

Valószínűesszámítás és statisztika – Folytonos valószínűségi változók, momentumok és nagy vatszám-tételek egy dimenzióban

Tartalom

- ▶ Folytonos valószínűségi változók, eloszlásfüggvény folytonos esetben, sűrűségfüggvény, és ezek kapcsolata
- ▶ Várható érték folytonos esetben
- ▶ Nemnegatív valószínűségi változók várható értéke
- ▶ Folytonos valószínűségi változók transzformáltjának eloszlása, transzformált várható értéke (momentumok, szórás)
- ▶ Valószínűségi változók szimulációja
- ▶ Markov- és Csebisev-egyenlőtlenség
- ▶ Nevezetes folytonos eloszlások: egyenletes, exponenciális, normális
- ▶ De Moivre–Laplace-tétel, centrális határeloszlás-tétel
- ▶ Nagy számok törvényei

Előismeretek: Differenciál- és integrálszámítás egy dimenzióban, diszkrét valószínűségi változók

Folytonos valószínűségi változók

Emlékeztető: Ha $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mező, akkor valószínűségi változónak az olyan $X: \Omega \rightarrow \mathbb{R}$ függvényt nevezzük, amelyre $\{X < x\} = \{\omega \in \Omega \mid X(\omega) < x\} \in \mathcal{F}$ minden $x \in \mathbb{R}$ esetén.

Eddig: diszkrét val. változók, $\text{Ran}(X)$ megszámlálható (=véges vagy megszámlálhatóan végtelen).

Példák **folytonos** véletlen mennyiségekre, ahol az értékkészlet nem megszámlálható:

- ▶ Geometriai valószínűségi mező, „egy intervallumon *találomra* választunk egy pontot”, valószínűség a “területtel” (intervallumhosszal) arányos $\rightarrow$ a választott pont egy folytonos értékkészletű val. változó
- ▶ Várakozási idők (pl. következő busz érkezéséig, villanykörte kiégéséig)
- ▶ A tőzsdén egy adott részvény értékének változása egy nap alatt (ennek az értéke lehet negatív is)
- ▶ Egy véletlenszerűen választott magyar állampolgár testmagassága, testsúlya stb.

Eloszlásfüggvény

Miután $\{X < x\}$ esemény, értelmezett a $\mathbb{P}(\{X < x\})$ valószínűség, amit egyszerűsítve $\mathbb{P}(X < x)$ -szel jelöltünk már eddig is.

Az $\mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \mathbb{P}(X < x)$ függvény igen hasznos, különösen akkor, amikor $\text{Ran}(X)$ nem véges, pláne ha nem is megszámlálható.

(4.1.2. Definíció.) Legyen X valószínűségi változó. Ekkor az $F_X: \mathbb{R} \rightarrow \mathbb{R}$,

$$F_X(x) \stackrel{\text{def}}{=} \mathbb{P}(X < x) \in [0, 1]$$

függvényt az X eloszlásfüggvényének nevezzük.

Egyszerű példa: szabályos érme feldobása, $\Omega = \{F, I\}$, $X(F) = 1$, $X(I) = 0$.

Rajzoljuk fel az X valószínűségi változó ($X(F) = 1$, $X(I) = 0$) eloszlásfüggvényét! (Lásd előadás.)

Figyeljük meg, hogy az eloszlásfüggvény nem folytonos – így lesz ez mindig, amikor $\text{Ran}(X)$ megszámlálható. Viszont **balról folytonos**.

Az eloszlásfüggvény tulajdonságai

Tetszőleges $a < b$ valós számokra

$$F_X(b) - F_X(a) = \mathbb{P}(X < b) - \mathbb{P}(X < a) = \mathbb{P}(a \leq X < b),$$

a valószínűség additivitása miatt.

Az eloszlásfüggvények karakterizálhatók is:

(4.1.3. Állítás.) Egy $F: \mathbb{R} \rightarrow \mathbb{R}$ eloszlásfüggvénye valamilyen valószínűségi változónak $\Leftrightarrow$

- (1) F (nem feltétlenül szigorúan) monoton nő,
- (2) F balról folytonos, azaz minden $x \in \mathbb{R}$ -re $\lim_{y \uparrow x} F(y) = F(x)$,
- (3) és $\lim_{x \rightarrow -\infty} F(x) = 0$ és $\lim_{x \rightarrow \infty} F(x) = 1$.

- ▶ (Bizonyítás: $\Rightarrow$ irány)
- ▶ (Megjegyzés: konvenció az eloszlásfüggvény definíciójával kapcsolatban különböző országokban)

Eloszlásfüggvény

(4.1.2. Definíció.) Legyen X valószínűségi változó. Ekkor az $F_X: \mathbb{R} \rightarrow \mathbb{R}$,

$$F_X(x) \stackrel{\text{def}}{=} \mathbb{P}(X \leq x) \in [0, 1]$$

függvényt az X eloszlásfüggvényének nevezzük.

Ez minden valószínűségi változó esetén definiált.

Diszkrét esetben praktikusabb volt a $p_X: k \mapsto p_X(k) = \mathbb{P}(X = k)$ **súlyfüggvény** használata, mint az eloszlásfüggvényé.

Ugyanis ha pl. X értékkészlete véges, akkor az F_X eloszlásfüggvény az X értékkészletének k pontjaiban felfelé **ugrik** (ugrás mérete: $\mathbb{P}(X = k)$), különben pedig konstans. (Lásd a pénzérmés példát.) Ugyanez igaz „szelíd” megszámlálhatóan végtelen értékkészlet esetén, pl. a geometriai, Poisson- és Zipf-eloszlásnál.

Probléma: ha az eloszlásfüggvény *folytonos* (mint pl. az $[0, 1]$ intervallumon „találomra” választott pont esetén), akkor minden fix $x \in \mathbb{R}$ esetén $\mathbb{P}(X = x) = 0$. Így a súlyfüggvény triviális, más mennyiségre van szükségünk $\rightarrow$ **sűrűségfüggvény**.

Eloszlásfüggvény

(4.1.2. Definíció.) Legyen X valószínűségi változó. Ekkor az $F_X: \mathbb{R} \rightarrow \mathbb{R}$,

$$F_X(x) \stackrel{\text{def}}{=} \mathbb{P}(X \leq x) \in [0, 1]$$

függvényt az X eloszlásfüggvényének nevezzük.

Ez minden valószínűségi változó esetén definiált.

Diszkrét esetben praktikusabb volt a $p_X: k \mapsto p_X(k) = \mathbb{P}(X = k)$ **súlyfüggvény** használata, mint az eloszlásfüggvényé.

Probléma: ha az eloszlásfüggvény *folytonos* (mint pl. az $[0, 1]$ intervallumon „találomra” választott pont esetén), akkor minden fix $x \in \mathbb{R}$ esetén $\mathbb{P}(X = x) = 0$. Így a súlyfüggvény triviális, más mennyiségre van szükségünk $\rightarrow$ **sűrűségfüggvény**.

(Megjegyzés: lehetetlen esemény vs. 0 valószínűségű esemény, lásd előadás)

(Megjegyzés: nem megszámlálható unióra nem additív a valószínűség)

Példa (a sűrűségfüggvény bevezetéséhez)

Kérdés: Válasszunk véletlenszerűen a $[0, 1]$ intervallumon egy pontot, jelölje ezt X , a négyzetét pedig Y . Mennyi a valószínűsége, hogy Y az $1/4$, illetve a $3/4$ egy kis környezetébe esik?

Vagy általában: egy $0 < x < 1$ pont kis környezetébe, mondjuk $(x - \Delta x, x + \Delta x)$ -be, ahol $\Delta x > 0$ olyan kicsi, hogy $x - \Delta x > 0$ és $x + \Delta x < 1$?

Jelölje F_Y az Y val. változó eloszlásfüggvényét. Ekkor

$$\begin{aligned}\mathbb{P}(Y \in (x - \Delta x, x + \Delta x)) &= \mathbb{P}(x - \Delta x < Y < x + \Delta x) \\ &= \mathbb{P}(Y < x + \Delta x, Y \not\leq x - \Delta x) = \mathbb{P}(Y < x + \Delta x) - \mathbb{P}(Y \leq x - \Delta x) \\ &= \mathbb{P}(Y < x + \Delta x) - \mathbb{P}(Y < x - \Delta x) \quad (\text{folyt.}) \\ &= F_Y(x + \Delta x) - F_Y(x - \Delta x)\end{aligned}$$

Mindkét oldalt $2\Delta x$ -szel osztva, majd Δx -szel 0-hoz tartva az F_Y eloszlásfüggvény **deriváltját** kapjuk:

$$\frac{\mathbb{P}(Y \in (x - \Delta x, x + \Delta x))}{2\Delta x} = \frac{F_Y(x + \Delta x) - F_Y(x - \Delta x)}{2\Delta x} \xrightarrow{\Delta x \downarrow 0} F'_Y(x).$$

Kis Δx esetén tehát: $\mathbb{P}(Y \in (x - \Delta x, x + \Delta x)) \approx F'_Y(x) \cdot 2\Delta x$.

Hogyan határozható meg F_Y ? Tényleg differenciálható? $\rightarrow$ Később.

Példa (a sűrűségfüggvény bevezetéséhez)

Kérdés: Válasszunk véletlenszerűen a $[0, 1]$ intervallumon egy pontot, jelölje ezt X , a négyzetét pedig Y . Mennyi a valószínűsége, hogy Y az $1/4$, illetve a $3/4$ egy kis környezetébe esik?

Vagy általában: egy $0 < x < 1$ pont kis környezetébe, mondjuk $(x - \Delta x, x + \Delta x)$ -be, ahol $\Delta x > 0$ olyan kicsi, hogy $x - \Delta x > 0$ és $x + \Delta x < 1$?

Jelölje F_Y az Y val. változó eloszlásfüggvényét. Ekkor

$$\begin{aligned}\mathbb{P}(Y \in (x - \Delta x, x + \Delta x)) &= \mathbb{P}(x - \Delta x < Y < x + \Delta x) \\&= \mathbb{P}(Y < x + \Delta x, Y \not\leq x - \Delta x) = \mathbb{P}(Y < x + \Delta x) - \mathbb{P}(Y \leq x - \Delta x) \\&= \mathbb{P}(Y < x + \Delta x) - \mathbb{P}(Y < x - \Delta x) \quad (\text{folyt.}) \\&= F_Y(x + \Delta x) - F_Y(x - \Delta x)\end{aligned}$$

Mindkét oldalt $2\Delta x$ -szel osztva, majd Δx -szel 0-hoz tartva az F_Y eloszlásfüggvény **deriváltját** kapjuk:

$$\frac{\mathbb{P}(Y \in (x - \Delta x, x + \Delta x))}{2\Delta x} = \frac{F_Y(x + \Delta x) - F_Y(x - \Delta x)}{2\Delta x} \xrightarrow{\Delta x \downarrow 0} F'_Y(x).$$

Kis Δx esetén tehát: $\mathbb{P}(Y \in (x - \Delta x, x + \Delta x)) \approx F'_Y(x) \cdot 2\Delta x$.

Hogyan határozható meg F_Y ? Tényleg differenciálható? $\rightarrow$ Később.

Sűrűségfüggvény

Tehát „folytonos” valószínűségi változó esetén az **eloszlásfüggvény deriváltja** (vagyis egy olyan függvény, amelynek **primitív függvénye** az eloszlásfüggvény) megadja, hogy kb. milyen valószínűséggel van a val. változó egy adott pont egy kis környezetében.

Ezzel a motivációval vezetjük be a következő fogalmakat.

(4.2.1. Definíció.) Egy X valószínűségi változót **folytonosnak** nevezünk, ha létezik olyan nemnegatív, valós $f_X: \mathbb{R} \rightarrow \mathbb{R}$ függvény, amire az $\int_{-\infty}^{\infty} f_X(z) dz$ improprius Riemann-integrál véges, és minden $x \in \mathbb{R}$ esetén

$$F_X(x) = \int_{-\infty}^x f_X(z) dz,$$

ahol (F_X az X val. változó eloszlásfüggvénye és) az integrál improprius Riemann-integrál.

Az f_X függvényt az X **sűrűségfüggvényének** hívjuk.

Megjegyzés: a folytonos val. változó ezen definíciója az *abszolút folytonos* val. változó fogalmának egy speciális esete, ami a gyakorlatban fontos példákat mind tartalmazza → lásd mértékelméleten alapuló valószínűségi tárgyak.

Sűrűségfüggvény

Tehát „folytonos” valószínűségi változó esetén az **eloszlásfüggvény deriváltja** (vagyis egy olyan függvény, amelynek **primitív függvénye** az eloszlásfüggvény) megadja, hogy kb. milyen valószínűséggel van a val. változó egy adott pont egy kis környezetében.

Ezzel a motivációval vezetjük be a következő fogalmakat.

(4.2.1. Definíció.) Egy X valószínűségi változót **folytonosnak** nevezünk, ha létezik olyan nemnegatív, valós $f_X: \mathbb{R} \rightarrow \mathbb{R}$ függvény, amire az $\int_{-\infty}^{\infty} f_X(z) dz$ improprius Riemann-integrál véges, és minden $x \in \mathbb{R}$ esetén

$$F_X(x) = \int_{-\infty}^x f_X(z) dz,$$

ahol (F_X az X val. változó eloszlásfüggvénye és) az integrál improprius Riemann-integrál.

Az f_X függvényt az X **sűrűségfüggvényének** hívjuk.

Megjegyzés: a folytonos val. változó ezen definíciója az *abszolút folytonos* val. változó fogalmának egy speciális esete, ami a gyakorlatban fontos példákat mind tartalmazza $\rightarrow$ lásd mértékelméleten alapuló valószínűségi tárgyak.

Sűrűségfüggvény

A Newton–Leibniz-tétel alapján, mivel a sűrűségfüggvény integrálja az eloszlásfüggvény, az eloszlásfüggvény deriváltja a sűrűségfv.

Itt nem probléma, ha az F_X eloszlásfüggvény deriváltja véges sok pontban nem létezik – ami a gyakorlatban igen hasznos lesz:

(4.2.2. Állítás.) Ha F_X folytonos és véges sok pont kivételével mindenhol deriválható, akkor X folytonos val. változó, és az

$$f(x) = \begin{cases} F'_X(x), & \text{ha } F_X \text{ differenciálható } x\text{-ben,} \\ 0, & \text{különben} \end{cases}$$

függvény sűrűségfüggvénye X -nek.

A 0 helyett itt írhatnánk pl. 42-t is, az integrálás szempontjából mindegy, ha véges sok pontban bárhogy megváltoztatjuk a sűrűségfüggvényt (ha nemnegatív marad).

(Egyszerű állítás.) Ha X folytonos, akkor F_X folytonos függvény.

Ez azért van, mert F_X monoton növekvő, tehát ha nem folytonos, akkor van egy ugrása fölfelé. Ekkor sűrűségfüggvénye nem létezhet.

Tehát: **semmilyen diszkrét val. változó nem folytonos!**

Sűrűségfüggvény

A Newton–Leibniz-tétel alapján, mivel a sűrűségfüggvény integrálja az eloszlásfüggvény, az eloszlásfüggvény deriváltja a sűrűségfv.

Itt nem probléma, ha az F_X eloszlásfüggvény deriváltja véges sok pontban nem létezik – ami a gyakorlatban igen hasznos lesz:

(4.2.2. Állítás.) Ha F_X folytonos és véges sok pont kivételével mindenhol deriválható, akkor X folytonos val. változó, és az

$$f(x) = \begin{cases} F'_X(x), & \text{ha } F_X \text{ differenciálható } x\text{-ben,} \\ 0, & \text{különben} \end{cases}$$

függvény sűrűségfüggvénye X -nek.

(Egyszerű állítás.) Ha X folytonos, akkor F_X folytonos függvény.

Ez azért van, mert F_X monoton növekvő, tehát ha nem folytonos, akkor van egy ugrása fölfelé. Ekkor sűrűségfüggvénye nem létezhet.

Tehát: **semmilyen diszkrét val. változó nem folytonos!**

Az Egyszerű állítás megfordítása nem igaz, illetve nem minden folytonos val. változó eloszlásfüggvénye differenciálható véges sok pont kivételével mindenhol. Ellenpéldák

→ lásd BME-TTK-s mértékelméletes/valszámos tárgyak.

A sűrűségfüggvény tulajdonságai

(4.2.4. Állítás.) Legyen X folytonos valószínűségi változó. Ekkor minden $a < b$ valós szám esetén

$$\mathbb{P}(a < X < b) = \int_a^b f_X(x) dx.$$

(Bizonyítás: lásd előadás)

A bizonyítás során láttuk: folytonos X esetén minden $a \in \mathbb{R}$ -re

$$\mathbb{P}(X = a) = \int_a^a f_X(x) dx = 0.$$

Emiatt minden a, b -re

$$\mathbb{P}(a < X < b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X \leq b).$$

Diszkrét esetben ez nincs így!!!

A sűrűségfüggvény karakterizációja

(4.2.5. Állítás.) Egy nemnegatív $f: \mathbb{R} \rightarrow \mathbb{R}$ függvényhez akkor és csak akkor létezik X folytonos valószínűségi változó, aminek az f sűrűségfüggvénye, ha f Riemann-integrálható és

$$\int_{-\infty}^{\infty} f(x) dx = 1.$$

Bizonyítás:

- ▶ Könnyű irány: ha X folytonos val. változó, akkor f_X -re teljesül az egyenlet, lásd előadás.
- ▶ A másik irányról (ami a jegyzetben nem szerepel): ha $f \geq 0$ Riemann-integrálható és $\int_{-\infty}^{\infty} f(x) dx = 1$, akkor $x \in \mathbb{R}$ -re

$$F(x) := \int_{-\infty}^x f(z) dz.$$

Könnyen ellenőrizhető: ekkor F teljesíti az eloszlásfüggvény tulajdonságait (jó határérték $\pm\infty$ -ben, mon. növe, balról folyt.).

