sem várom el a vizsgán.

Tétel. *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 $\frac{1}{n}\sum\limits_{k=1}^{n}\xi_k$ átlag akkor és csak akkor konvergál sztochasztikusan $n\to\infty$ esetében egy a számhoz, $-\infty < a < \infty$, ha*

$$\lim_{x\to\infty} x\left[F(-x)+(1-F(x))\right]=0, \quad \textit{és} \quad \lim_{u\to\infty}\int_{-u}^{u} xF(\,dx)=a.$$

Ha összehasonlítjuk a nagy számok gyenge és erős törvényét azt látjuk, hogy a nagy számok gyenge törvénye némileg enyhébb megkötések mellett érvényes mint a nagy számok erős törvénye. Érdemes látni független, egyforma eloszlású valószínűségi változók olyan sorozatát, amely teljesíti a nagy számok gyenge, de nem teljesíti a nagy számok erős törvényét. Ilyen példát fogalmazok meg a következő nem kötelező házi feladatban. Egyben jelzem, hogyan lehet bebizonyítani azt, hogy ez a példa teljesíti a nagy számok gyenge törvényét anélkül, hogy felhasználnánk a nem bizonyított tételt a nagy számok gyenge törvényének szükséges és elégséges feltételéről.

Nem kötelező házi feladat

Legyen ξ_k, $k=1,2,\dots$, független, egyforma eloszlású valószínűségi változók sorozata, az $f(x)=\frac{C}{x^2\log|x|}$, ha $|x|>2$. $f(x)=0$, ha $|x|\le 2$, képlettel megadott sűrűségfüggvénnyel. $\left(\int_{|x|>2}\frac{C\,dx}{x^2\log|x|}=1.\right)$ Definiáljuk az $S_n=\sum\limits_{k=1}^{m}\xi_k$, $n=1,2,\dots$, részletösszegeket. Lássuk be, hogy $E|\xi_1|=\infty$, ezért ezek a valószínűségi változók nem teljesítik a nagy számok erős törvényét. Másrészt az $\frac{S_n}{n}$ átlagok sztochasztikusan tartanak nullához, azaz minden $\varepsilon>0$ számra $P\left(\left|\frac{S_n}{n}\right|>\varepsilon\right)\to 0$, ha $n\to\infty$.

Segítség: Legyen $\bar{\xi}_k = \bar{\xi}_k^{(n)} = \xi_k I(|\xi_k| \le a_n)$, $\bar{\bar{\xi}}_k = \bar{\bar{\xi}}_k^{(n)} = \xi_k I(|\xi_k| > a_n)$, 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 $P(|S_n| > \varepsilon n) \le P\left(|\bar{S}_n| > \frac{\varepsilon}{2} n\right) + P\left(|\bar{\bar{S}}_n| > \frac{\varepsilon}{2} n\right) \le \frac{\operatorname{Var} \bar{S}_n}{\frac{\varepsilon^2}{4} n^2} + nP\left(\bar{\bar{\xi}}_1 \ne 0\right)$. Adjunk az a_n konstans alkalmas megválasztásával (például $a_n = n$) jó becslést a $P(|S_n| > n\varepsilon)$ valószínűségre.

Természetesen felvetődik a következő kérdés: Legyenek $\xi_1, \xi_2, \dots,$ független egyforma eloszlású valószínűségi változók, amelyekre $E\xi_1 = 0$. Legyen $S_n = \sum\limits_{k=1}^{n} \xi_k$, $n = 1, 2, \dots$. Ekkor a nagy számok gyenge törvénye szerint az $\frac{S_n}{n}$ átlagok sztochasztikusan tartanak nullához, a nagy számok erős törvénye szerint pedig egy valószínűséggel is nullához tartanak. Hogyan lehet élesíteni a nagy számok gyenge illetve erős törvényét? Pontosabban fogalmazva, milyen a_n, $n = 1, 2, \dots$ sorozatokra mondhatjuk, hogy $\frac{S_n}{a_n}$ sztochasztikusan tart nullához, ha $n \to \infty$? Milyen b_n, $n = 1, 2, \dots$ sorozatokra mondhatjuk, hogy $\frac{S_n}{b_n}$ egy valószínűséggel tart nullához, ha $n \to \infty$? Az egyszerűség kedvéért tegyük fel, hogy olyan valószínűségi változókat tekintünk, amelyeknek van második momentumuk.

