Valószínűségszámítás

10. előadás

Arató Miklós

2024.04.22.

Tartalomjegyzék

CHT

Tétel [Centrális határeloszlástétel független, azonos eloszlású változókra]: Legyenek $\xi_1, \xi_2, \ldots$ független, azonos eloszlásúak, $m := E\xi_1$ és $0 < \sigma^2 = D^2\xi_i < \infty$. Ekkor minden $x \in \mathbb{R}$-re

$$P\left(\frac{\sum_{i=1}^{n} \xi_i - n \cdot m}{\sigma \cdot \sqrt{n}} < x\right) \to \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}\, dt$$

Normális közelítés

Példa 1.

Egy országban meg szeretnék becsülni a koronavírussal fertőzöttek arányát úgy, hogy legalább $0,95$ valószínűséggel legfeljebb 1%-ot tévedjünk.

Példa 1.

Egy országban meg szeretnék becsülni a koronavírussal fertőzöttek arányát úgy, hogy legalább $0,95$ valószínűséggel legfeljebb 1%-ot tévedjünk. Jelölje N az ország lakosai, M a koronavírusosak számát, n pedig a megvizsgáltak számát, ekkor $p := \frac{M}{N}$-et akarjuk jól közelíteni. Legyen továbbá X_i értéke 1, ha az i-edik megvizsgált koronavírusos és 0 különben. Ekkor a

$$P\left(\left|\frac{X_1 + \ldots + X_n}{n} - \frac{M}{N}\right| \leq 0,01\right) \geq 0,95$$

egyenlőtlenségnek kell teljesülnie.

Példa 1. (folyt.)

A centrális határeloszlástétel szerint

$$P\left(\frac{-\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}} \leq \frac{\sum_{i=1}^{n} X_i - np}{\sqrt{n}\sqrt{p(1-p)}} \leq \frac{\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}}\right) \sim$$

$$\sim 2\cdot\Phi\left(\frac{\sqrt{n}\cdot 0,01}{\sqrt{p(1-p)}}\right) - 1 \geq 0,95,$$

azaz $\Phi\left(\frac{\sqrt{n}\cdot 0{,}01}{\sqrt{p(1-p)}}\right) \geq 0,975 = \Phi(1,96)$, tehát legyen $\frac{\sqrt{n}\cdot 0{,}01}{\sqrt{p(1-p)}} \geq 1,96$. Ezzel $\sqrt{n} \geq 1,96\cdot\frac{1}{0{,}01}\cdot\sqrt{p(1-p)}$, ha [1] $n \geq 10000$, tehát ha normális közelítéssel dolgozunk, kb. 10 000 embert kell megkérdezni.

[1] mivel a $\sqrt{p(1-p)}$ nem lehet nagyobb $0,5$-nél, így a jobb oldal felülről becsülhető 98-cal

Példa 2.

Mihez tart a következő: $e^{-n}\sum_{k=0}^{n-1}\frac{n^k}{k!}\xrightarrow{n\to\infty}$?

Ha végtelenig összegeznénk, nyilván 1 lenne az összeg, de most $n-1$-ig megyünk!

Példa 2. (folyt.)

Legyenek $\eta_i \sim$ 1-Poisson függetlenek, ezekre teljesülnek a centrális határeloszlástétel feltételei, azaz felírható:

$P\left(\frac{\sum_{i=1}^{n}\eta_i - n\cdot 1}{1\cdot\sqrt{n}} < x\right) \to \Phi(x)$, ahol $\sum_{i=1}^{n}\eta_i \sim n$-Poisson, így

$P\left(\sum_{i=1}^{n}\eta_i < n\right) = e^{-n}\cdot\sum_{k=0}^{n-1}\frac{n^k}{k!}$.

Ekkor speciálisan $x = 0$-ra $P\left(\frac{\sum_{i=1}^{n}\eta_i - n}{\sqrt{n}} < 0\right) \to \Phi(0) = \frac{1}{2}$.

Tehát a keresett összeg $\frac{1}{2}$-hez tart.

CHT független változókra

Tétel [Centrális határeloszlástétel független változókra]: Legyenek
$\xi_1, \xi_2, \ldots$ függetlenek. $m_k := E\xi_k$ és $0 < \sigma_k^2 = D^2\xi_k < \infty$, továbbá $S_n = \xi_1 + \ldots + \xi_n$, $D_n^2 = \sum_{k=1}^{n} \sigma_k^2$. Teljesül az ún. Ljapunov-feltétel:

$$\frac{1}{D_n^{2+\delta}} \cdot \sum_{k=1}^{n} E|\xi_k - m_k|^{2+\delta} \xrightarrow{n\to\infty} 0,$$

valamely $0 < \delta$-ra. Ekkor minden $x \in \mathbb{R}$-re

$$P\left(\frac{S_n - ES_n}{D_n} < x\right) \to \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{t^2}{2}}\, dt$$