A 4.1.3. Állítás visszairánya miatt (biz. lásd jegyzet) létezik olyan X val. változó, amelynek eloszlásfüggvénye F , így f pedig sűrűségfüggvénye X -nek.

A sűrűségfüggvény szemléletes jelentése

Ahogy már utaltunk rá, ha X folytonos, akkor $x \in \mathbb{R}$ -re

$$f_X(x) = \lim_{\Delta x \downarrow 0} \frac{\mathbb{P}(X \in (x - \Delta x, x + \Delta x))}{2\Delta x}$$

feltéve, hogy F_X differenciálható x -ben, ami **a legtöbb példában** véges sok pont kivételével teljesül.

Vagyis a $f_X(x)$ azt adja meg, hogy x egy kis környezetébe milyen valószínűséggel esik X , osztva a kis környezet hosszával.

Intuíció: végezzük az X -nek megfelelő véletlen kísérletet sokszor egymás után függetlenül. Ekkor az, hogy rögzített kis $\Delta x > 0$ esetén egy adott x -re $(x - \Delta x, x + \Delta x)$ -be milyen gyakorisággal (*sűrűséggel*) esnek a véletlen eredmények, kb. $f_X(x)$ -szel lesz arányos.

Ilyen értelemben a sűrűségfüggvény a súlyfüggvény megfelelő analógja a folytonos esetre vonatkozóan.

Intuíció pontosítása $\rightarrow$ lásd matematikai statisztika, empirikus sűrűségfüggvény/sűrűséghisztogram témaköre.

Első példa: az egyenletes eloszlás

Az $[a, b]$ (vagy (a, b)) intervallumon $(-\infty < a < b < \infty)$ vett egyenletes eloszlás annak a val. változónak az eloszlása, amit eddig “az $[a, b]$ intervallumon találomra választott X pont” eloszlásának neveztünk

→ geometriai valószínűségi mező, a valószínűség a „területtel” (itt intervallumhosszal) arányos.

Ez tehát azt jelenti, hogy

- ▶ $\mathbb{P}(a < X < b) = 1$,
- ▶ $\mathbb{P}(x < X < y)$ az (x, y) intervallum hosszával arányos, vagyis $\mathbb{P}(x < X < y) = \text{const} \cdot (y - x)$, ha $a \leq x < y \leq b$.

Ebből tehát $\text{const} = \frac{1}{b-a}$. Továbbá ha $a < x < b$, akkor az is adódik, hogy X sűrűségfüggvénye

$$f_X(x) = \lim_{\Delta x \downarrow 0} \frac{\mathbb{P}(X \in (x - \Delta x, x + \Delta x))}{2\Delta x} = \lim_{\Delta x \downarrow 0} \frac{2\Delta x}{(b-a)2\Delta x} = \frac{1}{b-a},$$

$f_X(x) = 0$, ha $x \leq a$ vagy $x \geq b$. (Ábra: lásd előadás.)

Folytonos X esetén tehát f_X nem mindig folytonos, F_X igen.

Első példa: az egyenletes eloszlás

Az (a, b) intervallumon egyenletes eloszlású X sűrűségfüggvénye

$$f_X(x) = \lim_{\Delta x \downarrow 0} \frac{\mathbb{P}(X \in (x - \Delta x, x + \Delta x))}{2\Delta x} = \lim_{\Delta x \downarrow 0} \frac{2\Delta x}{(b-a)2\Delta x} = \frac{1}{b-a},$$

ha $a < x < b$ és $f_X(x) = 0$ különben. Integrálással adódik, hogy

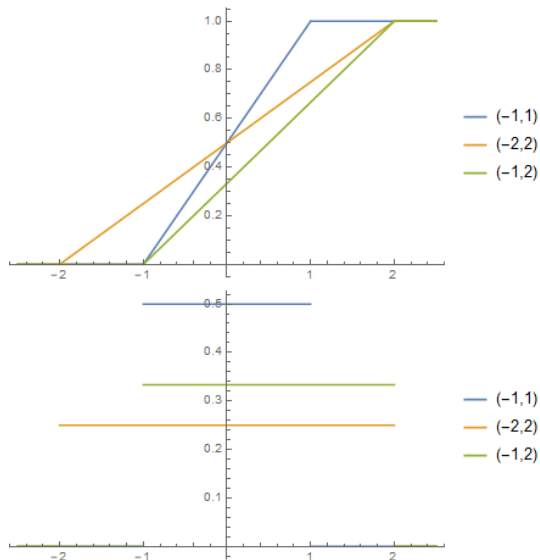
$$F_X(x) = \begin{cases} 0, & \text{ha } x \leq a, \\ \frac{x-a}{b-a}, & \text{ha } a < x < b, \\ 1, & \text{ha } x \geq b. \end{cases}$$

Jelölés: $X \sim U(a; b)$. Az $U(a; b)$ eloszlás esetén tehát a sűrűségfüggvény konstans ((a, b) -n, azon kívül pedig 0), az eloszlásfüggvény pedig lineáris ((a, b) -n, azon kívül konstans).

Megjegyzés: függvények megadása **a határokon folytonos esetben:**

- ▶ Sűrűségfüggvény: nyílt (vagyis $(-\infty, a)$, (a, b) , (b, ∞) , $(-\infty, \infty)$ alakú) intervallumon vagy ilyenek diszjunkt unióján legyen nem nulla, mert deriválni csak nyílt halmazon lehet.
- ▶ Eloszlásfüggvény: mivel folytonos, a határokat bármelyik szomszédos részbe besorolhatjuk, akár mind a kettőbe egyszerre (de legalább az egyikbe be kell sorolni).

Egyenletes eloszlás: eloszlás- és sűrűségfüggvény



Geometriai valószínűségi mezőn definiált valószínűségi változók (össznépi feladatmegoldó fesztivál)

Ha $X \sim U(a; b)$, akkor X transzformáltjai (pl. X^2, e^X stb.) már jellemzően nem egyenletes eloszlásúak, de eloszlásfüggvényük a geometriai valószínűség képlete alapján meghatározható, annak deriválásával pedig a sűrűségfüggvényük is.

Egyszerű példa: 5. feladatsor, 12. feladat.

Hasonlóan járunk el kétdimenziós geometriai valószínűségi mezőn definiált val. változó esetén is.

Kétdimenziós, de még viszonylag kellemesebb példa: 5. feladatsor, 13. (a) feladat ((b) esetleg gyakorlaton, vagy otthon önállóan).

Összetettebb példa: 5. feladatsor, 15. feladat.

Néha (akár geometriai val. mezőhöz kapcsolódó esetekben is) egyszerűbb a sűrűségfüggvényt meghatározni, mint az eloszlásfüggvényt. (Lásd 5. feladatsor, 16. feladat.)

5.12., 5.13. és 5.15. feladat

5.12. A $[0, 1]$ intervallumon taláломra kiválasztunk egy P pontot. Jelölje X a P pont és az intervallum P -hez közelebbi végpontja távolságának négyzetét.

(a) Legyen $\Omega = [0, 1]$ az eseményterünk. Írjuk fel X -et, mint $\Omega \rightarrow \mathbb{R}$ függvényt.

(b) Határozzuk meg X eloszlás- és sűrűségfüggvényét.

5.13. A síkon a $(0, 0)$, $(0, 1)$ és $(1, 1)$ csúcspontokkal rendelkező derékszögű, egyenlőszárú háromszög belsejében taláломra kiválasztunk egy (X, Y) pontot.

(a) Határozzuk meg X eloszlás- és sűrűségfüggvényét. (...)

5.15. Az egységnyi oldalú négyzet két átlellenes oldalán taláломra választunk egy a és egy b pontot. Jelöljük X -szel a két pont távolságának négyzetét.

(a) Határozzuk meg X eloszlásfüggvényét.

(b) Határozzuk meg X sűrűségfüggvényét.

(c) Hol a legnagyobb X sűrűségfüggvényének értéke?

Bertrand-paradoxon

A most következő diák részben Tóth Dávid németes valszám diasorából származnak.

Egy 1 sugarú körben válasszunk véletlenszerűen egy húrt. Mennyi a valószínűsége, hogy a húr legalább olyan hosszú, mint a körbe írható szabályos háromszög egy oldala?

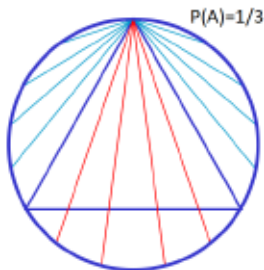
A következőkben megadunk három különböző „megoldást” (érvelést) három különböző eredménnyel, ezért nevezzük a problémát „paradoxonnak”.

Bertrand-paradoxon

1. megoldás: Válasszunk egyenletesen véletlenszerűen és egymástól függetlenül egy P és egy Q pontot a körvonalról, és legyen a húr az őket összekötő szakasz.

Sorsoljuk ki először a P pontot, majd rajzoljunk a körbe egy olyan szabályos háromszöget, aminek csúcsa ez a P .

Világos, hogy pontosan akkor kapunk a háromszög oldalánál hosszabb húrt, ha Q a háromszög másik két csúcsa közé esik. Így a keresett valószínűség $\frac{1}{3}$.

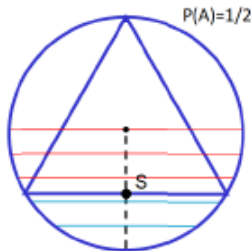


Bertrand-paradoxon

2. megoldás: Válasszuk ki egyenletesen véletlenszerűen egy sugarát a körnek, majd rajta egyenletesen véletlenszerűen (a sugár választástól függetlenül) egy P pontot. A húr legyen az, amelyik a sugárra merőleges, és a sugarat a P pontban metszi.

Sorsoljuk ki először a sugarat, majd rajzoljunk a körbe egy olyan szabályos háromszöget, aminek egyik oldala merőlegesen metszi a sugarat, jelölje a metszéspontot S .

Látható, hogy pontosan akkor kapunk a háromszög oldalánál hosszabb húrt, ha a P pont közelebb van a kör középpontjához, mint S . Mivel S felezi a sugarat, így a keresett valószínűség $\frac{1}{2}$.

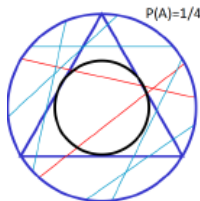


Bertrand-paradoxon

3. megoldás: Válasszunk ki egyenletesen véletlenszerűen egy P pontot a körlapon. Legyen a húr az, aminek éppen P a felezőpontja. Ilyen húr mindig csak egy van (hacsak P nem a középpont, de mivel ez nulla valószínűségű lehetőség, így ezzel az esettel nem kell foglalkoznunk).

Vegyünk egy szabályos háromszöget a körben, és nézzük ennek a beírható körét. A húr pontosan akkor lesz hosszabb a háromszög oldalánál, ha P a kisebb körbe esik.

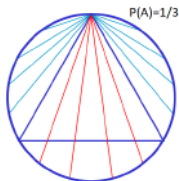
Az előző esetben láttuk, hogy ha a nagy kör sugara r , akkor a kis kör sugara $r/2$. A keresett valószínűség pedig a két kör területének aránya, azaz $\frac{(r/2)^2\pi}{r^2\pi} = \frac{1}{4}$.



Bertrand-paradoxon

Melyik megoldás helyes? Mindegyik.

Különböző modellek, különböző (mindegyik esetben geometriai) valószínűségi mezők $\rightarrow$ különböző eredmények.

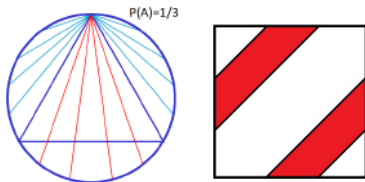


1. megoldás: két pontot választunk egymástól függetlenül, vagyis két szöget egyenletesen a $[0, 2\pi)$ intervallumon (a szögeket radiánban mérjük).

Bertrand-paradoxon

Melyik megoldás helyes? Mindegyik.

Különböző modellek, különböző (mindegyik esetben geometriai) valószínűségi mezők $\rightarrow$ különböző eredmények.



1. megoldás $\rightarrow$ Eseménytér: $\Omega = [0, 2\pi)^2 = [0, 2\pi) \times [0, 2\pi)$.

Ha α az elsőként választott szög, akkor a kedvező esetek halmaza $\{\alpha\} \times (\alpha + 2\pi/3, \alpha + 4\pi/3)$.

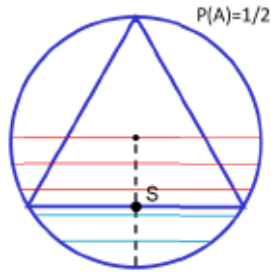
Ez azonban nem mindig esik bele teljesen az Ω négyzetbe, ezért $(\alpha + 2\pi/3, \alpha + 4\pi/3)$ -t „modulo 2π ” értjük, azaz az ide eső számok „ 2π -vel való osztási maradékát” vesszük (lásd ábra).

Így a kedvező terület $2\pi \cdot (2\pi/3) = 4\pi^2/3$ (két $4\pi/3$ befogójú derékszögű egyenlőszárú háromszög területe mínusz két $2\pi/3$ befogójúé), vagyis a keresett valószínűség $\frac{4\pi^2/3}{4\pi^2} = 1/3$.

Bertrand-paradoxon

A **2. megoldásnál** az eseménytér: $\Omega = [0, 2\pi) \times [0, 1]$ (a sugarat a végpontjának egy rögzített másik sugárral alkotott szögével, a sugáron választott pontot a középponttól való távolságával paraméterezzük, és a kettőt egymástól függetlenül választjuk). Kedvező kimenetek halmaza: $[0, 2\pi) \times [0, 1/2]$.

Így a keresett esemény valószínűsége $1/2$.

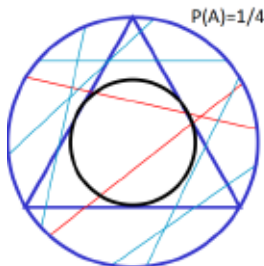


Bertrand-paradoxon

A **3. megoldásnál** az eseménytér: Ω az 1 sugarú, origó középpontú körlap.

A kedvező kimenetek halmaza az $1/2$ sugarú, ezzel koncentrikus kör.

Így a keresett esemény valószínűsége a két terület hányadosa, azaz $\frac{(1/2)^2\pi}{1^2\pi} = \frac{1}{4}$.



Néhány általános megfigyelés geometriai val. mezőknél

- ▶ Egy-egy pont választásának valószínűsége 0.
Hasonlóan: a síkban/térben egy egydimenziós szakasz vagy görbe választásának, illetve a térben egy kétdimenziós felület választásának valószínűsége is 0.
- ▶ Pl. az (a, b) intervallumon egyenletes eloszlású X esetén:
 $\mathbb{P}(X = x) = 0$, minden $x \in (a, b)$ -re. De $\{X = x\} \neq \emptyset$, hiszen $\bigcup_{x \in (a, b)} \{X = x\} = \{a < X < b\} = \Omega$.
0 valószínűségű, de nem lehetetlen események!
- ▶ A geometriai val. mező fogalma kiterjeszthető $\mathbb{R}^n$ bizonyos korlátos részhalmazaira is $\rightarrow$ „alacsonyabb dimenziós objektumok” választásának valószínűsége 0.
- ▶ A hosszúság/terület/térfogat/ n -dimenziós térfogat fogalma nem terjeszthető ki az eseménytér minden részhalmazára $\rightarrow$ ellenpélda a már említett Vitali-halmaz.
- ▶ Így nem lehet az eseménytér minden részhalmaza esemény, de a gyakorlatban fontos részhalmazok igen. (Pl. a síkban bármilyen sokszögbe, körbe/ellipszisbe stb. és ilyen halmazok megszámlálható metszetébe vagy uniójába esés esemény lesz.)

Példa: exponenciális eloszlás

(5.2.3. Definíció.) Egy Z valószínűségi változó **exponenciális eloszlású** $\lambda > 0$ paraméterrel, ha

$$f_Z(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{ha } x > 0, \\ 0, & \text{különben} \end{cases}$$

vagyis

$$F_Z(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ 1 - e^{-\lambda x}, & \text{ha } x > 0. \end{cases}$$

Jelölés: $X \sim \text{Exp}(\lambda)$.

Előfordulása: várakozási idők ($Z \geq 0$, 1 valószínűséggel), pl. egy villanykörte kiégéséig hátralévő idő.

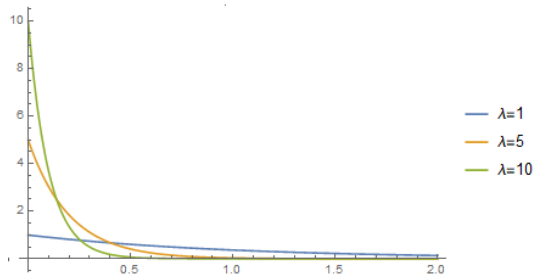
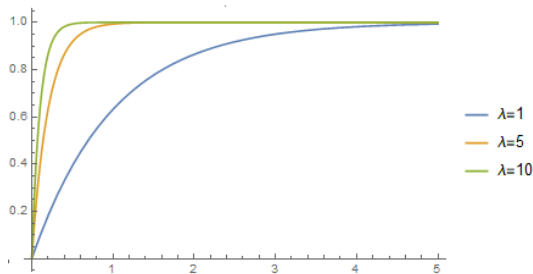
(Példa: össznépi gyakorlat, 6.1. feladat a), b))

Legyen X exponenciális eloszlású valószínűségi változó, amiről tudjuk, hogy $\mathbb{P}(X > 3) = e^{-6}$.

a) Mi X eloszlásának paramétere (λ)?

b) $\mathbb{P}(X < 2) = ?$

Exponenciális eloszlás: eloszlás- és sűrűségfüggvény



Exponenciális eloszlás: örökifjúság

Emlékeztető: a geometriai eloszlás örökifjúsága: ha $X \sim \text{Geo}(p)$, akkor

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s), \quad \forall s, t \in \{1, 2, \dots\}. \quad (1)$$

Figyeljük meg: ha s, t közül valamelyik nem egész szám, akkor (1) nem feltétlenül teljesül. Pl. legyen $t = 3/4$, $s = 1/2$, ekkor

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > 5/4 | X > 3/4) = \mathbb{P}(X \geq 2) = 1 - p,$$

de

$$\mathbb{P}(X > s) = \mathbb{P}(X > 1/2) = 1.$$

Viszont ha $X \sim \text{Exp}(\lambda)$, $\lambda > 0$, akkor (1) igaz bármely $s, t \in [0, \infty)$ esetén.