Az első kérdésre a válasz viszonylag egyszerű következménye a centrális határeloszlástételnek, és ezt tartalmazza a következő feladat. A második kérdésre a választ nehezebb megadni. Erre a kérdésre az alább megfogalmazott bizonyítás nélkül közölt iterált logaritmustétel adja meg a választ.

Feladat:

Legyenek $\xi_1, \xi_2, \dots,$ független egyforma eloszlású valószínűségi változók, amelyekre $E\xi_1 = 0$ és $E\xi_1^2 < \infty$. Legyen $S_n = \sum\limits_{k=1}^{n} \xi_k$, $n = 1, 2, \dots$. Ekkor pozitív valós számok valamely a_n sorozatára a $\frac{S_n}{a_n}$ valószínűségi változók akkor és csak akkor tartanak sztochasztikusan nullához $n \to \infty$ esetén, ha $\lim\limits_{n\to\infty} \frac{a_n}{\sqrt{n}} = \infty$.

Végül megfogalmazom a valószínűségszámítás egyik híres eredményét, az úgynevezett iterált logaritmus tételt

Iterált logaritmus tétel. *Legyenek $\xi_1(\omega), \xi_2(\omega), \dots,$ független, egyforma eloszlású valószínűségi változók valamilyen $(\Omega, \mathcal{A}, P)$ valószínűségi mezőn, amelyekre $E\xi_1(\omega) = 0$, $\operatorname{Var} \xi_1(\omega) = \sigma^2 < \infty$. Ekkor*

$$\limsup_{n\to\infty} \frac{\xi_1(\omega) + \cdots + \xi_n(\omega)}{\sqrt{2n\sigma^2 \log\log n}} = 1, \qquad \textit{majdnem minden } \omega \in \Omega \textit{ elemi eseményre,}$$

és

$$\liminf_{n\to\infty} \frac{\xi_1(\omega) + \cdots + \xi_n(\omega)}{\sqrt{2n\sigma^2 \log\log n}} = -1, \qquad \textit{majdnem minden } \omega \in \Omega \textit{ elemi eseményre.}$$

A valószínűségszámítás bevezető előadás rövid áttekintése.

Az előadássorozatot néhány fontos valószínűségszámítási probléma megfogalmazásával kezdtem, és ismertettem a valószínűségszámítás történetének néhány érdekes fordulópontját. A problémák megfogalmazásának célja az volt, hogy képet adjak arról, milyen kérdésekkel foglalkozik a valószínűségszámítás. Ugyancsak fontos része volt az előadásnak néhány kombinatorikai probléma és eredmény ismertetése, amelyek fontos szerepet játszanak valószínűségszámítási problémák vizsgálatában. A problémák és módszerek jobb megértése érdekében felelevenítettem néhány analízisbeli fogalmat, mindenekelőtt a Taylor sor fogalmát, illetve néhány a végtelen halmazok számosságával kapcsolatos fontos fogalmat és eredményt.

Ezután rátértem a valószínűségszámítás precíz, Kolmogorov által kidolgozott elméletének az ismertetésére. Ezen elméleten belül különösen fontos annak megértése, hogy hogyan lehet a valószínűségszámítás problémáit e formális, a halmazelmélet nyelvét használó modell keretei között tárgyalni. Ugyancsak fontos annak megértése, hogy a problémák átfogalmazása a Kolmogorov-féle modellbe hogyan segít konkrét feladatok megoldásában. A Kolmogorov-féle elméletet először olyan modellekben tárgyaltuk, amelyben a megfigyelésnek csak véges, esetleg megszámlálhatóan végtelen sok eredménye lehetséges. Annak érdekében, hogy ne csak véges vagy megszámlálhatóan végtelen sok lehetséges kimenetelt tartalmazó modellt tudjuk vizsgálni, szükség volt bizonyos mély mértékelméleti eredmények felidézésére. Példát mutattam arra is, hogy a valószínűségi mérték σ-additívitásából olyan azonosságok is következnek, amelyeket bár be lehet bizonyítani csupán az analízis eredményeinek a felhasználásával, de egy ilyen bizonyítás rendkívül sok számolást igényel.