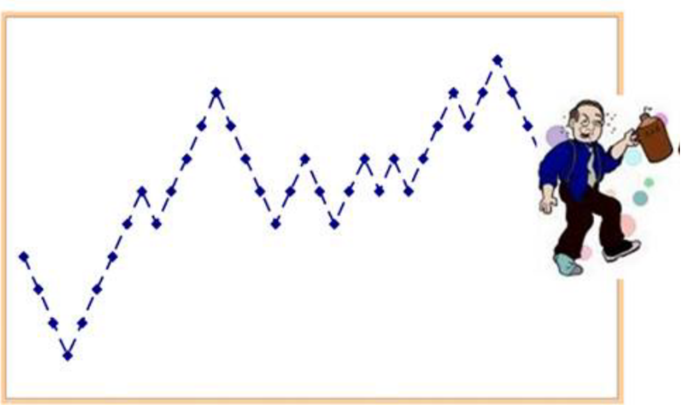
Közelítés hibája

Tétel [Berry-Esséen] Legyenek $\xi_1, \xi_2, \ldots$ független, azonos eloszlású valószínűségi változók, továbbá $E|\xi_1|^3 < \infty$ és $T_n := \frac{\sum\limits_{k=1}^{n} \xi_k - n \cdot m}{\sigma\sqrt{n}}$. Ekkor $\sup\limits_x |F_{T_n}(x) - \Phi(x)| \leq 0,474 \cdot \frac{E|\xi_1 - m|^3}{\sigma^3\sqrt{n}}$.

Tétel [Esséen] Legyenek $\xi_1, \xi_2, \ldots$ független valószínűségi változók, $E|\xi_k|^3 < \infty$, továbbá $S_n = \xi_1 + \ldots + \xi_n$, $D_n^2 = \sum\limits_{k=1}^{n} \sigma_k^2$, $F_n(x) = P\left(\frac{S_n - ES_n}{D_n} < x\right)$, $L_n = \frac{\sum_{k=1}^{n} E|X_k - m_k|^3}{B_n^{-3/2}}$.
Ekkor $\sup\limits_x |F_n(x) - \Phi(x)| \leq 0,56 L_n$

Bevezetés

Egy számegyenesen lépegetünk az egészeken 0-ból kiindulva, minden lépésben ugyanakkora eséllyel lépünk balra, mint jobbra.



ábra: Bolyongás a számegyenesen

Tulajdonságok

Legyenek $\xi_1, \xi_2, \dots$ független, azonos eloszlásúak:

$$P(\xi_i = 1) = P(\xi_i = -1) = \frac{1}{2}$$

Ekkor a helyzetünk n lépés után:

$$S_n = \sum_{i=1}^{n} \xi_i$$

Mit tanultunk eddig S_n-ről?

$$E\xi_i = \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot (-1) = 0, D^2\xi_i = E(\xi_i^2) = 1 \Rightarrow ES_n = 0, D^2(S_n) = n.$$

Teljes valószínűség tétel $\Rightarrow$

$$P(\exists n \geq 1 : S_n = 0) = 1$$

Tulajdonságok (folyt.)

Teljes várható érték tétel $\Rightarrow$

$$\tau = \min\{n \geq 1 : S_n = 0\} \Rightarrow E\tau = \infty$$

Nagy számok erős törvénye $\Rightarrow$

$$\frac{S_n}{n} \to 0 \text{ 1 valószínűséggel}$$

Csebisev-egyenlőtlenség $\Rightarrow$

$$P(S_n \geq cn) \leq \frac{n}{c^2 n^2} = \frac{1}{c^2 n}$$

CHT $\Rightarrow$

$$P(\frac{S_n}{\sqrt{n}} < x) \to \Phi(x) \Rightarrow P(y\sqrt{n} < S_n < x\sqrt{n}) \to \Phi(x) - \Phi(y)$$

Tulajdonságok (folyt.)

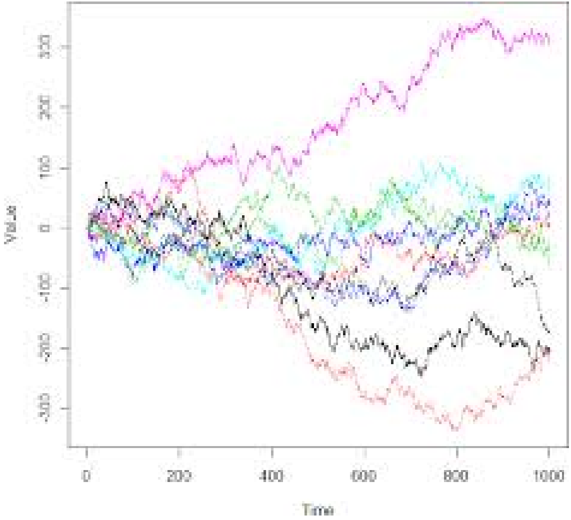
Moivre-Laplace lokális tétel $\Rightarrow$