(Bizonyítás: lásd előadás)

Modellezési kérdés: milyen várakozási idők esetén teljesül (1)?

Pl. a villanykörte esetén mérési eredmények szerint *nagyjából* igen.

Pl. egy ember elhalálozásáig hátralévő idő esetén nem!

A villanykörte nem nagyon öregszik, az ember igen.

Exponenciális eloszlás: örökifjúság

Emlékeztető: a geometriai eloszlás örökifjúsága: ha $X \sim \text{Geo}(p)$, akkor

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s), \quad \forall s, t \in \{1, 2, \dots\}. \quad (1)$$

Figyeljük meg: ha s, t közül valamelyik nem egész szám, akkor (1) nem feltétlenül teljesül. Pl. legyen $t = 3/4$, $s = 1/2$, ekkor

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > 5/4 | X > 3/4) = \mathbb{P}(X \geq 2) = 1 - p,$$

de

$$\mathbb{P}(X > s) = \mathbb{P}(X > 1/2) = 1.$$

Viszont ha $X \sim \text{Exp}(\lambda)$, $\lambda > 0$, akkor (1) igaz bármely $s, t \in [0, \infty)$ esetén.

(Bizonyítás: lásd előadás)

Modellezési kérdés: milyen várakozási idők esetén teljesül (1)?

Pl. a villanykörte esetén mérési eredmények szerint *nagyjából* igen.

Pl. egy ember elhalálozásáig hátralévő idő esetén nem!

A villanykörte nem nagyon öregszik, az ember igen.

Örökijfúság, exponenciális és geometriai eloszlás

Emlékeztető: (5.2.5. (1) Állítás.) Ha egy X nemkonstans valószínűségi változóra

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s), \quad \forall s, t \in \{1, 2, \dots\} \quad (2)$$

teljesül és $\text{Ran}(X) = \{1, 2, \dots\}$, akkor $X \sim \text{Geo}(p)$, ahol $p = \mathbb{P}(X = 1)$.

Most: (5.2.5. (2) Állítás.) Ha egy X nemkonstans valószínűségi változóra a

$$\mathbb{P}(X > s + t | X > t) = \mathbb{P}(X > s), \quad \forall s, t \geq 0 \quad (3)$$

örökijfúsági feltétel teljesül és $\text{Ran}(X) = [0, \infty)$, akkor X eloszlása exponenciális.

Bizonyítás: lásd előadás.

A bizonyítás során használjuk az alábbi lemmát, amit nem bizonyítunk:

(5.2.6. Lemma.) Ha $g: [0, 1) \rightarrow \mathbb{R}$ monoton csökkenő függvény, amire $g(t + s) = g(t) + g(s)$ teljesül tetszőleges $s, t \in [0, \infty)$ esetén, akkor $g(t) = -\lambda t$, ahol $\lambda = -g(1) \geq 0$.

Kapcsolat az exponenciális és a geometriai eloszlás között

Az örökifjúság miatt az exponenciális eloszlás a geometriai eloszlás folytonos analógja. A következő állítás alapján a geometriai eloszlás ki is fejezhető az exponenciálisból.

(5.2.7. Állítás.) Ha az X valószínűségi változó exponenciális eloszlású λ paraméterrel, akkor $[X]$ geometriai eloszlású $1 - e^{-\lambda}$ paraméterrel. (Bizonyítás)

Figyeljük meg, hogy az exponenciális és a Poisson-eloszlás között szintén van kapcsolat: ha $X \sim \text{Exp}(\lambda)$ és $Y \sim \text{Pois}(\lambda)$, akkor

$$\mathbb{P}(X \geq 1) = e^{-\lambda} = \mathbb{P}(Y = 0).$$

A két eloszlást az ún. Poisson-folyamat kapcsolja össze $\rightarrow$ lásd MSc-s tárgyak: Sztochasztika és Tömegkiszolgálás (1).

Intuíció: tegyük fel, hogy bizonyos ismétlődő események (pl. telefonhívások, földrengések stb.) független, $\text{Exp}(\lambda)$ eloszlású időközönként követik egymást. Ekkor egy t hosszúságú időintervallumba eső események száma $\text{Pois}(\lambda t)$ eloszlású lesz, továbbá diszjunkt intervallumokba eső események számai függetlenek lesznek.

Ekkor $\mathbb{P}(X \geq 1) = \mathbb{P}(Y = 0)$ azt fejezi ki, hogy egy $t = 1$ hosszú időintervallumban nincs esemény.

Kapcsolat az exponenciális és a geometriai eloszlás között

Az örökifjúság miatt az exponenciális eloszlás a geometriai eloszlás folytonos analógja. A következő állítás alapján a geometriai eloszlás ki is fejezhető az exponenciálisból.

(5.2.7. Állítás.) Ha az X valószínűségi változó exponenciális eloszlású λ paraméterrel, akkor $\lceil X \rceil$ geometriai eloszlású $1 - e^{-\lambda}$ paraméterrel. (Bizonyítás)

Figyeljük meg, hogy az exponenciális és a **Poisson**-eloszlás között szintén van kapcsolat: ha $X \sim \text{Exp}(\lambda)$ és $Y \sim \text{Pois}(\lambda)$, akkor

$$\mathbb{P}(X \geq 1) = e^{-\lambda} = \mathbb{P}(Y = 0).$$

A két eloszlást az ún. **Poisson-folyamat** kapcsolja össze $\rightarrow$ lásd MSc-s tárgyak: Sztochasztika és Tömegkiszolgálás (1).

Intuíció: tegyük fel, hogy bizonyos ismétlődő események (pl. telefonhívások, földrengések stb.) független, $\text{Exp}(\lambda)$ eloszlású időközönként követik egymást. Ekkor egy t hosszúságú időintervallumba eső események száma $\text{Pois}(\lambda t)$ eloszlású lesz, továbbá diszjunkt intervallumokba eső események számai függetlenek lesznek.

Ekkor $\mathbb{P}(X \geq 1) = \mathbb{P}(Y = 0)$ azt fejezi ki, hogy egy $t = 1$ hosszú időintervallumban nincs esemény.

Nem minden val. változó diszkrét vagy folytonos...

Példa: az X val. változót úgy kapjuk, hogy egy szabályos pénzérmét feldobunk, és ha fej az eredmény, akkor a kedvenc véletlenszám-generátorunkkal $(0, 1)$ -en választunk egyenletesen egy számot és ezt tekintjük X -nek, ha pedig írás, akkor feldobunk egy szabályos dobókockát és az eredményt tekintjük X -nek. Ekkor

- ▶ X nem folytonos, mert ha az lenne, akkor $\mathbb{P}(X = 6) = 0$ teljesülne, de $\mathbb{P}(X = 6) = \frac{1}{2} \cdot \frac{1}{6} = \frac{1}{12}$ (szorzási szabály).
- ▶ X nem is diszkrét, mert a $(0, 1)$ intervallumon minden értéket felvehet, tehát értékkészlete nem megszámlálható.

Feltéve, hogy fejet dobunk, X úgy viselkedik, mint egy folytonos val. változó, és feltéve, hogy írást, úgy, mint egy diszkrét.

Ha $x \in (0, 1)$ és Δx olyan kicsi, hogy $0 < x - \Delta x < x + \Delta x < 1$, akkor

$$\mathbb{P}(X \in (x - \Delta x, x + \Delta x)) = \mathbb{P}(\text{fej})\mathbb{P}(X \in (x - \Delta x, x + \Delta x) | \text{fej}) = \Delta x.$$

Így a „sűrűségfüggvény” a $(0, 1)$ -en konstans $\frac{1}{2}$, azonban ez nem igazi sűrűségfüggvény, mert ha 0-tól 1-ig integráljuk, csak $\frac{1}{2}$ -et ($= \mathbb{P}(\text{fej})$) kapunk.

Várható érték folytonos esetben

Emlékeztető: ha X **nemnegatív** diszkrét valószínűségi változó, amelyre

$$\sum_{k \in \text{Ran}(X)} |k| \mathbb{P}(X = k) < \infty,$$

akkor

$$\mathbb{E}(X) = \sum_{k \in \text{Ran}(X)} k \mathbb{P}(X = k).$$

Ennek analógja a folytonos esetben: ahelyett, hogy a súlyfüggvényt k -val szorozva összegeznénk, a sűrűségfüggvényt x -szel szorozva kiintegráljuk $\rightarrow$ ez is egy „valószínűséggel súlyozott átlag”.

(4.3.3. Állítás.) Legyen X folytonos valószínűségi változó, amelyre

$$\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty.$$

Ekkor

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Várható érték folytonos esetben

(4.3.3. Állítás.) Legyen X folytonos valószínűségi változó, amelyre

$$\int_{-\infty}^{\infty} |x| f_X(x) dx < \infty.$$

Ekkor

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$

Az állítást – bizonyos mértékelméleti alapok hiányában – nem bizonyítjuk.

Megjegyzés: Ezen tárgy szempontjából a 4.3.3. Állításban szereplő képletet nyugodtan tekinthetjük a várható érték *definíciójának* folytonos esetben.

Azonban nem igaz, hogy diszkrét és folytonos esetben más lenne a várható érték definíciója. A jegyzetben szereplő 3.4.1. Definíció **minden** nemnegatív val. változó várható értékét definiálja, ami lefedi a diszkrét és a folytonos esetet is. Ez a definíció később következik.

(Össznépi gyakorlat: 6.8. feladat, az 5.15. eredménye alapján)

Nevezetes eloszlások várható értéke

(Állítás.) Ha $X \sim \text{Exp}(\lambda)$ valamely $\lambda > 0$ -ra, akkor $\mathbb{E}(X) = \frac{1}{\lambda}$.

(Bizonyítás: lásd előadás/jegyzet, parc. int.)

Vagyis minél nagyobb λ , annál kisebb a várható érték.

Intuíció: ha bizonyos események között a várakozási idők független $\text{Exp}(\lambda)$ -eloszlásúak, akkor egy 1 hosszú időintervallumba átlagosan λ db esemény fog esni $\rightarrow$ lásd Poisson-folyamat. Emiatt az exponenciális eloszlás λ paraméterét szokás az eloszlás **rátájának** is nevezni.

(Állítás.) Ha X egyenletes eloszlású az (a, b) intervallumon (ahol $-\infty < a < b < \infty$), akkor $\mathbb{E}(X) = \frac{a+b}{2}$.

(Bizonyítás: lásd előadás/jegyzet)

Vagyis az $U(a; b)$ eloszlású val. változó „átlagosan” az intervallum középpontjában van. (Szimmetria!)

Nevezetes eloszlások várható értéke

(Állítás.) Ha $X \sim \text{Exp}(\lambda)$ valamely $\lambda > 0$ -ra, akkor $\mathbb{E}(X) = \frac{1}{\lambda}$.

(Bizonyítás: lásd előadás/jegyzet, parc. int.)

Vagyis minél nagyobb λ , annál kisebb a várható érték.

Intuíció: ha bizonyos események között a várakozási idők független $\text{Exp}(\lambda)$ -eloszlásúak, akkor egy 1 hosszú időintervallumba átlagosan λ db esemény fog esni $\rightarrow$ lásd Poisson-folyamat. Emiatt az exponenciális eloszlás λ paraméterét szokás az eloszlás **rátájának** is nevezni.

(Állítás.) Ha X egyenletes eloszlású az (a, b) intervallumon (ahol $-\infty < a < b < \infty$), akkor $\mathbb{E}(X) = \frac{a+b}{2}$.

(Bizonyítás: lásd előadás/jegyzet)

Vagyis az $U(a; b)$ eloszlású val. változó „átlagosan” az intervallum középpontjában van. (Szimmetria!)

Várható érték: pozitív rész, negatív rész

Kérdés: ha $\int_{-\infty}^{\infty} |x| f_X(x) dx = \infty$, lehet-e $\mathbb{E}(X)$ -et legalább $+\infty$ vagy $-\infty$ értékkel definiálni?

Erre a kérdésre ugyanazt válaszolhatjuk, mint diszkrét esetben:

- ▶ Ha $\mathbb{P}(X \geq 0) = 1$: $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx \in [0, \infty]$ jóldefiniált.
- ▶ Ha $\mathbb{P}(X \leq 0) = 1$: $\mathbb{E}(X) \in [-\infty, 0]$ jóldefiniált.
- ▶ Ha X pozitív és negatív értékeket is felvesz: ekkor, ahogy láttuk, $X^+ = \max(X, 0)$ és $X^- = \max(-X, 0)$ nemnegatív valószínűségi változók, továbbá

$$X = X^+ - X^- \text{ és } |X| = X^+ + X^-.$$

Ha $\mathbb{E}(X^+) < \infty$ vagy $\mathbb{E}(X^-) < \infty$, akkor legyen

$$\mathbb{E}(X) \stackrel{\text{def.}}{=} \mathbb{E}(X^+) - \mathbb{E}(X^-),$$

ami vagy egy valós szám, vagy $+\infty$, vagy $-\infty$. Ha $\mathbb{E}(X^+) = \mathbb{E}(X^-) = \infty$, akkor a várható értéket nem értelmezzük. (Példa: Cauchy-eloszlás.)

- ▶ Ismét látjuk: $\mathbb{E}(|X|) < \infty$ pontosan akkor, ha $\mathbb{E}(X)$ létezik és véges.

A várható érték linearitása

A várható érték linearitása nemcsak diszkrét, hanem tetszőleges val. változók esetén igaz, amelyeknek létezik várható értéke:

(Állítás.) Legyenek X, Y valószínűségi változók, amelyekre $\mathbb{E}(|X|) < \infty$ és $\mathbb{E}(|Y|) < \infty$. Ekkor $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$ és $c \in \mathbb{R}$ esetén $\mathbb{E}(cX) = c\mathbb{E}(X)$.

Ugyanez igaz pl. nemnegatív val. változókra is, használva az $\infty + \infty = \infty + c = \infty$ konvenciót (ahol $c \in \mathbb{R}$):

(4.3.1. Állítás.) Legyenek X, Y nemnegatív valószínűségi változók. Ekkor $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$.

Ezek az állítások akkor is igazak pl., ha X diszkrét és Y folytonos. Pl. ha X egy kockadobás eredménye és $Y \sim \text{Exp}(1)$, akkor $\mathbb{E}(X + Y) = 4,5$.

Transzformált várható értéke

Folytonos X esetén is fontos mennyiség például X szórásnégyzete:

$$\mathbb{D}^2(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Kérdés: Hogyan számoljuk ki $\mathbb{E}(X^2)$ -et?

1. lehetőség: alkalmazzuk az alábbi állítást $g(x) = x^2$ -re.

(4.3.5. Állítás részlete.) Legyen X folytonos valószínűségi változó és $g: \mathbb{R} \rightarrow \mathbb{R}$ folytonos függvény. Ha

$$\int_{-\infty}^{\infty} |g(x)| \cdot f_X(x) \, dx < \infty,$$

akkor

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) \, dx.$$

Megjegyzés: Ha $g(x) \geq 0$ minden $x \in \text{Ran}(X)$ esetén, akkor $\mathbb{E}(g(X)) \in [0, \infty]$ mindig jóldefiniált. Pl. $\mathbb{E}(X^2)$ minden X -re jóldefiniált, legfeljebb végtelen.

Transzformált várható értéke

Folytonos X esetén is fontos mennyiség például X szórásnégyzete:

$$\mathbb{D}^2(X) = \mathbb{E}((X - \mathbb{E}(X))^2) = \mathbb{E}(X^2) - \mathbb{E}(X)^2.$$

Kérdés: Hogyan számoljuk ki $\mathbb{E}(X^2)$ -et?

1. lehetőség: alkalmazzuk az alábbi állítást $g(x) = x^2$ -re.

(4.3.5. Állítás részlete.) Legyen X folytonos valószínűségi változó és $g: \mathbb{R} \rightarrow \mathbb{R}$ folytonos függvény. Ha

$$\int_{-\infty}^{\infty} |g(x)| \cdot f_X(x) \, dx < \infty,$$

akkor

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) \cdot f_X(x) \, dx.$$

Megjegyzés: Ha $g(x) \geq 0$ minden $x \in \text{Ran}(X)$ esetén, akkor $\mathbb{E}(g(X)) \in [0, \infty]$ mindig jóldefiniált. Pl. $\mathbb{E}(X^2)$ minden X -re jóldefiniált, legfeljebb végtelen.

Transzformált eloszlása folytonos esetben

Kérdés: Hogyan számoljuk ki $\mathbb{E}(X^2)$ -et?

2. lehetőség: Meghatározzuk a $g(X)$ val. változó $f_{g(X)}$ sűrűségfüggvényét, majd a várható érték képlete szerint kiszámoljuk a várható értékét:

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} x \cdot f_{g(X)}(x) dx.$$

Ha csak $\mathbb{E}(g(X))$ érdekel minket, akkor sokszor könnyebb az előző dián szereplő képletet használni. Ha viszont kíváncsiak vagyunk a $g(X)$ val. változó eloszlására is, akkor mindenképpen meg kell határoznunk $f_{g(X)}$ -et. **Eljárás:**

1. Meghatározzuk $g(X)$ értékkészletét.
2. Definíció szerint kiszámoljuk az $F_{g(X)}$ eloszlásfüggvényt.
3. Ennek a deriváltja (...) az $f_{g(X)}$ sűrűségfüggvény.

Figyelem: $f_{g(X)}$ -et általában nem tudjuk közvetlenül meghatározni.

Adósság: egyenletes eloszlás transzformáltja (négyzete)

Példa: $Y = X^2$ eloszlása, ahol X egyenletes eloszlású $(0, 1)$ -en.