Elegendő a fogalmak és eredmények jelentésének megértése, illetve azt látni, hogyan lehet ezt feladatok megoldásában használni. Külön tárgyaltam több olyan feladatot, amelyek geometriai meggondolások segítségével egyszerűen megoldhatóak. Ezen megoldások hátterében természetes és szemléletes, de mély mértékelméleti eredmények vannak. Bár e mértékelméleti ereméyek bizonyítását nem tárgyaltam, egy ilyen elveken alapuló (jó) feladatmegoldást teljes értékűnek fogok tekinteni a vizsgán.

Ezután tárgyaltam és bebizonyítottam a Borel–Cantelli lemmát. Itt is elegendő az eredményt megérteni, illetve természetes esetekben ezt az eredmény tudni kell használni. Az előadásban tárgyalt következő fontos fogalom a feltételes valószínűség volt. Nagyon fontos megérteni, hogy hogyan kell ezt a fogalmat konkrét esetekben használni. Az itt kimondott eredményeket jobban meg lehet érteni, ha konkrét feladatokat vizsgálunk.

Bevezettem események függetlenségének a fogalmát is. Az itt kimondott fogalmak, eredmények ismerete rendkívül fontos. A tárgyalás során szóba kerültek mértékelméleti fogalmak és eredmények. Az egyszerűbb fogalmakat, mint például a valószínűségi mező, a σ-algebra fogalmát ismerni kell, de már olyan fogalmaknak az ismeretét, mint például a majdnem-mindenütt teljesülő tulajdonságok, Lebesgue integrál, Fubini tétel elegendő körülbelül tudni, azon a szinten, ami ahhoz kell, hogy a leírt valószínűségi eredményeket értsék és tudják használni. Ezenbelül a Fubini tétel ismerete fontos, mert ezt nagyon sok feladat megoldásában használják. (Az itt használt eredmények bizonyítása és részletes tárgyalása a mértékelmélet és nem a valószínűségszámítás témája. Arról, hogy milyen szigorúsággal követelik meg ezen eredmények ismeretét a mértékelmélet vizsgán nem én döntök.)

Több független eseményekkel kapcsolatos feladatot tárgyaltam. Ilyen feladatok megoldását a vizsgán tudni kell. Ugyancsak tárgyaltam az e problémakör általánosításának tekinthető úgynevezett szita-formulát, és néhány ennek segítségével megoldható problémát. Bár ez a problémakör a valószínűségszámítás temészetes részének tekinthető, ennek ismeretét, illetve (a csak) ezen formula segítségével megoldható feladatok megoldását nem követelem meg a vizsgán.

Ezután konkrét először diszkrét, majd folytonos eloszlású valószínűségi változókat vizsgáltam. Definiáltam valószínűségi változók függetlenségét, várható értékét, szórásnégyzetét és különböző valószínűségi változók kovarianciafüggvényét. Először diszkrét valószínűségi változók esetében definiáltam ezeket a fogalmakat, és csak később vezettem be őket folytonos eloszlású valószínűségi változók esetében. Ennek a tárgyalásmódnak elsősorban metodológiai okai voltak. Ugyanis diszkrét eloszlású valószínűségi változók vizsgálata nem igényli mély mértékelméleti fogalmak és eredmények ismeretét, és sok érdekes és fontos valószínűségszámítás jelenség már ezek segítségével is tárgyalható.

Folytonos eloszlású valószínűségi változók tárgyalásában a legfontosabb fogalmak, mint például a várható érték és szórásnégyzet precíz definiciójához szükséges a Lebesgue integrál fogalma. Vázoltam a Lebesgue integrál definicióját, de e fogalom pontos ismerete (ezen tantárgy keretein belül) nem szükséges. Viszont ki kell tudni számolni véletlen összegek várható értéket, szórásnégyzetet, és független valószínűségi változók összegének a sűrűségfüggvényét (konvolució segítségével). Tudni kell, hogy a felhasznált tételek milyen feltételek mellett érvényesek. E feladatok megoldásának érdekében jól meg kell érteni, hogy hogyan érdemes valószínűségi problémák jó modelljét bevezetni, és hogyan kell velük számolni. Ezenkívül tudni kell az analízisben tanult integrál és differenciálszámítás legfontosabb eredményeit is.