$$\tilde{\xi}_i = \frac{\xi_i + 1}{2} \Rightarrow P(\tilde{\xi}_i = 0) = P(\tilde{\xi}_i = 1) = \frac{1}{2} \Rightarrow \tilde{S}_n = \sum_{i_1}^{n} \tilde{\xi}_i \sim B(n, \frac{1}{2})$$

$$P(S_n = k) = P\left(\tilde{S}_n = \frac{k+n}{2}\right) \sim \frac{1}{\sqrt{2\pi n \frac{1}{4}}} \exp\left(-\frac{\left(\frac{k+n}{2} - \frac{n}{2}\right)^2}{2n\frac{1}{4}}\right) =$$

$$= \sqrt{\frac{2}{\pi n}} \exp\left(-\frac{k^2}{2n}\right), \text{ ha } -A\sqrt{n} \le k \le A\sqrt{n}$$

Trajektóriák



Pontos valószínűségek

Legyen $0 \leq k \leq n$-ra

$u_{2k} = P(S_{2k} = 0), f_{2k} = P(S_2 \neq 0, \dots, S_{2k-2} \neq 0, S_{2k} = 0) = P(\tau = 2$

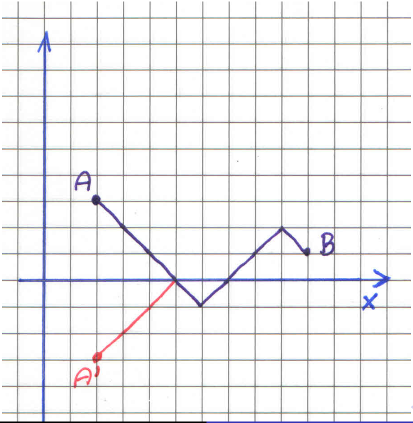
Nyilvánvaló, hogy

$$u_{2k} = \frac{\binom{2k}{k}}{2^{2k}}$$

Állítás: $f_{2k} = \frac{u_{2(k-1)}}{2k} = \frac{\binom{2k-2}{k-1}}{k2^{2k-1}}$
Biz.: $f_{2k} = 2P(S_1 > 0, S_2 > 0, \dots, S_{2k-1} > 0, S_{2k} = 0)$
A, B egész koordinátájú pontok az $x \geq 0$ félsíkon. A-ból B-be vezető út: olyan trajektória, melyet a véletlen bolyongás be tud járni.
Tükrözési elv: legyen A és B az x tengely azonos oldalán lévő két pont. Jelölje A' az A tükörképét az x tengelyre. Ekkor az A-ból B-be vezeti azon utak száma, amelyeknek az x tengellyel van közös pontjuk, megegyezik az A'-ből B-be vezető utak számával

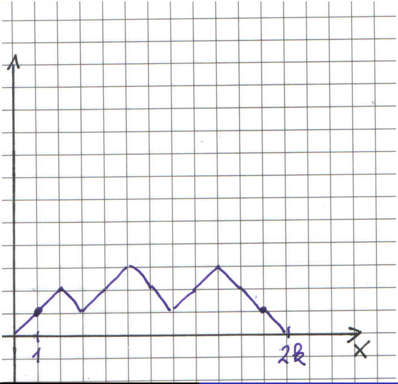
Tükrözési elv bizonyítása

Az $A \to B$ tengelyt metsző útnál tükrözzük az x tengelyre az első metszéspontig terjedő szakaszt. Így egy $A' \to B$ utat kapunk. A megfeleltetés kölcsönösen egyértelmű, tehát a két típusú utak száma azonos.



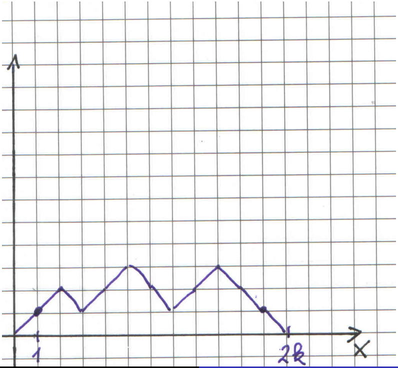
Állítás bizonyításának folytatása

$$P(S_1 > 0, S_2 > 0, \ldots, S_{2k-1} > 0, S_{2k} = 0) =$$
$$= \frac{x \text{ tengelyt nem érintő utak száma } (1,1)\text{-ből } (2k-1,1)\text{-be}}{2^{2k}}$$



Állítás bizonyításának folytatása

$$\{x \text{ tengelyt nem érintő utak száma } (1,1)\text{-ből } (2k-1,1)\text{-be}\} = \{\text{utak száma } (1,1)\text{-ből } (2k-1,1)\text{-be}\} - \{x \text{ tengelyt érintő utak száma } (1,1)\text{-ből } (2k-1,1)\text{-be}\}$$



Állítás bizonyításának folytatása

$$f_{2k} = \frac{\binom{2k-2}{k-1} - \binom{2k-2}{k}}{2^{2k}} = \frac{\binom{2k-2}{k-1}}{k2^{2k-1}}$$