1. $\text{Ran}(Y) = (0, 1)$, mert $\{x^2 : x \in (0, 1)\} = (0, 1)$.

$$2. F_Y(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ \sqrt{x}, & \text{ha } 0 < x < 1, \\ 1, & \text{ha } x \geq 1. \end{cases}$$

(Számolás: lásd előadás)

$$3. \text{Így } f_Y(x) = \begin{cases} \frac{1}{2\sqrt{x}}, & \text{ha } 0 < x < 1, \\ 0, & \text{ha } x \leq 0 \text{ vagy } x \geq 1. \end{cases}$$

Vagyis a diasor elején feltett kérdésre a válasz:

$$\mathbb{P}(Y \in (3/4 - \Delta x, 3/4 + \Delta x)) \approx \frac{1}{\sqrt{3}} \cdot 2\Delta x$$

és

$$\mathbb{P}(Y \in (1/4 - \Delta x, 1/4 + \Delta x)) \approx 1 \cdot 2\Delta x,$$

ha Δx kicsi. Vegyük észre: Y sűrűségfüggvénye nem korlátos!

(További példa eloszlástrafóra: össznépi gyak., 6.13. feladat)

Szórásnégyzet tulajdonságai

Igaz a folytonos/általános esetben is: ha $c \in \mathbb{R}$ és $\mathbb{E}(X^2) < \infty$, akkor $\mathbb{D}^2(cX) = c^2\mathbb{D}^2(X)$, $\mathbb{D}^2(X + c) = \mathbb{D}^2(X)$.

Emlékeztető: (6.4.2. Állítás) Ha X és Y független valószínűségi változók és $\mathbb{E}(X^2), \mathbb{E}(Y^2) < \infty$, akkor

$$\mathbb{D}^2(X + Y) = \mathbb{D}^2(X) + \mathbb{D}^2(Y).$$

Ezt csak diszkrét esetben láttuk be (és ott is csak részben), de a folytonos esetben hasonló a bizonyítás.

A diszkrét bizonyítás kulcsa: ha X, Y függetlenek, akkor $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$. Ugyanez igaz *bármilyen* független valószínűségi változók esetén.

- ▶ Azt már tudjuk, hogy mi a függetlenség általános valószínűségi változók esetén: X, Y függetlenek $\Leftrightarrow \mathbb{P}(X < x, Y < y) = \mathbb{P}(X < x)\mathbb{P}(Y < y)$, $\forall x, y \in \mathbb{R}$.
- ▶ Nem tanultuk még, hogy $\mathbb{E}(XY)$ -t hogyan lehet kiszámolni a folytonos esetben (és hogy egyáltalán mit nevezünk „folytonos esetnek” többdimenziós eloszlásoknál) $\rightarrow$ hamarosan.

Nevezetes folytonos eloszlások szórása

(Példa.) Ha X egyenletes eloszlású az (a, b) intervallumon, akkor

$$\mathbb{D}^2(X) = \frac{(b-a)^2}{12}.$$

(Tehát $\mathbb{D}(X) = \frac{b-a}{2\sqrt{3}}$.)

(Bizonyítás – $\mathbb{E}(X^2)$ kiszámolása: lásd előadás)

(Példa: össznépi gyakorlat, 6.6. feladat, $\mathbb{D}^2(X+c) = \mathbb{D}^2(X)$)

(10.1.1. Példa.) Ha $Z \sim \text{Exp}(\lambda)$ (ahol $\lambda > 0$), akkor $\mathbb{D}^2(Z) = \frac{1}{\lambda^2}$ (tehát $\mathbb{D}(Z) = \frac{1}{\lambda}$).

(Bizonyítás: azt már tudjuk, hogy $\mathbb{E}(Z)^2 = \frac{1}{\lambda^2}$, ezért elég belátni, hogy $\mathbb{E}(Z^2) = \frac{2}{\lambda^2}$.)

Ennek kiszámítása $\mathbb{E}(Z)$ kiszámításához hasonló: $2 \times$ parc. int. ugyanazzal a szereposztással, illetve a második parc. int.

elkerülhető, ha felismerjük az $\text{Exp}(\lambda)$ eloszlás sűrűségfüggvényét. Lásd 6.14. feladat.)

Nemnegatív valószínűségi változók várható értéke

Nemnegatív val. változó esetén létezik egy elegáns alternatív kiszámítási módja a várható értéknek, ami esetenként könnyebb, mint az általános képlet.

(Állítás.) Ha X **nemnegatív** valószínűségi változó, akkor

$$\mathbb{E}(X) = \int_0^{\infty} \mathbb{P}(X \geq t) dt = \int_0^{\infty} (1 - F_X(t)) dt.$$

(Bizonyítás folytonos X esetén: nincs a jegyzetben, ezért itt)

$$\begin{aligned}\mathbb{E}(X) &= \int_0^{\infty} x f_X(x) dx = \int_0^{\infty} \int_0^x 1 dt f_X(x) dx \\ &= \int_0^{\infty} \int_t^{\infty} f_X(x) dx dt = \int_0^{\infty} [F_X(x)]_{x=t}^{\infty} dt \\ &= \int_0^{\infty} (\lim_{x \rightarrow \infty} F_X(x) - F_X(t)) dt = \int_0^{\infty} (1 - F_X(t)) dt.\end{aligned}$$

Az integrálfelcserélés azért lehetséges, mert az integrandus ≥ 0 (lásd Fubini-tétel). A többi lépés az eloszlás- és sűrűségfüggvény ismert tulajdonságaiból következik.

Nemnegatív valószínűségi változók várható értéke

Nemnegatív val. változó esetén létezik egy elegáns alternatív kiszámítási módja a várható értéknek, ami esetenként könnyebb, mint az általános képlet.

(Állítás.) Ha X **nemnegatív** valószínűségi változó, akkor

$$\mathbb{E}(X) = \int_0^{\infty} \mathbb{P}(X \geq t) dt = \int_0^{\infty} (1 - F_X(t)) dt.$$

- ▶ Az állítás nincs a jegyzetben, de a tárgy anyagának része (sőt már a régi tárgyénak is része volt).
- ▶ A bizonyításban szereplő integrálfelcserélést gyakran *Fubini flip*-ként említi az angol nyelvű irodalom.
- ▶ Diszkrét esetben is igaz az állítás (példa: gyakorlat, 6.7. feladat). Bizonyítás: integrálok helyett szummákkal.
- ▶ A geometriai eloszlás várható értékének a jegyzetben lévő („szummafelcserélős”) kiszámításánál már használtuk ezt a trükköt.
- ▶ (Példa: exponenciális eloszlás)

Folytonos valószínűségi változók szimulációja

Az eloszlástranszformációk egy informatikai alkalmazása:

Tegyük fel, hogy van egy véletlenszám-generátorunk, amely képes közelítőleg független, közelítőleg $U(0; 1)$ eloszlású (pszeudo)véletlen számokat generálni.

Az nem ezen tárgy anyaga, hogy a pszeudovéletlen számok generálása hogyan működik. Tekinthejtük black boxnak.

Kérdés: Hogyan szimuláljunk ennek segítségével (kb.) független, (kb.) egy másik, megadott folytonos eloszlást követő val. változókat? Pl. $\text{Exp}(\lambda)$ eloszlásúakat?

Válasz: Az $U(0; 1)$ eloszlású val. változókat helyettesítsük be a keresett eloszlás eloszlásfüggvényének inverzébe.

Ennek a jóságát a következőkben két lépésben igazoljuk.

Folytonos valószínűségi változók szimulációja

Az eloszlástranszformációk egy informatikai alkalmazása:

Tegyük fel, hogy van egy véletlenszám-generátorunk, amely képes közelítőleg független, közelítőleg $U(0; 1)$ eloszlású (pszeudo)véletlen számokat generálni.

Az nem ezen tárgy anyaga, hogy a pszeudovéletlen számok generálása hogyan működik. Tekinthetjük black boxnak.

Kérdés: Hogyan szimuláljunk ennek segítségével (kb.) független, (kb.) egy másik, megadott folytonos eloszlást követő val. változókat? Pl. $\text{Exp}(\lambda)$ eloszlásúakat?

Válasz: Az $U(0; 1)$ eloszlású val. változókat helyettesítsük be a keresett eloszlás eloszlásfüggvényének inverzébe.

Ennek a jóságát a következőkben két lépésben igazoljuk.

Folytonos valószínűségi változók szimulációja

(Állítás.) Ha X folytonos val. változó, akkor az $F_X(X)$ val. változó $U(0; 1)$ eloszlású.

(Bizonyítás abban az esetben, ha az f_X sűrűségfüggvény egy I nyílt $((-\infty, \infty), (-\infty, b), (a, \infty)$ vagy (a, b) alakú) intervallumon mindenhol pozitív, azon kívül pedig 0.

Ekkor F_X szigorúan monoton nő az I intervallumon, „ I bal végpontjában” határértéke 0, „jobb végpontjában” határértéke pedig 1. Így az $F_X^{-1}: (0, 1) \rightarrow I$ inverzfüggvény jóldefiniált, szigorúan monoton növekvő és I -n minden értéket felvesz.)

(Következmény.) Ha X folytonos val. változó és $U \sim U(0; 1)$, akkor $F_X^{-1}(U)$ eloszlásfüggvénye F_X .

(Bizonyítás az előbb tárgyalt esetben)

($F_X^{-1}(U)$ eloszlásfv.-e $F_X \Leftrightarrow F_X^{-1}(U)$ és X eloszlása azonos.)

Megjegyzés: Az Állítás és a Következmény igaz tetszőleges folytonos val. változóra is, de óvatosabbnak kell lenni annak meghatározásánál, hogy az F_X és az F_X^{-1} függvények milyen értelmezési tartományon szig. mon. növekvők.

Folytonos valószínűségi változók szimulációja: példa

Ha $X \sim \text{Exp}(\lambda)$, $\lambda > 0$, akkor $F_X(x) = 1 - e^{-\lambda x}$.

X sűrűségfüggvénye az $I = (0, \infty)$ nyílt intervallumon pozitív, azon kívül 0.

$\lim_{x \downarrow 0} F_X(x) = 0$, $\lim_{x \uparrow \infty} F_X(x) = 1$, F_X szigorúan monoton nő $(0, \infty)$ -en.

Ekkor $x > 0$ és $u \in (0, 1)$ esetén

$$\begin{aligned} x = F_X^{-1}(u) &\Leftrightarrow F_X(x) = u &\Leftrightarrow 1 - e^{-\lambda x} = u \\ &\Leftrightarrow -\lambda x = \ln(1 - u) &\Leftrightarrow x = \frac{-\ln(1 - u)}{\lambda}. \end{aligned}$$

Tehát ha $U \sim U(0; 1)$ egy számítógéppel nyert véletlen szám, akkor $\frac{-\ln(1-U)}{\lambda} \sim \text{Exp}(\lambda)$.

(Példa Excelben szimulált exponenciális eloszlású véletlen számokra: lásd előadás)

Diszkrét valószínűségi változók szimulációja

Diszkrét X val. változó szimulációja esetén az előbbi eljárás nem használható, mert F_X nem invertálható (hiszen szakaszonként konstans, a szakaszok között pedig ugrik).

Kérdés: Hogyan szimuláljuk X -et egy $U \sim U(0; 1)$ val. változó segítségével?

Válasz: Ha $\text{Ran}(X) = \{k_1, k_2, \dots\}$ (k_i páronként különbözőek), osszuk fel a $(0, 1)$ intervallumot $I_1, I_2, \dots$ diszjunkt intervallumok uniójára úgy, hogy I_i hossza megegyezzen $\mathbb{P}(X = k_i)$ -vel. Ezután szimuláljuk X -et úgy, hogy ha $U \in I_i$, akkor az $X = k_i$ értéket adjuk vissza.

Pl. dobókocka: $U \in (0, 1/6] \Rightarrow X = 1$, $U \in (1/6, 2/6] \Rightarrow X = 2$, $\dots$, $U \in (5/6, 1) \Rightarrow X = 6$. Ez tökéletes szimuláció (legalábbis annyira, amennyire U tökéletesen $U(0; 1)$ eloszlású).

Ha X értékkészlete végtelen (pl. geometriai, Poisson- vagy Zipf-eloszlás), csak véges sok lehetséges értéket tudunk így szimulálni.

Diszkrét valószínűségi változók szimulációja

Diszkrét X val. változó szimulációja esetén az előbbi eljárás nem használható, mert F_X nem invertálható (hiszen szakaszonként konstans, a szakaszok között pedig ugrik).

Kérdés: Hogyan szimuláljuk X -et egy $U \sim U(0; 1)$ val. változó segítségével?

Válasz: Ha $\text{Ran}(X) = \{k_1, k_2, \dots\}$ (k_i páronként különbözőek), osszuk fel a $(0, 1)$ intervallumot $I_1, I_2, \dots$ diszjunkt intervallumok uniójára úgy, hogy I_i hossza megegyezzen $\mathbb{P}(X = k_i)$ -vel. Ezután szimuláljuk X -et úgy, hogy ha $U \in I_i$, akkor az $X = k_i$ értéket adjuk vissza.

Pl. dobókocka: $U \in (0, 1/6] \Rightarrow X = 1$, $U \in (1/6, 2/6] \Rightarrow X = 2$, $\dots$, $U \in (5/6, 1) \Rightarrow X = 6$. Ez tökéletes szimuláció (legalábbis annyira, amennyire U tökéletesen $U(0; 1)$ eloszlású).

Ha X értékkészlete végtelen (pl. geometriai, Poisson- vagy Zipf-eloszlás), csak véges sok lehetséges értéket tudunk így szimulálni.

Eddig tart a zh anyaga.

- ▶ A Bertrand-paradoxon és a Randomized Quicksort még nem volt, később fogjuk őket pótolni és a zh-hoz még nem kellene.
- ▶ A kiegészítő anyagként megjelölt részek sem kellene (a vizsgához sem).
- ▶ A gyakorlat anyagából a 6. feladatsor az utolsó, ami benne lesz a zh-ban (a 6. feladatsorról még az is, ami esetleg csak a 7. héten hangzik el gyakorlaton).
- ▶ „Elméleti feladat” (definíció, tétel kimondása) csak a vizsgán lesz, de természetesen a zh-feladatokban is szükség van az előadáson elhangzott definíciókra, állításokra. Ezeket meg kell nevezni és a szükséges tulajdonságokra is hivatkozni kell.
Pl. „Az A , B és C események teljes eseményrendszert alkotnak, mert páronként kizáróak és uniójuk Ω (+ ennek indoklása), ezért a teljes valószínűség tétele szerint: ... ”
Természetesen lehet rövidebben is fogalmazni, de akkor is kell a teljes eseményrendszer, annak jelentése és a TVT.
- ▶ Számológépet lehet hozni (és érdemes is). Nevezetes eloszlás táblázatot mi adunk, sajátot használni nem lehet.

Várható érték általános esetben

Eddig: 1. Ha X diszkrét, akkor $\mathbb{E}(X) = \sum_{k \in \text{Ran}(X)} k\mathbb{P}(X = k)$,
2. ha X folytonos, akkor $\mathbb{E}(X) = \int_{-\infty}^{\infty} xf_X(x)dx$ (mindkét esetben feltéve, hogy $\mathbb{E}(|X|)$ létezik),
és 3. ha X nemnegatív, akkor $\mathbb{E}(X) = \int_0^{\infty} (1 - F_X(x))dx \in [0, \infty]$.
Kérdés: Mi köze ezeknek a várhatóérték-fogalmaknak egymáshoz?
Van-e a várható értéknek olyan definíciója, amely mindhárom esetet magába foglalja?

Várható érték általános esetben

- Eddig: 1. Ha X diszkrét, akkor $\mathbb{E}(X) = \sum_{k \in \text{Ran}(X)} k \mathbb{P}(X = k)$,
2. ha X folytonos, akkor $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx$ (mindkét esetben feltéve, hogy $\mathbb{E}(|X|)$ létezik),
és 3. ha X nemnegatív, akkor $\mathbb{E}(X) = \int_0^{\infty} (1 - F_X(x)) dx \in [0, \infty]$.

Kérdés: Mi köze ezeknek a várhatóérték-fogalmaknak egymáshoz?
Van-e a várható értéknek olyan definíciója, amely mindhárom esetet magába foglalja?

1. válasz: $\mathbb{E}(X) = \int_0^{\infty} (1 - F_X(x)) dx \in [0, \infty]$ bármilyen nemnegatív val. változóra igaz.

(Ezt már kimondtuk, bár csak folytonos esetben láttuk be.)

Ez alapján: akármilyen X val. változóra $X^+ = \max\{X, 0\}$ és $X^- = \max\{-X, 0\}$ nemnegatív val. változók, ezért $\mathbb{E}(X^+), \mathbb{E}(X^-)$ jóldefiniált.

Ha legalább az egyik véges, akkor

$\mathbb{E}(X) = \mathbb{E}(X^+) - \mathbb{E}(X^-) \in [-\infty, \infty]$ is jóldefiniált. Ha mindkettő végtelen, akkor $\mathbb{E}(X)$ -et nem értelmezzük.

Az általános esetben is igaz: mindkettő véges $\Leftrightarrow \mathbb{E}(|X|) < \infty$.

Várható érték általános esetben

Eddig: 1. Ha X diszkrét, akkor $\mathbb{E}(X) = \sum_{k \in \text{Ran}(X)} k \mathbb{P}(X = k)$,
2. ha X folytonos, akkor $\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_X(x) dx$ (mindkét esetben feltéve, hogy $\mathbb{E}(|X|)$ létezik),
és 3. ha X nemnegatív, akkor $\mathbb{E}(X) = \int_0^{\infty} (1 - F_X(x)) dx \in [0, \infty]$.

Kérdés: Mi köze ezeknek a várhatóérték-fogalmaknak egymáshoz?
Van-e a várható értéknek olyan definíciója, amely mindhárom esetet magába foglalja?

Másik válasz: szintén X^+ -szal és X^- -szal, de ezek várható értékét máshogy definiáljuk (amiről belátható, hogy az előző definícióval ekvivalens).