Bevezettem a legfontosabb diszkrét és folytonos eloszlásokat, Ezekkel tudni kell számolni, illetve ezen eloszlások valószínűségi tartalmát is tudni és érteni kell.

Tárgyaltam a Markov és Csebisev egyenlőtlenséget. Ezután az előadás fontos témája volt az eloszlások fogalma, legfontosabb tulajdonságai, a valószínűségszámításban szereplő különböző konvergenciafogalmak (eloszlásban való konvergencia, sztochasztikus konvergencia, egy valószínűséggel való konvergencia) valamint azok jelentése és kapcsolata. Ugyancsak fontos megérteni a nagy számok gyenge és erős törvényét, azt hogy mit mondanak ki ezek az eredmények. A bizonyítások részleteinek ismerete nem szükséges.

Az előadás legfontosabb témája a centrális határeloszlástétel volt. Ennek tárgyalása érdekében bevezettem a karakterisztikus függvény fogalmát. Ennek erészletes tárgyalására azonban nem maradt idő. Ezenkívül a Fourier sorok elmélete, amely jobban megvilágítja az itt szereplő fogalmakat és a bizonyítások hátterét sokaknak nem szerepelt az eddigi tananyagában. Ezért ennek az elméleti résznek az ismerete a vizsgán nem szükséges. Viszont nagyon fontos tudni azt, hogy hogyan és mire lehet használni a centrális határeloszlástételt. Ebbe beletartoznak az előadáson tárgyalt statisztikai problémák is.

Végül a több-dimenziós centrális határeloszlástételt és a χ-négyzet próbát tárgyaltam. Ezen belül ismertettem bizonyos fogalmak (eloszlás függvény, sűrűségfüggvény, eloszlásban való konvergencia, várható érték, szórásnégyzet) több-dimenziós megfelelőit. Ezen fogalmakat és a hozzájuk kötödő eredményeket ismerni kell. Viszont a bizonyításokat nem kérem, elegendő ezek bizonyos durva áttekintését ismerni.

Végül megjegyzem, hogy a vizsga elsősorban feladatok megoldásából fog állni. Ezért a felkészüléshez azt javaslom, hogy ne csak a definiciókat, tételeket, hanem a tárgyalt feladatokat is tanulmányozzák. Ez utóbbiak is fontos részei az előadássorozatomnak, amelynek részletes leírása megtalálható a homepage-emen.

Homepage-em címe: http://www.renyi.hu/~major

Kiegészítés. A nagy számok erős törvényének a bizonyítása.

Ebben a kiegészítésben bebizonyítom a nagy számok törvényének elégségességről szóló felét az alább megfogalmazott Kolmogorov egyenlőtlenség segítségével. Pontosabban, azt látom be, hogy ha $\xi_1,\xi_2,\dots,$ független egyforma eloszlású valószínűségi változók sorozata, amely teljesíti az $E|\xi_1|<\infty$ feltételt, akkor ez a sorozat teljesíti a nagy számok erős törvényét is.

Kolmogorov egyenlőtlenség. *Legyenek $\xi_1,\dots,\xi_n$ független, (nem feltétlenül egyforma eloszlású) valószínűségi változók, $E\xi_k=0$, $E\xi_k^2=\sigma_k^2$. Tekintsük e valószínűségi változók $S_k=\sum\limits_{p=1}^k \xi_p$, $k=1,\dots,n$, részletösszegeit. Ekkor*

$$P\left(\sup_{1\le k\le n}|S_k|\ge x\right)\le\frac{ES_n^2}{x^2}=\frac{\sum\limits_{k=1}^n\sigma_k^2}{x^2}$$

minden $x>0$-ra.