(3.4.1. Definíció.) Legyen X egy nemnegatív valószínűségi változó. Ekkor definíció szerint

$$\mathbb{E}(X) = \sup_{Z: Z \text{ egyszerű val. változó, } Z \leq X} \mathbb{E}(Z) \in [0, \infty].$$

Vagyis X várható értéke a nála (ω -nként) nem nagyobb egyszerű Z val. változók várható értékeinek „lehető legnagyobb értéke”, precízebben legkisebb felső korlátja, azaz szuprénuma.

Várható érték általános esetben

$$X \geq 0 \quad \Rightarrow \quad \mathbb{E}(X) = \sup_{Z: Z \text{ egyszerű val. változó, } Z \leq X} \mathbb{E}(Z) \in [0, \infty].$$

Egyszerű következmények (részben már találkoztunk velük):

- ▶ A várható érték monotonitása: Ha $\mathbb{P}(0 \leq X \leq Y) = 1$ (ahol az egyenlőtlenség ω -nként értendő), akkor $\mathbb{E}(X) \leq \mathbb{E}(Y)$.
(Itt $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ is lehet végtelen. Konvenció: $\infty > c$ minden $c \in \mathbb{R}$ -re és $\infty \geq \infty$).
- ▶ Ugyanez igaz annak feltevése nélkül is, hogy X és Y nemnegatívak, amennyiben a várható értékük definiált (az nem baj, ha $+\infty$ vagy $-\infty$).
- ▶ Speciális eset: pl. ha $c \in \mathbb{R}$ és $\mathbb{P}(X \geq c) = 1$, akkor $c \leq \mathbb{E}(X) \leq \infty$. Ha $\mathbb{P}(X \leq c) = 1$, akkor $-\infty \leq \mathbb{E}(X) \leq c$.
- ▶ Ha $a < b$ és $\mathbb{P}(a \leq X \leq b) = 1$, akkor $\mathbb{E}(X) \in [a, b]$.
- ▶ Ezeket a zh-/vizsgafeladatok megoldásánál érdemes ellenőrizni!

Markov-egyenlőtlenség

Előfordulhat, hogy egy val. változó pontos eloszlását nem ismerjük, de a momentumait (várható értékét, szórását stb.) becsülni tudjuk (pl. mintavétel alapján $\rightarrow$ lásd hamarosan a statisztika résznél).

A következőkben olyan egyenlőtlenségekkel ismerkedünk meg, amelyek egy val. változó momentumai alapján adnak becslést arra, hogy a val. változó milyen valószínűséggel vesz fel “extrém” (nagy/kicsi) értékeket.

(9.1.1. Állítás (Markov-egyenlőtlenség).) Legyen X nemnegatív értékű valószínűségi változó. Ekkor minden $a > 0$ esetén

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}.$$

(Bizonyítás)

Igaz diszkrét és folytonos (és általános) esetben is.

A Markov-egyenlőtlenséget néha $a > 0$ helyett csak $a > \mathbb{E}(X)$ -re mondják ki (mert ha $a \leq X$, akkor triviális egyenlőtlenséget kapunk).

A Markov-egyenlőtlenség általánosításai

(9.1.1. Állítás (Markov-egyenlőtlenség).) Legyen $X \geq 0$, ekkor

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}, \quad \forall a > 0.$$

Mit tehetünk, ha X negatív értékeket is felvesz, vagy ha a Markov-egyenlőtlenség nem ad túl jó becslést?

Transzformáljuk X -et valamilyen alkalmas nemnegatív valószínűségi változóvá. Az így kapott egyenlőtlenség, amit magyarul szokás Tóth Bálint nyomán *turbó Markov-egyenlőtlenség*nek is nevezni:

(Következmény.) Legyen X valószínűségi változó és $g: \underbrace{\text{Ran}(X)}_{\subseteq \mathbb{R}} \rightarrow [0, \infty)$ folytonos, **szigorúan monoton növekvő** függvény.

Ekkor minden $a > \mathbb{E}(g(X))$ esetén

$$\mathbb{P}(X \geq a) = \mathbb{P}(g(X) \geq g(a)) \leq \frac{\mathbb{E}(g(X))}{g(a)}.$$

A szigorú monoton növekvés fontos – különben az első egyenlőség nem feltétlenül teljesül!

A Markov-egyenlőtlenség általánosításai

(9.1.1. Állítás (Markov-egyenlőtlenség).) Legyen $X \geq 0$, ekkor

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}, \quad \forall a > 0.$$

Mit tehetünk, ha X negatív értékeket is felvesz, vagy ha a Markov-egyenlőtlenség nem ad túl jó becslést?

Transzformáljuk X -et valamilyen alkalmas nemnegatív valószínűségi változóvá. Az így kapott egyenlőtlenség, amit magyarul szokás Tóth Bálint nyomán *turbó Markov-egyenlőtlenség*nek is nevezni:

(Következmény.) Legyen X valószínűségi változó és $g: \underbrace{\text{Ran}(X)}_{\subseteq \mathbb{R}} \rightarrow [0, \infty)$ folytonos, **szigorúan monoton növekvő** függvény.

Ekkor minden $a > \mathbb{E}(g(X))$ esetén

$$\mathbb{P}(X \geq a) = \mathbb{P}(g(X) \geq g(a)) \leq \frac{\mathbb{E}(g(X))}{g(a)}.$$

A szigorú monoton növekvés fontos – különben az első egyenlőség nem feltétlenül teljesül!

Csebisev-egyenlőtlenség

(„Turbó Markov-egyenlőtlenség.”) Legyen X valószínűségi változó és $g: \text{Ran}(X) \rightarrow [0, \infty)$ folytonos, **szigorúan monoton növekvő** függvény. Ekkor minden $a > \mathbb{E}(g(X))$ esetén

$$\mathbb{P}(X \geq a) = \mathbb{P}(g(X) \geq g(a)) \leq \frac{\mathbb{E}(g(X))}{g(a)}.$$

(9.1.2. Következmény (Csebisev-egyenlőtlenség).) Legyen Y valószínűségi változó, ekkor minden $a > 0$ esetén

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq a) \leq \frac{\mathbb{D}^2(Y)}{a^2}.$$

Biz.: Ez nem más, mint a turbó Markov, az alábbi választással:

- ▶ $X = |Y - \mathbb{E}(Y)|$,
- ▶ $g(x) = x^2$ (ami $\text{Ran}(X)$ -en, azaz $[0, \infty)$ -en szig. mon. nő).

A turbó Markov és a szórásnégyzet definíciója miatt

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq a) = \mathbb{P}((Y - \mathbb{E}(Y))^2 \geq a^2) \leq \frac{\mathbb{E}((Y - \mathbb{E}(Y))^2)}{a^2} = \frac{\mathbb{D}^2(Y)}{a^2}.$$

Csebisev-egyenlőtlenség

(„Turbó Markov-egyenlőtlenség.”) Legyen X valószínűségi változó és $g: \text{Ran}(X) \rightarrow [0, \infty)$ folytonos, **szigorúan monoton növekvő** függvény. Ekkor minden $a > \mathbb{E}(g(X))$ esetén

$$\mathbb{P}(X \geq a) = \mathbb{P}(g(X) \geq g(a)) \leq \frac{\mathbb{E}(g(X))}{g(a)}.$$

(9.1.2. Következmény (Csebisev-egyenlőtlenség).) Legyen Y valószínűségi változó, ekkor minden $a > 0$ esetén

$$\mathbb{P}(|Y - \mathbb{E}(Y)| \geq a) \leq \frac{\mathbb{D}^2(Y)}{a^2}.$$

Biz.: Ez nem más, mint a turbó Markov, az alábbi választással:

- ▶ $X = |Y - \mathbb{E}(Y)|$,
- ▶ $g(x) = x^2$ (ami $\text{Ran}(X)$ -en, azaz $[0, \infty)$ -en szig. mon. nő).

A Csebisev-egyenlőtlenség bármilyen val. változóra jó (bár $\mathbb{E}(Y^2) = \infty$ esetén triviális becslést ad.)

Csebisev-egyenlőtlenség: példák

Egy adott adatbázis szerver átlagosan 50 lekérést fogad egy időegység alatt. A lekérések számának szórása a tapasztalatok szerint 5. Adjunk alsó becslést annak a valószínűségére, hogy 40-nél több, de 60-nál kevesebb lesz a lekérések száma egy időegység alatt. (7.14. feladat, lásd gyakorlat)

A feladat tipikus példája annak, amikor egy val. változó eloszlását nem ismerjük, csak a momentumairól (itt a szórásáról) tudunk valamit.

A Csebisev-egyenlőtlenség nem mindig ad nagyon jó becslést, sőt $\mathbb{E}(Y^2) < \infty$ esetén is adhat triviális becslést (lásd gyakorlat, 7.15. feladat).

Paraméteres Csernov-egyenlőtlenség

(„Turbó Markov-egyenlőtlenség.”) Legyen X valószínűségi változó és $g: \text{Ran}(X) \rightarrow [0, \infty)$ folytonos, **szigorúan monoton növekvő** függvény. Ekkor minden $a > \mathbb{E}(g(X))$ esetén

$$\mathbb{P}(X \geq a) = \mathbb{P}(g(X) \geq g(a)) \leq \frac{\mathbb{E}(g(X))}{g(a)}.$$

A $g(x) = x^2$ választáshoz képest még jobb becslést kaphatunk, ha a gyorsabban növekvő g -t választunk:

(9.1.4. Következmény (Paraméteres Csernov-egyenlőtlenség).)

Legyen X valószínűségi változó. Ekkor minden $a, t > 0$ esetén

$$\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(e^{tX})}{e^{ta}}.$$

Biz.: ez valóban a turbó Markov, ahol $X = X$, $g(x) = e^{tx}$.

(9.1.5. Példa: Legyen $X \sim \text{Pois}(5)$. Adjunk felső becslést

$\mathbb{P}(X \geq 10)$ -re. $\rightarrow$ A paraméteres Csernov-egyenlőtlenséggel, t -t optimalizálva jóval pontosabb felső korlátot kapunk, mint a Csebisev-egyenlőtlenséggel.)

(Mese a nagy eltérésekről... $\rightarrow$ előadás)

„A” nevezetes eloszlás

Tekintsük az alábbi véletlen mennyiségeket:

emberek lábmérete, születési súlya, vérnyomása, magassága (nők/férfiak külön), napi átlaghőmérséklet (az év adott napján), páratartalom, mérési hiba egy kísérletben, zaj.

Sok tényező apró, független tényező folytonos eredménye.

Elnevezések: normális eloszlás, Gauss-eloszlás, haranggörbe (a sűrűségfüggvénye).

Leggyakoribb és legnagyobb jelentőségű eloszlás a valószínűségelméletben és a statisztikában.

Ez a rész jelentős mértékben tartalmaz Pintér Márta tavalyelőtti diasorából átvett szövegeket.

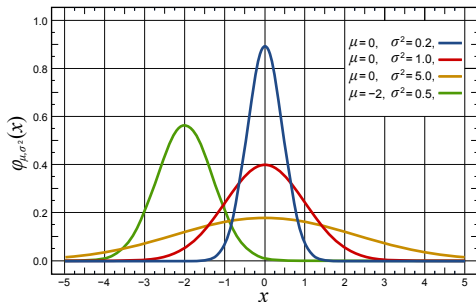
Normális eloszlás

(8.1.1. Definíció.) Egy Y folytonos valószínűségi változó **normális eloszlású** $\mu \in \mathbb{R}$ és $\sigma^2 > 0$ paraméterekkel, ha sűrűségfüggvénye

$$f_Y(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

Jelölés: $X \sim N(\mu; \sigma^2)$. $N(0; 1)$ neve standard normális.

(Megjegyzés: σ^2 négyzetgyökét σ -val jelöljük, de az eloszlás paramétere σ^2 , nem σ .)



A sűrűségfüggvény μ -re szimmetrikus, és minél nagyobb σ , annál szétterültebb.

Standard normális eloszlás: sűrűségfüggvény

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \quad x \in \mathbb{R}.$$

Ez nemnegatív függvény. A következő állításból következik, hogy valóban sűrűségfüggvény, vagyis $\int_{-\infty}^{\infty} f_Y(x) dx = 1$:

(8.1.2. Állítás.) $\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}} dx = \sqrt{2\pi}.$

(Bizonyítás: polárkoordinátákkal)

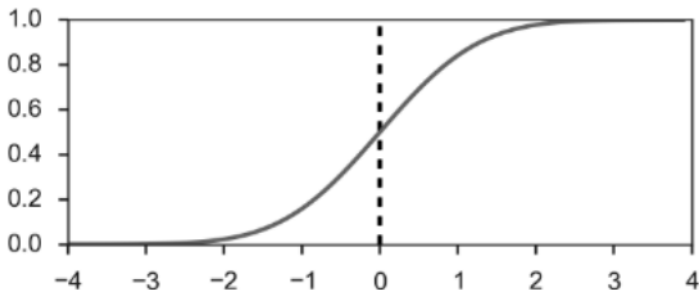
φ tulajdonságai: páros függvény, határértéke $\pm\infty$ -ben 0, mindenhol pozitív, maximuma 0-ban

Standard normális eloszlás: eloszlásfüggvény

$$\Phi(x) = \int_{-\infty}^x \varphi(t) dt = \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt.$$

Ez elemi függvényekkel nem kifejezhető.

Ehelyett közelítéseket használunk.



Φ tulajdonságai: eloszlásfüggvény tulajdonságai igazak rá,
szigorúan monoton nő, $\Phi(x) = 1 - \Phi(-x)$.

Standard normális eloszlás: táblázat

A standard normális eloszlás eloszlásfüggvényének táblázata

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9985	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998

x	0,00	0,01	0,02
0,0	0,5000	0,5040	0,5080
0,1	0,5398	0,5438	0,5478
0,2	0,5793	0,5832	0,5871
0,3	0,6179	0,6217	0,6255
0,4	0,6554	0,6591	0,6628
0,5	0,6915	0,6950	0,6985

$$\Phi(0,21) = 0,5832$$

$$\Phi(0,315) = 0,6236$$

$$\Phi(-0,11) = 1 - 0,5438 = 0,4562$$

Standard normális eloszlás: táblázat

„Gyakorlaton a korábbi években a következő eljárás alakult ki a táblázat használatához:

- a) Ha kettőnél több tizedesjegyre tudjuk x -et és $\Phi(x)$ a kérdés, akkor kerekítsünk a legközelebbi két tizedesjegyes x -re; és arra számoljunk $\Phi(x)$ -et, illetve ha pont félúton van az x , akkor átlagoljuk a szomszédos két tizedesjegyre kerekített számok Φ értékét.
 - b) Ha visszafelé kell kikeresniük egy értéket a normális eloszlásos táblázatból, azaz $\Phi(x) = y$, és y -t tudjuk, de az érték két táblázatban szereplő szám közé esik, azaz $\Phi(x_1) < y < \Phi(x_2)$ ahol x_1, x_2 szomszédosak a táblázatban, akkor a közelebbi x -et vegyük inverz értéknek, illetve ha pont félúton van az y , akkor pedig a kisebbik x -et.”
- (Pintér Márta)

Nem-standard normális

A nem-standard normális eloszlásra megadott függvények is sűrűségfüggvények.

Ugyanis a nem-standard normális eloszlás lineáris transzformációval nyerhető a standardból:

Legyen $\mu, \sigma \in \mathbb{R}$, $\sigma > 0$, X folytonos valószínűségi változó, amelynek eloszlásfüggvénye minden pontban differenciálható, és legyen $Y = \sigma X + \mu$. Ekkor

$$f_Y(x) = \frac{1}{\sigma} f_X\left(\frac{x - \mu}{\sigma}\right).$$

(Bizonyítás: standard eloszlástrafó-feladat, lásd előadás.

Minden folytonos val. változóra igaz a zöld feltétel nélkül is, a közvetett függvény deriválása helyett integrálhelyettesítéssel.

Lásd 8.2.1. Lemma, bizonyítás a jegyzetben.)

Ha $X \sim N(0; 1)$, akkor $Y = \sigma X + \mu \sim N(\mu; \sigma^2)$, mert

$$\frac{1}{\sigma} f_X\left(\frac{x - \mu}{\sigma}\right) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\frac{x-\mu}{\sigma})^2}{2}} = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}.$$

Standardizálás

Visszafelé: legyen $Y \sim N(\mu; \sigma^2)$ és legyen $X = \frac{Y-\mu}{\sigma}$. Ez standard normális eloszlású, ugyanis a korábbiak szerint

$$f_X(x) = \sigma f_Y(\sigma x + \mu) = \sigma \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(\sigma x + \mu - \mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Ezzel beláttuk:

(8.2.2. Következmény.) Legyen $\mu \in \mathbb{R}$ és $\sigma > 0$. Egy Y valószínűségi változó pontosan akkor $N(\mu; \sigma^2)$ eloszlású, ha létezik $X \sim N(0; 1)$, amire $Y = \sigma X + \mu$, azaz $X = \frac{Y-\mu}{\sigma}$.

Normális eloszlás várható értéke, szórásnégyzete

(8.2.3. Állítás.) Legyen $Y \sim N(\mu; \sigma^2)$. Ekkor $\mathbb{E}(Y) = \mu$ és $\mathbb{D}^2(Y) = \sigma^2$.

A bizonyításhoz először megmutatjuk, hogy elég a standard normális eloszlásra ($\mu = 0$, $\sigma = 1$) belátni az állítást.

Ugyanis tegyük fel, hogy ezt beláttuk, és legyen $Y \sim N(\mu; \sigma^2)$. Ekkor van olyan $X \sim N(0; 1)$, amelyre $Y = \sigma X + \mu$, tehát

$$\mathbb{E}(Y) = \mathbb{E}(\sigma X + \mu) = \sigma \mathbb{E}(X) + \mu = \mu$$

és

$$\mathbb{D}^2(Y) = \mathbb{D}^2(\sigma X + \mu) = \sigma^2 \mathbb{D}^2(X) = \sigma^2.$$

(Bizonyítás $X \sim N(0; 1)$ esetén: $\mathbb{E}(X) = 0$, $\mathbb{E}(X^2) = 1$ (parc. int.), lásd előadás/jegyzet)

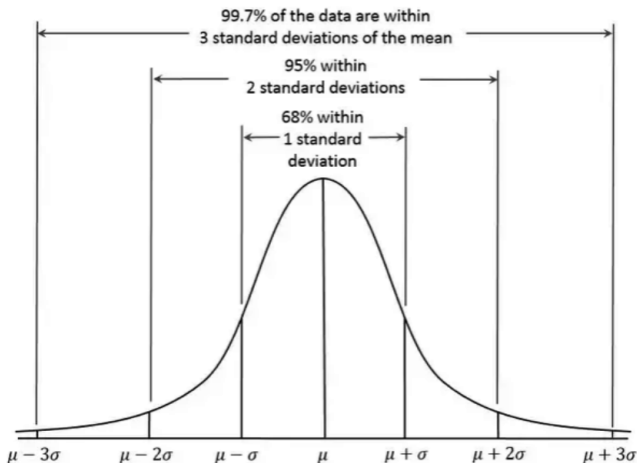
Standardizálás

Normális eloszlás standardizálása: ha $Y \sim N(\mu; \sigma^2)$, akkor

$$\frac{Y - \mathbb{E}(Y)}{\mathbb{D}(Y)} \sim N(0; 1).$$

Megjegyzés: Más eloszlású val. változót is standardizálhatunk (levonva a várható értékét, és leosztva a szórásával), feltéve, hogy a 2. momentuma véges. Ez a standardizált persze tipikusan nem lesz standard normális eloszlású, de várható értéke 0 lesz, szórása 1.

Normális eloszlás szórása szemléletesen



- ▶ Az esetek kb. 68%-át lefedi a $[\mu - \sigma, \mu + \sigma]$ intervallum,
- ▶ f_Y a $\mu \pm \sigma$ pontoknál vált konvexitást.

Normális eloszlás

Példa: Egy zajos csatornán egy ± 1 értékű jelet próbálunk átjuttatni. A zaj miatt a megkapott jel értéke nem feltétlenül ± 1 , hanem a jel értéke plusz egy $N(0; \sigma^2)$ eloszlású véletlen szám. Határozzuk meg σ -t, ha tudjuk, hogy annak az esélye, hogy egy “+1” jel esetén negatív érték jut át a csatornán, 0,0226.

Zaj: $Y \sim N(0; \sigma^2)$.

$$\mathbb{P}(1 + Y < 0) = \mathbb{P}(Y < -1) = \mathbb{P}\left(\frac{Y}{\sigma} < \frac{-1}{\sigma}\right) = \Phi\left(\frac{-1}{\sigma}\right) = 0,0226,$$

tehát $\sigma = \frac{1}{2,00}$.

Normális eloszlás

Példa: A minta hőmérséklete ($^{\circ}\text{C}$) $Y \sim N(-2; 1,69)$.

Mi a valószínűsége, hogy a minta hőmérséklete nagyobb, mint 0°C ?

Standardizált: $X = \frac{Y+2}{\sqrt{1.69}} \sim N(0; 1)$.

$$\mathbb{P}(Y > 0) = \mathbb{P}\left(\frac{Y+2}{1.3} > \frac{2}{1.3}\right) = 1 - \Phi\left(\frac{2}{1.3}\right).$$

$$\mathbb{P}(Y > 0) \approx 0,0620 \approx 6\%$$

Példa: normális eloszlások exponenciális momentumai

A normális eloszlás esetén sem mindig mindent a táblázatból olvasunk ki, hanem esetenként közvetlenül a sűrűségfüggvénnyel dolgozunk.

(Példa: exponenciális momentumok – össznépi gyakorlat, 7.6. feladat)

A példa üzenete – a konkrét számoláson túl – : bizonyos integrálokat azért könnyű kiszámolni, mert normális eloszlások sűrűségfüggvényére/várható értékére stb. vezethetők vissza. A valszámos megfontolások olykor az analízisben is segíthetnek.

(Másik példa: 7.5. feladat, lásd gyak.)

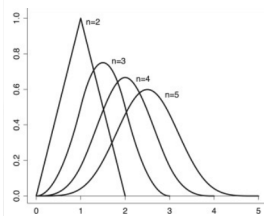
Kitérő: konvolúció

Ha X és Y független val. változók, akkor $X + Y$ eloszlását X és Y **konvolúciójának** hívjuk. Folytonos eset: $X + Y$ sűrűségfüggvénye

$$f_{X+Y}(x) = \int_{-\infty}^{\infty} f_X(z)f_Y(x-z)dz, \quad x \in \mathbb{R}.$$

- ▶ A binomiális eloszlás független indikátorok összege.
- ▶ $X \sim \text{Pois}(\lambda)$, $Y \sim \text{Pois}(\mu) \Rightarrow X + Y \sim \text{Pois}(\lambda + \mu)$.
- ▶ $\text{Geo}(p)$ eloszlásúak összege “negatív binomiális”.
- ▶ Egyenletes eloszlások összege: Irwin-Hall eloszlás (lásd ábra).
- ▶ Exponenciálisok összege “gamma-eloszlás”.
- ▶ Ha $X \sim N(\mu_1; \sigma_1^2)$ és $Y \sim N(\mu_2; \sigma_2^2)$, akkor $X + Y \sim N(\mu_1 + \mu_2; \sigma_1^2 + \sigma_2^2)$.

(Ebben csak az a meglepő, hogy az összeg is normális eloszlású. A paraméterek az $\mathbb{E}$ és $\mathbb{D}^2$ azonosságából adódnak.)



de Moivre–Laplace-tétel

Kérdés: Miért jön elő a normális eloszlás alkalmazásokban, mérési eredményeknél?

A következő tétel szerint a binomiális eloszlás standardizáltja rögzített p és nagy n esetén közelíthető a normális eloszlással:

(8.3.1. Tétel: de Moivre–Laplace-tétel.) Legyen $p \in (0, 1)$ és $S_n \sim B(n; p)$. Ekkor minden $a < b$ valós számra

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(a < \frac{S_n - \mathbb{E}(S_n)}{\mathbb{D}(S_n)} < b\right) = \int_a^b \varphi(x) dx = \Phi(b) - \Phi(a).$$

Ahogy láttuk: $\frac{S_n - \mathbb{E}(S_n)}{\mathbb{D}(S_n)} = \frac{S_n - np}{\sqrt{np(1-p)}}.$

Megjegyzés: A konvergencia sebességéről is lehet tudni konkrétumot.

de Moivre–Laplace-tétel

(8.3.1. Tétel: de Moivre–Laplace-tétel.) Legyen $p \in (0, 1)$ és $S_n \sim B(n; p)$. Ekkor minden $a < b$ valós számra

$$\lim_{n \rightarrow \infty} \mathbb{P}\left(a < \frac{S_n - \mathbb{E}(S_n)}{\mathbb{D}(S_n)} < b\right) = \int_a^b \varphi(x) dx = \Phi(b) - \Phi(a).$$

Poisson-eloszlásnál láttuk, hogy a binomiális határeloszlása a Poisson.

Most: binomiális határeloszlása a normális.

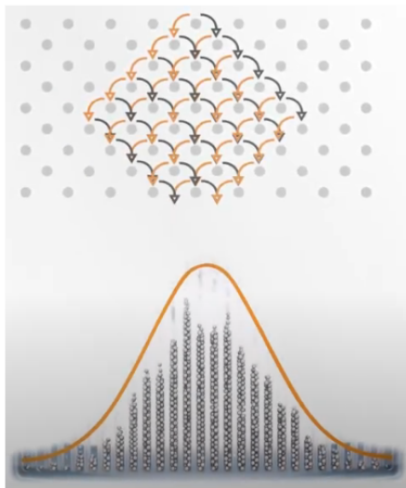
- ▶ Ott nem vontunk le (n -től függő) várható értéket, itt igen.
- ▶ Ott nem osztottunk le (n -től függő) szórással, itt igen.
- ▶ Ott $p_n \rightarrow 0$ és $np_n \rightarrow \lambda$ (pl. $p_n = \lambda/n$) volt feltétel, itt p konstans.

Tehát: n a Poisson-approximációnál is nagy, de ott $p = p_n$ nullához tart, a normális approximációnál/CHT-nál nem.

Galton-deszka

Egy golyó vég-pozíciójának eloszlása: $B(n; 1/2)$ (a bal széléről számolva).

A binomiális eloszlás közelíti a normális eloszlás sűrűségfüggvényét.



Példa: de Moivre–Laplace-tétel

(8.3.2. Példa.) A matematikusok 31,4% százaléka szandált hord. Száz találomra választott matematikust nézve, közelítőleg mi az esélye, hogy kevesebb, mint 25 pár szandált találunk rajtuk?

Szandálpárok száma: $S_n \sim B(n; p)$, $n = 100$, $p = 0.314$.

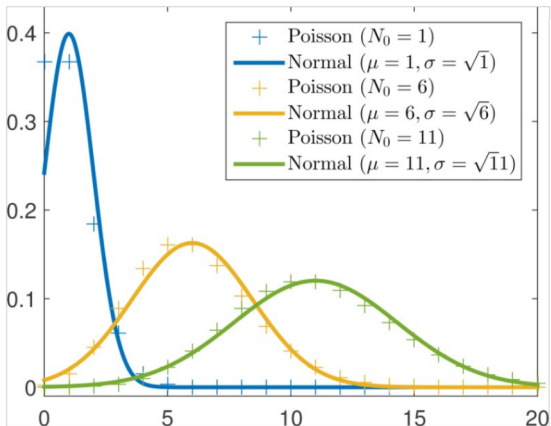
$$\mathbb{E}(S_n) = 31,4, \mathbb{D}(S_n) = \sqrt{100 \cdot 31,4 \cdot 68,6} \approx 4,6412.$$

$\frac{S_n - 31,4}{4,6412}$ közelítőleg $N(0; 1)$ eloszlású.

$$\mathbb{P}(X < 25) = \mathbb{P}\left(\frac{X - 31,4}{4,6412} < \frac{25 - 31,4}{4,6412}\right) \approx \Phi(-1,3790) = 1 - \Phi(1,3790) \approx 0,0839 \approx 8\%.$$

Poisson vs. normális

Nemcsak a binomiális, hanem pl. a Poisson-eloszlás standardizáltja is közelíti a normális eloszlást, ha a Poisson-kísérletet sokszor ismételjük.



Mintaátlag viselkedése

Általában: legyenek $X_1, X_2, \dots, X_n$ független, **azonos eloszlású** val. változók, amelyekre $\mathbb{E}(X_i^2) < \infty$.

(Azonos eloszlásúak = mindnek ugyanaz az eloszlásfüggvénye.)

Rövidítés: fae.=független és azonos eloszlású.

Angolul i.i.d.=independent and identically distributed.

Legyen $S_n = \sum_{i=1}^n X_i$. Ekkor:

$$\mathbb{E}(S_n) = \mathbb{E}(\sum_{i=1}^n X_i) \stackrel{\mathbb{E} \text{ lineáris}}{=} \sum_{i=1}^n \mathbb{E}(X_i) \stackrel{\text{a.e.}}{=} n\mathbb{E}(X_1).$$

$$\mathbb{D}^2(S_n) = \mathbb{D}^2(\sum_{i=1}^n X_i) \stackrel{\text{fgtl.}}{=} \sum_{i=1}^n \mathbb{D}^2(X_i) \stackrel{\text{a.e.}}{=} n\mathbb{D}^2(X_1).$$

Tekintsük az $\overline{X}_n = \frac{S_n}{n}$ **mintaátlagot**. Látjuk: $\mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1)$,

$$\mathbb{D}(\overline{X}_n) = \mathbb{D}(S_n/n) = \frac{1}{n}\mathbb{D}(S_n) = \frac{\mathbb{D}(X_1)}{\sqrt{n}} \longrightarrow 0.$$

Intuíció (pontatlan!): nagy n -re

- ▶ S_n/n közel van konstans $\mathbb{E}(X_1)$ -hez $\rightarrow$ **nagy számok törvénye**,
- ▶ az $\frac{S_n - n\mathbb{E}(X_1)}{\sqrt{n\mathbb{D}(X_1)}}$ standardizált kb. $N(0; 1)$ eloszlású $\rightarrow$ **centrális határeloszlás-tétel**.

Ezek formalizálásához: konvergenciafajták.

Mintaátlag viselkedése

Általában: legyenek $X_1, X_2, \dots, X_n$ független, **azonos eloszlású** val. változók, amelyekre $\mathbb{E}(X_i^2) < \infty$.

Legyen $S_n = \sum_{i=1}^n X_i$. Ekkor:

$$\mathbb{E}(S_n) = \mathbb{E}(\sum_{i=1}^n X_i) \stackrel{\text{E lineáris}}{=} \sum_{i=1}^n \mathbb{E}(X_i) \stackrel{\text{a.e.}}{=} n\mathbb{E}(X_1).$$

$$\mathbb{D}^2(S_n) = \mathbb{D}^2(\sum_{i=1}^n X_i) \stackrel{\text{fgtl.}}{=} \sum_{i=1}^n \mathbb{D}^2(X_i) \stackrel{\text{a.e.}}{=} n\mathbb{D}^2(X_1).$$

Tekintsük az $\overline{X}_n = \frac{S_n}{n}$ **mintaátlagot**. Látjuk: $\mathbb{E}(\overline{X}_n) = \mathbb{E}(X_1)$,

$$\mathbb{D}(\overline{X}_n) = \mathbb{D}(S_n/n) = \frac{1}{n}\mathbb{D}(S_n) = \frac{\mathbb{D}(X_1)}{\sqrt{n}} \longrightarrow 0.$$

Intuíció (pontatlan!): nagy n -re

- ▶ S_n/n közel van konstans $\mathbb{E}(X_1)$ -hez $\rightarrow$ **nagy számok törvénye**,
- ▶ az $\frac{S_n - n\mathbb{E}(X_1)}{\sqrt{n}\mathbb{D}(X_1)}$ standardizált kb. $N(0; 1)$ eloszlású $\rightarrow$ **centrális határeloszlás-tétel**.

Ezek formalizálásához: konvergenciafajták.

Eloszlásbeli konvergencia

(9.3.1. Definíció.) Legyen $X_1, X_2, \dots$ val. változók egy végtelen sorozata; X_n eloszlásfüggvénye: F_{X_n} .

A fenti sorozat eloszlásban konvergál egy F_Z eloszlásfüggvényű Z val. változóhoz, ha $n \rightarrow \infty$ esetén

$$F_{X_n}(x) \rightarrow F_Z(x)$$

minden olyan $x \in \mathbb{R}$ esetén, ahol F_Z folytonos.

Jelölés: $X_n \xrightarrow{d} Z$.

Tehát a de Moivre–Laplace-tétel eloszlásbeli konvergenciát állít.

Történelmileg az első központi határeloszlás típusú tétel.

Speciális esete a hamarosan következő általánosabb centrális határeloszlás- tételnek (CHT).

Előtte még nézzünk pár példát az eloszlásbeli konvergenciára, hogy egy kicsit feldolgozzuk a definíciót.

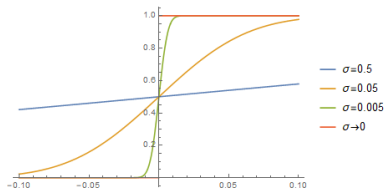
Eloszlásbeli konvergencia, 1. példa: normális eloszlás nullához tartó szórással

Legyen $n \in \mathbb{N}$ esetén X_n normális eloszlású 0 várható értékkel és $1/n$ szórásnégyzettel.

Láttuk: ahogy a szórás 0-hoz tart (n pedig végtelenhez), a sűrűségfüggvény egyre keskenyebb és csúcsosabb.

Az eloszlásfüggvényekre pedig a következő teljesül:

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n < x) = 0, \forall x < 0, \quad \lim_{n \rightarrow \infty} F_{X_n}(x) = 1, \forall x > 0.$$



Tehát $x = 0$ kivételével $F_{X_n}(x)$ a következőhöz tart:

$$F(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ 1, & \text{ha } x > 0. \end{cases}$$

Eloszlásbeli konvergencia, 1. példa: normális eloszlás nullához tartó szórással

Legyen $n \in \mathbb{N}$ esetén X_n normális eloszlású 0 várható értékkel és $1/n$ szórásnégyzettel.

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n < x) = 0, \forall x < 0, \quad \lim_{n \rightarrow \infty} F_{X_n}(x) = 1, \forall x > 0.$$

Tehát $x = 0$ kivételével $F_{X_n}(x)$ a következőhöz tart:

$$F(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ 1, & \text{ha } x > 0. \end{cases}$$

Ez az ún. **Heaviside-függvény** (balról folytonos változata).

Vegyük észre: F is egy eloszlásfüggvény, mégpedig az 1 valószínűséggel konstans 0 val. változó eloszlásfüggvénye.

Tehát X_n eloszlásban tart az X konstans 0 val. változóhoz.

F az $x = 0$ pontban nem folytonos, így definíció szerint $F_{X_n}(0)$ -nak nem kell $F(0)$ -hoz tartania. Nem is tart hozzá, hiszen

$$F_{X_n}(0) = \mathbb{P}(X_n < 0) = 1/2 \text{ minden } n\text{-re!}$$

Eloszlásbeli konvergencia, 1. példa: normális eloszlás nullához tartó szórással

Legyen $n \in \mathbb{N}$ esetén X_n normális eloszlású 0 várható értékkel és $1/n$ szórásnégyzettel.

$$\lim_{n \rightarrow \infty} F_{X_n}(x) = \lim_{n \rightarrow \infty} \mathbb{P}(X_n < x) = 0, \forall x < 0, \quad \lim_{n \rightarrow \infty} F_{X_n}(x) = 1, \forall x > 0.$$

Tehát $x = 0$ kivételével $F_{X_n}(x)$ a következőhöz tart:

$$F(x) = \begin{cases} 0, & \text{ha } x \leq 0, \\ 1, & \text{ha } x > 0. \end{cases}$$

Ez az ún. **Heaviside-függvény** (balról folytonos változata).