A Kolmogorov egyenlőtlenség bizonyítása. Rögzítsünk egy $x>0$ számot, és vezessük be a következő $\tau=\tau(n,x)$ véletlen időpontot: $\tau(\omega)=k$ valamely $1\le k\le n$ számra, ha $|S_k(\omega)|\ge x$, és $|S_j(\omega)|<x$ minden $1\le j<k$ számra, és $\tau(\omega)=n$, ha $|S_j(\omega)|<x$ minden $1\le j\le n$ indexre. A τ véletlen időpont (megállási szabály) szemléletes tartalma a következő: Az első olyan $k\le n$ időpontot keressük, amelyre $S_k\ge x$. Ha van ilyen időpont, akkor τ ezzel az időponttal egyenlő. Ha nincs ilyen időpont, akkor $\tau=n$. Azt állítom, hogy érvényes az

$$ES_\tau^2\le ES_n^2 \tag{b}$$

egyenlőtlenség. Ennek bizonyítása érdekében vezessük be a $\tau_j=\max(\tau,j)$, $1\le j\le n$, valószínűségi változókat. Mivel $\tau_1=\tau$, és $\tau_n=n$, ezért a (b) reláció bizonyításához elég megmutatni, hogy $ES^2_{\tau_j}\le ES^2_{\tau_{j+1}}$ minden $1\le j<n$ indexre. Viszont $S_{\tau_j}(\omega)=S_{\tau_{j+1}}(\omega)$, ha $\tau>j$, és $S_{\tau_j}(\omega)=S_j(\omega)$, $S_{\tau_{j+1}}(\omega)=S_{j+1}(\omega)=S_j(\omega)+\xi_{j+1}(\omega)$, ha $\tau\le j$. Ezért elég megmutatni, hogy

$$\begin{aligned}ES^2_{\tau_{j+1}}(\omega)-ES^2_{\tau_j}(\omega)&=E\left[\left(S_{\tau_j}(\omega)+\xi_{j+1}(\omega)\right)^2-S^2_{\tau_{j+1}}(\omega)\right]I_A(\omega)\\&=2EI_A(\omega)S_j(\omega)\xi_{j+1}(\omega)+E\xi^2_{j+1}(\omega)I_A(\omega)\ge0,\end{aligned}$$

ahol A az $\{\omega\colon\ \tau(\omega)\le j\}$ esemény, és $I_A(\omega)$ az A halmaz indikátor függvénye. Ennek érdekében vegyük észre, hogy a $\xi_{j+1}(\omega)$ és $I_A(\omega)S_j(\omega)$ valószínűségi változók függetlenek, mivel a második valószínűségi változó a $\xi_1(\omega),\dots,\xi_j(\omega)$ valószínűségi változók függvénye. Innen következik, hogy

$$EI_A(\omega)S_j(\omega)\xi_{j+1}(\omega)=EI_A(\omega)S_j(\omega)E\xi_{j+1}(\omega)=0.$$

Másrészt $E\xi^2_{j+1}(\omega)I_A(\omega)\ge 0$. Innen következik a bizonyítandó egyenlőtlenség és a (b) reláció.

A Kolmogorov egyenlőtlenség egyszerűen következik a (b) relációból és a Csebisev egyenlőtlenségből. Valóban, mivel $\left\{\omega\colon \sup\limits_{1\le k\le n}|S_k(\omega)|\ge x\right\}=\{\omega\colon\ |S_\tau(x)|\ge x\}$, ezért

$$P\left(\sup_{1\le k\le n}|S_k|\ge x\right)=P\left(|S_\tau|\ge x\right)\le\frac{ES_\tau^2}{x^2}\le\frac{ES_n^2}{x^2}.$$

Érdemes megjegyezni, hogy a Kolmogorov egyenlőtlenség ugyanazt a felső becslést adja annak valószínűségére, hogy az S_k részletösszegek szuprémuma nagyobb mint valamilyen $x>0$ szám, mint amit a Csebisev egyenlőtlenség ad annak az eseménynek a valószínűségére, hogy az utolsó tag S_n nagyobb, mint x. Természetesen

$$P\left(\sup_{1\le k\le n}|S_k|>x|\right)\ge P\left(S_n|>x|\right),$$

de mint ahogy a Kolmogorov és Csebisev egyenlőtlenség összehasonlítása is sugallja a fenti egyenlőtlenség baloldala nem sokkal nagyobb mint a jobboldala.