Vegyük észre: F is egy eloszlásfüggvény, mégpedig az 1 valószínűséggel konstans 0 val. változó eloszlásfüggvénye.

Tehát X_n eloszlásban tart az X konstans 0 val. változóhoz.

Megjegyzés: A De Moivre–Laplace tételben diszkrét val. változók konvergáltak eloszlásban egy folytonoshoz.

Itt folytonos val. változók konvergálnak egy diszkrétéhez.

Eloszlásbeli konvergencia, 2. példa: a binomiális eloszlás Poisson-approximációja

Legyen $\lambda > 0$, és $n \in \mathbb{N}$ esetén legyen $X_n \sim B(n, \lambda/n)$. Legyen továbbá $X \sim \text{Pois}(\lambda)$.

Korábban láttuk: minden $k \in \mathbb{N}$ esetén

$$\lim_{n \rightarrow \infty} \mathbb{P}(X_n = k) = \mathbb{P}(X = k). \quad (4)$$

Mivel X_n és X értékkészlete is a $\{0, 1, 2, \dots\}$ halmaz egy-egy részhalmaza, F_{X_n} és F_X is konstans a $(-\infty, 0]$, $(0, 1]$, $(1, 2]$, $(2, 3]$ stb. intervallumokon.

Emiatt az eloszlásbeli konvergencia következik a (4) egyenletből (világos? Elég azt belátni, hogy $F_{X_n}(k) \rightarrow F_X(k)$ minden $k \in \{0, 1, \dots\}$ esetén. **Bár ezekben a pontokban F_X éppen nem folytonos, de ha ezekben a pontokban teljesül a konvergencia, akkor a szakaszonkénti konstansság miatt mindenhol teljesül.**).

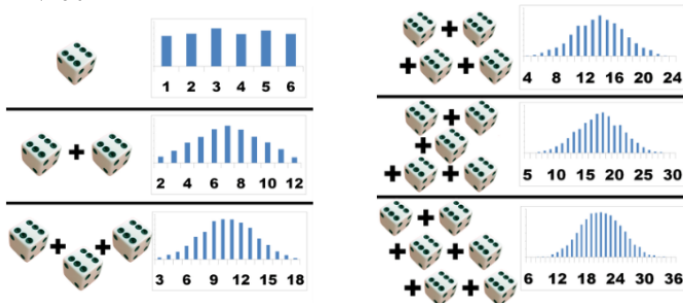
Itt diszkrét valószínűségi változók konvergálnak eloszlásban egy szintén diszkrétéhez.

Centrális határeloszlás-tétel (CHT)

(9.3.2. Tétel, CHT.) Legyen $X_1, X_2, \dots$ független, azonos eloszlású val. változók egy végtelen sorozata. Tegyük fel, hogy $\mathbb{E}(X_i^2) < \infty$, vagyis a szórásnégyzetük véges, és legyen $Z \sim N(0; 1)$. Ekkor

$$\frac{X_1 + \dots + X_n - n\mathbb{E}(X_1)}{\sqrt{n\mathbb{D}(X_1)}} \xrightarrow{d} Z,$$

ha $n \rightarrow \infty$.



Bizonyításról hamarosan.

Centrális határeloszlás-tétel (CHT)

$$\frac{X_1 + \dots + X_n - n\mathbb{E}(X_1)}{\sqrt{n\mathbb{D}(X_1)}} \xrightarrow{d} Z \sim N(0; 1).$$

Kérdés: Hogy lehet ezt felírni valószínűségek konvergenciájaként?

$$\mathbb{P}\left(\underbrace{\frac{X_1 + \dots + X_n - n\mathbb{E}(X_1)}{\sqrt{n\mathbb{D}(X_1)}}}_{= \frac{\overline{X}_n - \mathbb{E}(\overline{X}_n)}{\mathbb{D}(\overline{X}_n)},} < a\right) \xrightarrow{n \rightarrow \infty} \Phi(a), \quad a \in \mathbb{R}.$$

mivel $\mathbb{D}^2(\overline{X}_n) = \frac{n}{n^2} \mathbb{D}^2(X_1)$

A CHT-ból következik a de Moivre–Laplace-tétel: ha $X_1, \dots, X_n$ független, p valószínűségű események indikátorai, akkor $X_1 + \dots + X_n \sim B(n; p)$.

Példa: CHT

(7.13. feladat. Az előző példa törölve, mert az össznépi gyakorlatról idén visszakerült a rendes gyakorlatra)

Egy kereskedelmi rádió 100 zeneszámot ismételtet úgy, hogy a következő számot mindig a korábbiaktól függetlenül választja, azonos valószínűséggel a 100 lehetséges szám közül (így ugyanaz a szám akár egymás után többször is előfordulhat). A 100 szám közül az egyik Angéla kedvenc száma. Angéla bekapcsolja a rádiót, és addig hallgatja, amíg a kedvenc számát századszorra végig nem hallgatta, aztán kikapcsolja a rádiót.

1. Jelölje X azon zeneszámok számát, amiket Angéla a kedvenc száma **első** végighallgatásáig meghallgat (tehát beleértve a kedvenc száma első elhangzását is). Milyen eloszlású X és milyen paraméterrel/paraméterekkel? *Az eloszlás típusát nem kell megindokolni.*
2. Közelítsük alkalmas approximációval annak a valószínűségét, hogy Angéla a rádió kikapcsolása előtt legalább 10 200 számot hallgat végig.

(Vizsgakurzus vizsgája, 2024.06.27., 2. feladat – megoldás lásd ea.)

Ordó-jelölés nullához tartó függvényekre

A továbbiakhoz a képletek bonyolultságának csökkentése érdekében kiterjesztjük a nagy ordó-jelölést a nullához tartó függvényekre.

Emlékeztető algeből: ha $f, g: \mathbb{N} \rightarrow \mathbb{N}$, azt mondjuk, hogy $f(n) = O(g(n))$ (vagy $f(n) \in O(g(n))$), ha létezik olyan $c > 0$ és $n_0 \in \mathbb{N}$, hogy minden $n \geq n_0$ -ra $f(n) \leq cg(n)$.

Ugyanez a definíció érvényes szó szerint olyankor is, amikor f, g értékkészlete $\mathbb{R}$ ($f(n)$ és $g(n)$ nem feltétlenül mindig egész szám) és g nemnegatív.

Amivel felülről becslünk (itt $g(n)$), az legyen nemnegatív.

Érdekes eset: amikor $f(n)$ és $g(n)$ nullához tart.

Pl. $1/n^2 = O(1/n)$, $2024/n = O(1/\sqrt{n})$ stb.

Hasonlóan kiterjeszthető az Ω - és Θ -jelölés is erre az esetre.

Berry–Esseen-tétel, kiegészítő anyag

$$\mathbb{P}\left(\frac{X_1 + \dots + X_n - n\mathbb{E}(X_1)}{\sqrt{n}\mathbb{D}(X_1)} < a\right) \xrightarrow{n \rightarrow \infty} \Phi(a), \quad a \in \mathbb{R}.$$

A bal oldalon álló kifejezés $\Phi(a)$ -hoz konvergál, de **milyen gyorsan?**
Ökölszabály: $n \geq 30$ esetén nyugodtan használhatjuk a CHT-t (de Moivre–Laplace-t is).

A konvergencia sebessége, ha az X_i -knek véges a 3. momentuma:
(Berry–Esseen-tétel, néha Cramér–Berry–Esseen.) Létezik olyan $C > 0$, hogy ha $X_1, X_2, \dots$ fae. val. változók, amelyekre $\mathbb{E}(|X_1|^3) < \infty$, akkor minden n -re

$$\left| \mathbb{P}\left(\frac{X_1 + \dots + X_n - n\mathbb{E}(X_1)}{\sqrt{n}\mathbb{D}(X_1)} < a\right) - \Phi(a) \right| \leq \frac{C\mathbb{E}(|X_1|^3)}{\sqrt{n}}.$$

- ▶ C univerzális, tehát nem függ X_1 eloszlásától!
- ▶ Tehát ha $\mathbb{E}(|X_1|^3) < \infty$, akkor a CHT-val való közelítés hibájának abszolút értéke a tétel szerint $O(1/\sqrt{n})$.

Berry–Esseen-tétel, kiegészítő anyag

$$\mathbb{P}\left(\frac{X_1 + \dots + X_n - n\mathbb{E}(X_1)}{\sqrt{n}\mathbb{D}(X_1)} < a\right) \xrightarrow{n \rightarrow \infty} \Phi(a), \quad a \in \mathbb{R}.$$

A bal oldalon álló kifejezés $\Phi(a)$ -hoz konvergál, de **milyen gyorsan?**
Ökölszabály: $n \geq 30$ esetén nyugodtan használhatjuk a CHT-t (de Moivre–Laplace-t is).

A konvergencia sebessége, ha az X_i -knek véges a 3. momentuma:
(Berry–Esseen-tétel, néha Cramér–Berry–Esseen.) Létezik olyan $C > 0$, hogy ha $X_1, X_2, \dots$ fae. val. változók, amelyekre $\mathbb{E}(|X_1|^3) < \infty$, akkor minden n -re

$$\left| \mathbb{P}\left(\frac{X_1 + \dots + X_n - n\mathbb{E}(X_1)}{\sqrt{n}\mathbb{D}(X_1)} < a\right) - \Phi(a) \right| \leq \frac{C\mathbb{E}(|X_1|^3)}{\sqrt{n}}.$$

- ▶ C univerzális, tehát nem függ X_1 eloszlásától!
- ▶ Tehát ha $\mathbb{E}(|X_1|^3) < \infty$, akkor a CHT-val való közelítés hibájának abszolút értéke a tétel szerint $O(1/\sqrt{n})$.

A CHT bizonyításáról

Általános eset (amikor csak azt tesszük fel, hogy $\mathbb{E}(X_1^2) < \infty$):
BME-TTK Valszám 2, kb. 3–4 előadás.

Most: bizonyítás **vázlat** azon extra feltétel teljesülése esetén, hogy $\mathbb{E}(e^{tX_1}) < \infty$ teljesül valamilyen $t > 0$ -ra és valamilyen $t < 0$ -ra is.

- ▶ Ez jó a legismertebb nevezetes eloszlások esetén: pl. ha X_1 indikátorváltozó (vö. de Moivre–Laplace-tétel), vagy Poisson, geometriai, diszkrét egyenletes, exponenciális, egyenletes, normális.
- ▶ Nem jó pl. a Zipf(a)-eloszlás esetén semmilyen a -ra, bár van olyan a , amelyre $\mathbb{E}(X_1^2) < \infty$ és ezért a CHT igaz.

A bizonyításhoz bevezetjük a **momentumgeneráló függvény** fogalmát, ami sok mindenre használható a CHT bizonyításán kívül is.

Momentumgeneráló függvény

(9.3.4. Definíció.) Az X val. vált. momentumgeneráló függvénye:

$$t \in \mathbb{R} \mapsto M_X(t) \stackrel{\text{def.}}{=} \mathbb{E}(e^{tX}) \in (0, \infty].$$

- ▶ $M_X(0) = \mathbb{E}(1) = 1$, bármilyen X -re.
- ▶ $\mathbb{E}(e^{tX})$ a transzformált várható értékére vonatkozó képlettel számítható ki (rögzített t -re).
- ▶ Előfordulhat, hogy $M_X(t) = \infty$ minden $t \neq 0$ -ra.
(Példa: Cauchy-eloszlás, $f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$.)
- ▶ Miért hívjuk momentumgeneráló függvénynek?

Mert az n -edik deriváltja a 0-ban X n -edik momentuma:
(Állítás.) Tegyük fel, hogy $t \mapsto M_X(t)$ véges a 0 egy nyílt környezetében, azaz $(-\varepsilon, \varepsilon)$ -on valamely $\varepsilon > 0$ -ra. Ekkor

$$\left. \frac{d^n}{dt^n} \right|_{t=0} M_X(t) = \mathbb{E}(X^n).$$

(Nem bizonyítjuk.)

- ▶ A momentumgeneráló függvényt hívják Laplace-transzformáltnak is $\rightarrow$ a Fourier-transzformálttal rokon fogalom.

Momentumgeneráló függvény

(9.3.4. Definíció.) Az X val. vált. momentumgeneráló függvénye:

$$t \in \mathbb{R} \mapsto M_X(t) \stackrel{\text{def.}}{=} \mathbb{E}(e^{tX}) \in (0, \infty].$$

- ▶ $M_X(0) = \mathbb{E}(1) = 1$, bármilyen X -re.
- ▶ $\mathbb{E}(e^{tX})$ a transzformált várható értékére vonatkozó képlettel számítható ki (rögzített t -re).
- ▶ Előfordulhat, hogy $M_X(t) = \infty$ minden $t \neq 0$ -ra.
(Példa: Cauchy-eloszlás, $f_X(x) = \frac{1}{\pi} \frac{1}{1+x^2}$.)
- ▶ Miért hívjuk momentumgeneráló függvénynek?

Mert az n -edik deriváltja a 0-ban X n -edik momentuma:
(Állítás.) Tegyük fel, hogy $t \mapsto M_X(t)$ véges a 0 egy nyílt környezetében, azaz $(-\varepsilon, \varepsilon)$ -on valamely $\varepsilon > 0$ -ra. Ekkor

$$\left. \frac{d^n}{dt^n} \right|_{t=0} M_X(t) = \mathbb{E}(X^n).$$

- ▶ A momentumgeneráló függvényt hívják Laplace-transzformálnak is $\rightarrow$ a Fourier-transzformálttal rokon fogalom.

Momentumgeneráló függvény: példák

Diszkrét példa (angol kis zh-n volt hasonló tavalý, amikor voltak kis zh-k):

- ▶ Ha $M_X(t) = \frac{1}{2}e^t + \frac{1}{4}e^{2t} + \frac{1}{4}e^{4t}$, adjuk meg X várható értékét.
- ▶ Ha $\mathbb{P}(X = 1) = 1/2$ és $\mathbb{P}(X = 2) = \mathbb{P}(X = 4) = 1/4$, adjuk meg X momentumgeneráló függvényét.

Pl. a geometriai és az exponenciális eloszlásnál kis $t > 0$ -ra $M_X(t) < \infty$, de nagy t -re $M_X(t) = \infty$. Ebben a határesetben még jó lesz a CHT-biz. vázlatunk.

(Exponenciális eset $\rightarrow$ lásd előadás.

Ebből kijön pl. $\mathbb{E}(X) = \frac{1}{\lambda}$ parc. int. nélkül.)

Standard normális eloszlás $\rightarrow$ már korábban kiszámoltuk, hogy $M_Z(t) = \mathbb{E}(e^{tZ}) = e^{t^2/2}$.

Poisson(λ)-eloszlású $X \rightarrow$ egy IMSc-feladatban ki kellett számolni $\mathbb{E}(2024^X)$ -et. Ugyanígy kiszámítható $\mathbb{E}(e^{\alpha X})$ értéke minden $\alpha \in \mathbb{R}$ -re, ami mindig véges. Sőt már ki is számoltuk előadáson is (mikor? ;)).

Független valószínűségi változók összegének momentumgeneráló függvénye

A momentumgeneráló függvény egy igen hasznos tulajdonsága:

(Állítás.) Ha $t \in \mathbb{R}$ és Z_1, Z_2 független valószínűségi változók, amelyekre $M_{Z_1}(t), M_{Z_2}(t) < \infty$, akkor

$$M_{Z_1+Z_2}(t) = M_{Z_1}(t)M_{Z_2}(t).$$

(Szavakban: „konvolúció momentumgeneráló függvénye a momentumgeneráló függvények szorzata”.)

Biz.: $M_{Z_1+Z_2}(t) = \mathbb{E}(e^{t(Z_1+Z_2)}) = \mathbb{E}(e^{tZ_1}e^{tZ_2}) = \mathbb{E}(e^{tZ_1})\mathbb{E}(e^{tZ_2}) = M_{Z_1}(t)M_{Z_2}(t)$, mert Z_1 és Z_2 függetlensége miatt e^{tZ_1} és e^{tZ_2} is függetlenek.

Miért jó? $\rightarrow$ Mert a $Z_1 + Z_2$ konvolúció eloszlás-/sűrűségfüggvényét nehéz kiszámolni, de Z_1 és Z_2 momentumgeneráló függvényét gyakran könnyebb.

Momentumgeneráló függvény és eloszlásbeli konvergencia

(9.3.5. Állítás, nem bizonyítjuk.) Legyenek Y, Z illetve $Y_1, Y_2, \dots$ valószínűségi változók. Tegyük fel, hogy $M_Y(t)$ és $M_Z(t)$ minden $t \in \mathbb{R}$ esetén véges.

(1) Ha $M_Y(t) = M_Z(t)$ minden $t \in \mathbb{R}$ esetén, akkor Y és Z azonos eloszlásúak.

(2) Ha $\lim_{n \rightarrow \infty} M_{Y_n}(t) = M_Z(t)$ minden $t \in \mathbb{R}$ esetén, akkor $Y_n \xrightarrow{d} Z$.

Független normálisak összege normális, és a paraméterek összeadódnak

Példa: Lássuk be azt a korábban már kimondott állítást, hogy ha X, Y függetlenek és $X \sim N(\mu_1; \sigma_1^2)$ és $Y \sim N(\mu_2; \sigma_2^2)$, akkor $X + Y \sim N(\mu_1 + \mu_2; \sigma_1^2 + \sigma_2^2)$.
(Ahol $\mu_1, \mu_2 \in \mathbb{R}$, $\sigma_1^2, \sigma_2^2 > 0$.)

Ugyanis: már beláttuk (össznépi 7.6. feladat), hogy $t \in \mathbb{R}$ esetén $\mathbb{E}(e^{tX}) = e^{\mu_1 t + \frac{t^2 \sigma_1^2}{2}}$ és $\mathbb{E}(e^{tY}) = e^{\mu_2 t + \frac{t^2 \sigma_2^2}{2}}$.

Mivel X, Y függetlenek, ezért összegük momentumgeneráló függvénye a momentumgeneráló függvényeik szorzata:

$$\mathbb{E}(e^{t(X+Y)}) = \mathbb{E}(e^{tX})\mathbb{E}(e^{tY}) = e^{(\mu_1 + \mu_2)t + (\sigma_1^2 + \sigma_2^2)\frac{t^2}{2}}.$$

Ez pont az $N(\mu_1 + \mu_2; \sigma_1^2 + \sigma_2^2)$ eloszlás momentumgeneráló függvénye a t helyen.

Az előző dián lévő Állítás miatt ebből következik, hogy $X + Y \sim N(\mu_1 + \mu_2; \sigma_1^2 + \sigma_2^2)$.

Momentumgeneráló függvény és CHT (kiegészítő anyag)

Az eddigiek alapján most belátjuk a CHT-t az általunk vizsgált esetben (feltételezve, hogy az eddig felsorolt, nem bizonyított állítások is igazak).

A vizsgáláshoz a következő bizonyításvázlatot nem kell tudni, de a momentumgeneráló függvényekkel kapcsolatos eddigi definíciók/állítások az anyag részét képezik.

CHT: A célunk azt belátni, hogy $Y_n = \frac{\sum_{i=1}^n X_i - n\mathbb{E}(X_1)}{\sqrt{n\mathbb{D}(X_1)}}$
 $= \frac{\sum_{i=1}^n (X_i - \mathbb{E}(X_1))}{\sqrt{n\mathbb{D}(X_1)}}$ eloszlásban $Z \sim N(0; 1)$ -hez tart.

Feltesszük, hogy $M_{X_1}(t)$ véges a 0 egy nyílt környezetében.

Mostantól azt is feltesszük, hogy $\mathbb{E}(X_1) = 0$ és $\mathbb{D}^2(X_1) = 1$.

Ugyanis ha erre az esetre belátjuk a CHT-t, abból standardizálással az általános eset is adódik.

Ebben a speciális esetben Y_n alakja egyszerűbb: $Y_n = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$.

Momentumgeneráló függvény és CHT (kiegészítő anyag)

CHT: A célunk azt belátni, hogy $Y_n = \frac{\sum_{i=1}^n X_i - n\mathbb{E}(X_1)}{\sqrt{n\mathbb{D}(X_1)}}$
 $= \frac{\sum_{i=1}^n (X_i - \mathbb{E}(X_1))}{\sqrt{n\mathbb{D}(X_1)}}$ eloszlásban $Z \sim N(0; 1)$ -hez tart.

Feltesszük, hogy $M_{X_1}(t)$ véges a 0 egy nyílt környezetében.

Mostantól azt is feltesszük, hogy $\mathbb{E}(X_1) = 0$ és $\mathbb{D}^2(X_1) = 1$.

Ugyanis ha erre az esetre belátjuk a CHT-t, abból standardizálással az általános eset is adódik.

Ebben a speciális esetben Y_n alakja egyszerűbb: $Y_n = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$.

Momentumgeneráló függvény és CHT (kieg. anyag)

A 9.3.5. Állítás alapján tehát elég azt belátni, hogy $Y_n = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$ -re

$$\lim_{n \rightarrow \infty} M_{Y_n}(t) = M_Z(t) = e^{t^2/2},$$

ahol $Z \sim N(0; 1)$. Láttuk: $M_Z(t) = \mathbb{E}(e^{tZ}) = e^{t^2/2}$.

Kérdés: mi Y_n momentumgeneráló függvénye, ahol $Y_n = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$?

Ha t közel van 0-hoz, akkor $\mathbb{E}(e^{tX_1}) < \infty$, és $X_1, \dots, X_n$ függetlensége és azonos eloszlása miatt

$$\mathbb{E}(e^{tY_n}) = \mathbb{E}(e^{t \sum_{i=1}^n \frac{X_i}{\sqrt{n}}}) = \mathbb{E}\left(\prod_{i=1}^n e^{t \frac{X_i}{\sqrt{n}}}\right) = \mathbb{E}\left(e^{\frac{t}{\sqrt{n}} X_1}\right)^n = M_{X_1}(t/\sqrt{n})^n.$$

Cél: $M_{X_1}(t/\sqrt{n})^n \xrightarrow{n \rightarrow \infty} M_Z(t) = \mathbb{E}(e^{tZ}) = e^{t^2/2}$, minden $t \in \mathbb{R}$ -re.

Ezzel ekvivalens cél, ha logaritmust veszünk:

$$n \ln M_{X_1}\left(\frac{t}{\sqrt{n}}\right) \xrightarrow{n \rightarrow \infty} t^2/2, \text{ ha } t \in \mathbb{R}.$$

Momentumgeneráló függvény és CHT (kieg. anyag)

A 9.3.5. Állítás alapján tehát elég azt belátni, hogy $Y_n = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$ -re

$$\lim_{n \rightarrow \infty} M_{Y_n}(t) = M_Z(t) = e^{t^2/2},$$

ahol $Z \sim N(0; 1)$. Láttuk: $M_Z(t) = \mathbb{E}(e^{tZ}) = e^{t^2/2}$.

Kérdés: mi Y_n momentumgeneráló függvénye, ahol $Y_n = \frac{\sum_{i=1}^n X_i}{\sqrt{n}}$?

Ha t közel van 0-hoz, akkor $\mathbb{E}(e^{tX_1}) < \infty$, és $X_1, \dots, X_n$ függetlensége és azonos eloszlása miatt

$$\mathbb{E}(e^{tY_n}) = \mathbb{E}(e^{t \sum_{i=1}^n \frac{X_i}{\sqrt{n}}}) = \mathbb{E}\left(\prod_{i=1}^n e^{t \frac{X_i}{\sqrt{n}}}\right) = \mathbb{E}\left(e^{\frac{t}{\sqrt{n}} X_1}\right)^n = M_{X_1}(t/\sqrt{n})^n.$$

Cél: $M_{X_1}(t/\sqrt{n})^n \xrightarrow{n \rightarrow \infty} M_Z(t) = \mathbb{E}(e^{tZ}) = e^{t^2/2}$, minden $t \in \mathbb{R}$ -re.

Ezzel ekvivalens cél, ha logaritmust veszünk:

$$n \ln M_{X_1}\left(\frac{t}{\sqrt{n}}\right) \xrightarrow{n \rightarrow \infty} t^2/2, \text{ ha } t \in \mathbb{R}.$$

Momentumgeneráló függvény és CHT (kieg. anyag)

Végső célunk tehát: $n \ln M_{X_1}(\frac{t}{\sqrt{n}}) \xrightarrow{n \rightarrow \infty} t^2/2$, ha $t \in \mathbb{R}$.

Fejtsük Taylor-sorba $L(x) = \ln M_{X_1}(x)$ -et $x_0 = 0$ körül.

Ahogy láttuk, M_{X_1} -nek az n -edik deriváltja a 0-ban $\mathbb{E}(X_1^n)$.

Továbbá $M_{X_1}(0) = 1$, $\mathbb{E}(X_1) = 0$, $\mathbb{E}(X_1^2) = \mathbb{D}^2(X_1) + \mathbb{E}(X_1)^2 = 1$.

Szorzat- és láncszabállyal

$$\begin{aligned} L(x) = \ln M_{X_1}(x) &\underset{\text{Taylor}}{=} \underbrace{\ln M_{X_1}(0)}_{=1} + (\ln M_{X_1})'(0)x + (\ln M_{X_1})''(0)\frac{x^2}{2} + O(x^3) \\ &= 0 + \ln' M_{X_1}(0) \cdot \underbrace{M'_{X_1}(0)}_{=\mathbb{E}(X_1)=0} x \\ &\quad + \underbrace{(\ln' M_{X_1}(0) \cdot M''_{X_1}(0))}_{=1/M_{X_1}(0)} + \underbrace{\ln'' M_{X_1}(0) \cdot M'_{X_1}(0)}_{=\mathbb{E}(X_1^2)=1} \frac{x^2}{2} + O(x^3) = x^2/2 + O(x^3) \end{aligned}$$

amint $x \rightarrow 0$. Tehát tényleg azt kapjuk (minden $t \in \mathbb{R}$ -re), hogy

$$n \ln M_{X_1}(t/\sqrt{n}) = nL(t/\sqrt{n}) = n\left(\frac{t^2}{2n} + O\left(\frac{1}{n^{3/2}}\right)\right) \xrightarrow{n \rightarrow \infty} \frac{t^2}{2}.$$

Momentumgeneráló függvény és CHT (kieg. anyag)

Végső célunk tehát: $n \ln M_{X_1}(\frac{t}{\sqrt{n}}) \xrightarrow{n \rightarrow \infty} t^2/2$, ha $t \in \mathbb{R}$.

Fejtsük Taylor-sorba $L(x) = \ln M_{X_1}(x)$ -et $x_0 = 0$ körül.

Ahogy láttuk, M_{X_1} -nek az n -edik deriváltja a 0-ban $\mathbb{E}(X_1^n)$.

Továbbá $M_{X_1}(0) = 1$, $\mathbb{E}(X_1) = 0$, $\mathbb{E}(X_1^2) = \mathbb{D}^2(X_1) + \mathbb{E}(X_1)^2 = 1$.

Szorzat- és láncszabállyal

$$\begin{aligned} L(x) = \ln M_{X_1}(x) &\underset{\text{Taylor}}{=} \underbrace{\ln M_{X_1}(0)}_{=1} + (\ln M_{X_1})'(0)x + (\ln M_{X_1})''(0)\frac{x^2}{2} + O(x^3) \\ &= 0 + \ln' M_{X_1}(0) \cdot \underbrace{M'_{X_1}(0)}_{=\mathbb{E}(X_1)=0} x \\ &\quad + \underbrace{(\ln' M_{X_1}(0) \cdot M''_{X_1}(0))}_{=1/M_{X_1}(0)} + \underbrace{\ln'' M_{X_1}(0) \cdot M'_{X_1}(0)}_{=\mathbb{E}(X_1)=0} \frac{x^2}{2} + O(x^3) = x^2/2 + O(x^3) \end{aligned}$$

amint $x \rightarrow 0$. Tehát tényleg azt kapjuk (minden $t \in \mathbb{R}$ -re), hogy

$$n \ln M_{X_1}(t/\sqrt{n}) = nL(t/\sqrt{n}) = n\left(\frac{t^2}{2n} + O\left(\frac{1}{n^{3/2}}\right)\right) \xrightarrow{n \rightarrow \infty} \frac{t^2}{2}.$$

Nagy számok törvénye, konvergenciafajták

A relatív gyakoriság tart a valószínűséghez.

“Független, azonos eloszlású kísérletek átlaga a várható értékhez tart.”

Mit értünk az alatt, hogy “tart”?

Egyfajta konvergenciátípust már láttunk: eloszlásban való konvergencia. $F_{X_n}(x) \rightarrow F_Z(x)$.

Kétféle kerül még elő az idén: $Z, X_1, X_2, \dots$ val. változókra

- **Sztocasztikus (avagy valószínűségben való) konvergencia:** X_n valószínűségben konvergál Z -hez, ha $\forall \varepsilon > 0$ -ra teljesül, hogy

$$\lim_{n \rightarrow \infty} \mathbb{P}(|X_n - Z| > \varepsilon) = 0.$$

- **1 valószínűséggel való (avagy majdnem biztos) konvergencia:** X_n 1 valószínűséggel/majdnem biztosan konvergál Z -hez, ha

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} X_n = Z\right) = \mathbb{P}\left(\left\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} X_n(\omega) = Z(\omega)\right\}\right) = 1.$$

Ez a legerősebb a háromból, míg az eloszlásban való konvergencia a leggyengébb.

Kapcsolat a konvergenciafajták között (kiegészítő anyag)

Mit értünk a konvergenciafajták erősségén/gyengeségén?

- ▶ Ha X_n majdnem biztosan/1 valószínűséggel konvergál Z -hez, akkor sztochasztikusan/valószínűségben is konvergál.
- ▶ Az előbbi állítás megfordítása nem igaz. Igaz viszont: ha X_n sztochasztikusan konvergál Z -hez, akkor létezik egy olyan $X_{n_1}, X_{n_2}, \dots$ végtelen **részsorozat** ($1 \leq n_1 < n_2 < \dots$), amely 1 valószínűséggel Z -hez konvergál.

Ezért néha szokás a sztochasztikus konvergenciát „gyenge”, a majdnem biztos konvergenciát pedig „erős” konvergenciának is nevezni. Erre utal a következőkben a *nagy számok erős és gyenge törvénye*.

A fenti állításokat itt nem tudjuk bizonyítani (lásd mértékelmélet kurzus.)

Általában: „majdnem biztosan” = „1 valószínűséggel”. Pl.: egy $U(0; 1)$ eloszlású X val. változó értéke majdnem biztosan nem $1/2$. $\{X = 1/2\}$ nem lehetetlen, de 0 valószínűségű esemény.

Nagy számok törvénye(i)

(9.2.1. Tétel.) Legyen $X_1, X_2, \dots$ független, azonos eloszlású valószínűségi változók egy sorozata.

Tegyük fel, hogy $\mathbb{E}(X_n) = \mu \in \mathbb{R}$ és $\mathbb{D}(X_n) = \sigma > 0$ minden n -re. Jelölje $\overline{X}_n$ az $\frac{X_1 + \dots + X_n}{n}$ átlagot.

- ▶ **Nagy számok gyenge törvénye:** Tetszőleges $\varepsilon > 0$ esetén

$$\lim_{n \rightarrow \infty} \mathbb{P}(|\overline{X}_n - \mu| \geq \varepsilon) = 0.$$

- ▶ **Nagy számok erős törvénye:**

$$\mathbb{P}\left(\lim_{n \rightarrow \infty} \overline{X}_n = \mu\right) = \mathbb{P}\left(\left\{\omega \in \Omega \mid \lim_{n \rightarrow \infty} \overline{X}_n(\omega) = \mu\right\}\right) = 1.$$

- ▶ (Gyenge törvény biz. Csebisevvel: lásd előadás/jegyzet)
- ▶ A gyenge törvény (indikátor X_i -k esetén) Bernoulli, az erős törvény Kolmogorov nevéhez fűződik.
- ▶ A törvények igazak akkor is, ha a szórások nem végesek, csak $\mathbb{E}(|X_1|) < \infty$ teljesül. Erős törvény bizonyítása az általános esetben: BME-TTK, Valszám 2, kb. két előadás.

Nagy számok törvénye, speciális eset: relatív gyakoriság

Legyenek $A_1, \dots, A_n$ együttesen független, p valószínűségű események.

$$\mathbb{P}(A_1 \cup \dots \cup A_n) = \mathbb{P}\left(\sum_{i=1}^n \mathbb{1}_{A_i} > 0\right).$$

A tétel miatt $\overline{X_n} = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{A_i}$ -re:

- ▶ Annak a valószínűsége, hogy p -től $\overline{X_n}$ egy akármilyen kicsi rögzített ε -nal is eltér, nullához tart. (A gyenge törvény miatt.)
- ▶ 1 valószínűséggel olyan kimenetelen vagyunk, ahol az $n \mapsto \overline{X_n}(\omega)$ számsorozat p -hez tart. (Az erős törvény miatt.)
- ▶ Következmény: „whatever can happen, will happen”.
1 valószínűséggel olyan kimenetelen vagyunk, ahol az $A_1, A_2, \dots$ események közül legalább az egyik (sőt, végtelen sok) bekövetkezik.

(Lásd még: majom az írógépnél, pl.

https://hu.wikipedia.org/wiki/V%C3%A9gtelen_sok_majom_%C3%A9s_%C3%ADr%C3%B3g%C3%A9p_t%C3%A9le)

Miért 1 valószínűséggel?

Jogos kérdés: a nagy számok erős törvényében miért kell az “1 valószínűséggel” feltétel? Nem lehetne esetleg kihagyni?

Válasz: Nem, még diszkrét valószínűségi változók esetén sem. Mértékelméleti hiányosságok miatt itt nem beszélünk arról, hogy a törvényben szereplő végtelen sok valószínűségi változót $(X_1, X_2, \dots)$ milyen $(\Omega, \mathcal{F}, \mathbb{P})$ valószínűségi mezőn lehet együttesen definiálni.

Intuitíció: Ha pl. $X_1, X_2, \dots$ független indikátorváltozók, ahol $\mathbb{P}(X_i = 1) = \mathbb{P}(X_i = 0) = 1/2 \in (0, 1)$, akkor egy megfelelő Ω eseménytér elemei végtelen 0–1 sorozatok, ahol az i -edik bit az X_i értékét adja meg.

Ez az eseménytér tartalmaz olyan elemeket is, ahol a mintaátlag nem tart $1/2$ -hez, pl. a $00000\dots$ sorozatnál 0 -hoz, az $101101110\dots$ sorozatnál 1 -hez, a $0010100101\dots$ sorozatnál $2/5$ -höz tart.

Így **nem lehetetlen esemény**, hogy a mintaátlag nem tart $1/2$ -hez. Az erős törvény szerint azonban **0 valószínűségű esemény**.

Kitérő: konvergenciák

Konvergencia-típusok, amiket említettünk:

- ▶ egy valószínűségű konvergencia $\mathbb{P}(\lim_{n \rightarrow \infty} \dots) = 1$
- ▶ sztochasztikus konvergencia $\mathbb{P}(|\dots| \geq \varepsilon) \xrightarrow{n \rightarrow \infty} 0$
- ▶ eloszlásbeli konvergencia $F_{X_n}(x) \rightarrow F_Z(x)$.

Tegyük fel, hogy $\mathbb{E}(X_i) = 0$ és $\mathbb{D}(X_i) = 1$. Ekkor

- ▶ NSZT: $\overline{X_n} \rightarrow 0$ (1. és 2. értelemben)
- ▶ CHT: $\sqrt{n} \cdot \overline{X_n} \rightarrow Z \sim N(0; 1)$ (3. értelemben)

NSZT további alkalmazásai: lásd a statisztika résznél (empirikus eloszlásfüggvény konvergenciája).