A nagy számok törvényének bizonyításában hasznos a következő technikai jellegű lemma.

Kronecker lemma: *Ha az a_n és q_n, $n=1,2,\dots$ számsorozatok olyanok, hogy a $\sum\limits_{n=1}^{\infty}a_n$ összeg konvergens, a q_n sorozat monoton nő és $\lim\limits_{n\to\infty}q_n=\infty$, akkor $\lim\limits_{n\to\infty}\frac{1}{q_n}\sum\limits_{k=1}^{n}a_kq_k=$ 0.*

A Kronecker lemma bizonyítása. Vezessük be az $R_n=\sum\limits_{j=n}^{\infty}a_j$, $n=1,2,\dots$, összegeket, és írjuk fel az $\frac{1}{q_n}\sum\limits_{k=1}^{n}a_kq_k=\frac{1}{q_n}\sum\limits_{k=1}^{n}(R_k-R_{k+1})q_k=\frac{1}{q_n}\sum\limits_{k=1}^{n}R_k(q_k-q_{k-1})-R_{n+1}$ azonosságot. (A fenti típusú számolás, amelyet Abel-féle átrendezésnek hívnak, a parciális integrálás diszkrét megfelelője.) Felhasználva, hogy létezik olyan $R<\infty$ szám, amelyre $|R_n|<R$ minden $n=1,2,\dots$ indexre, minden $\varepsilon>0$ számhoz létezik olyan

$n_0 = n_0(\varepsilon)$ index, amelyre $|R_n| \le \varepsilon$, ha $n \ge n_0$, és $q_k - q_{k-1} \ge 0$ minden $k = 1, 2, \dots,$ számra, az előző azonosság alapján felírhatjuk a következő egyenlőtlenséget.

$$\frac{1}{q_n}\left|\sum_{k=1}^{n} a_k q_k\right| \le \frac{1}{q_n}\sum_{k=1}^{n_0} R(q_k - q_{k-1}) + \frac{1}{q_n}\sum_{k=n_0+1}^{n} \varepsilon(q_k - q_{k-1}) + \varepsilon \le R\frac{q_{n_0}}{q_n} + 2\varepsilon,$$

ha $n \ge n_0$. Mivel ez az egyenlőtlenség minden $\varepsilon > 0$ számra érvényes, és $\lim_{n\to\infty} q_n = \infty$, innen $n \to \infty$ határátmenettel következik a Kronecker lemma állítása.

A nagy számok erős törvényének be nem bizonyított részének igazolásához (ahhoz, hogy véges várható érték létezése esetén igaz a nagy számok erős törvénye) elegendő a következő állítást bebizonyítani.

Tétel a nagy számok erős törvényéről. *Ha ξ_n, $n = 1, 2, \dots,$ független egyforma eloszlású valószínűségi változók sorozata, $E|\xi_1| < \infty$, $E\xi_1 = 0$, $S_n = \sum_{k=1}^{n} \xi_k$, $n = 1, 2, \dots,$ akkor $\lim_{n\to\infty} \frac{S_n}{n} = 0$ egy valószínűséggel.*

A nagy számok erős törvényéről szóló tétel bizonyítása. Vezessük be a tételben szereplő ξ_n valószínűségi változók következő felbontását: $\xi_n = \xi'_n + \xi''_n$, ahol $\xi'_n = \xi_n I(|\xi_n| \le n)$, $\xi''_n = \xi_n I(|\xi_n| > n)$, és $I(A)$ az A halmaz indikátorfüggvényét jelöli. Megmutatom, hogy a tétel állításának bizonyítását redukálni lehet a $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \xi'_n = 0$ egy valószínűséggel, sőt a $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\xi'_n - E\xi'_n) = 0$ egy valószínűséggel reláció bizonyítására.

Az első redukció igazolásához elég észrevenni, hogy

$$\sum_{k=1}^{\infty} P(\xi''_k \ne 0) = \sum_{k=1}^{\infty} P(|\xi_k| > k) = \sum_{k=1}^{\infty} P(|\xi_1| > k) \le 1 + E|\xi_1| < \infty,$$

ezért a Borel–Cantelli lemma alapján egy valószínűséggel csak véges sok k-ra teljesül a $\xi''_k \ne 0$ esemény. Innen $\frac{1}{n} \sum_{k=1}^{n} \xi''_k \to 0$ egy valószínűséggel.

A második redukció igazolásához elég megmutatni, hogy

$$\left|\frac{1}{n}\sum_{k=1}^{n} E\xi'_k\right| = \left|-\frac{1}{n}\sum_{k=1}^{n} E\xi''_k\right| \le \frac{1}{n}\sum_{k=1}^{n} E|\xi_1| I(|\xi_1| \ge k) \to 0,$$

ha $n \to \infty$, mert $E|\xi_1| < \infty$, és ezért $E|\xi_1| I(|\xi_1| > k) \to 0$, ha $k \to \infty$.

A redukált állítást érdemes a Kronecker lemma segítségével visszavezetni a következő állítás igazolására. A tétel feltételeinek teljesülése esetén a $\sum_{n=1}^{\infty} \frac{\xi'_n - E\xi'_n}{n}$ összeg egy valószínűséggel konvergens. Valóban, ebből az állításból és a Kronecker lemmából $a_n = \frac{\xi'_n - E\xi'_n}{n}$ és $q_n = n$ választással következik, hogy $\lim_{n\to\infty} \sum_{k=1}^{n} \frac{\xi'_k - E\xi'_k}{n} = 0$ egy valószínűséggel.

Ezen állítás igazolásához elég belátni, hogy tetszőleges $\varepsilon > 0$ számra

$$P\left(\sup_{N\ge n}\left|\sum_{k=n}^{N}\frac{\xi_k'-E\xi_k'}{k}\right|>\varepsilon\right)\to 0 \quad \text{ha } n\to\infty.$$

Innen ugyanis következik, hogy

$$P\left(\sup_{N,M\colon N\ge n,\, M\ge n}\left|\sum_{k=M}^{N}\frac{\xi_k'-E\xi_k'}{k}\right|>2\varepsilon\right)\to 0 \quad \text{ha } n\to\infty.$$

Mivel ez az egyenlőtlenség alkalmazható minden $\varepsilon=\frac{1}{k}$, $k=1,2,\dots,$ számra, innen következik, hogy a $\sum\limits_{k=1}^{n}\frac{\xi_k'-E\xi_k'}{k}$ sorozat egy valószínűséggel Cauchy sorozat, és a $\sum\limits_{k=1}^{\infty}\frac{\xi_k'-E\xi_k'}{k}$ összeg egy valószínűséggel konvergens.

Az igazolandó egyenlőtlenséget a Kolmogorov egyenlőtlenség segítségével igazolhatjuk. Ugyanis a Kolmogorov egyenlőtlenség alapján

$$P\left(\sup_{N\ge n}\left|\sum_{k=n}^{N}\frac{\xi_k'-E\xi_k'}{k}\right|>\varepsilon\right)\le\frac{1}{\varepsilon^2}\sum_{k=n}^{\infty}\frac{\operatorname{Var}\xi_k'}{k^2}\le\frac{1}{\varepsilon^2}\sum_{k=n}^{\infty}\frac{E{\xi_k'}^2}{k^2}.$$

Ezért a kívánt egyenlőtlenség igazolásához elég megmutatni, hogy

$$\sum_{k=1}^{\infty}\frac{E{\xi_k'}^2}{k^2}<\infty.$$

Ez az egyenlőtlenség következik az alábbi számolásból.

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{1}{k^2}E{\xi_k'}^2&\le\sum_{k=1}^{\infty}\frac{1}{k^2}E\xi_1^2 I(|\xi_1|>k)\le\text{const.}\sum_{j=1}^{\infty}j^2P(j\le|\xi_1|<j+1)\left(\sum_{k=j}^{\infty}\frac{1}{k^2}\right)\\&\le\text{const.}\sum_{j=1}^{\infty}jP(j\le|\xi_1|<j+1)\le\text{const.}\,E|\xi_1|<\infty.\end{aligned}$$