Wahrscheinlichkeitstheorie

8. Woche

26-27. Oktober 2022

Eloszlásfüggvény

Emlékeztető:

Ha $X : \Omega \to \mathbb{R}$ egy valószínűségi változó, akkor azt az $F_X : \mathbb{R} \to [0;1]$ függvényt, melyet az

$$F_X(t) = \mathbb{P}(X < t)$$

formula definiál minden $t \in \mathbb{R}$-re, az X valószínűségi változó *eloszlásfüggvényének* nevezzük.

Valószínűségi változók együttes eloszlása

- A diszkrét esetben a súlyfüggvény írta le egy változó eloszlását,
- általános esetben az eloszlásfüggvényt használjuk erre a célra.
- A diszkrét esetben a súlyfüggvény többváltozós általánosítása kézenfekvő módon írta le több valószínűségi változó együttes eloszlását,
- az általános esetben analóg módon kaphatjuk meg azt az eloszlásfüggvény általánosításával.

Valószínűségi változók együttes eloszlása

Definíció. Legyen $\underline{X} = (X_1, \ldots, X_n)$ egy valószínűségi vektorváltozó, vagyis $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ azonos valószínűségi mezőn értelmezett tetszőleges valószínűségi változók, ekkor az *együttes eloszlásfüggvényüket* az

$$F_{\underline{X}}(t_1, \ldots, t_n) = \mathbb{P}(X_1 < t_1, \ldots, X_n < t_n)$$

formula definiálja.

Hasonlóan az egy változós esethez, az együttes eloszlásfüggvény megadása elegendő ahhoz, hogy minden, az $X_1, \ldots, X_n$ változók segítségével leírható esemény valószínűségét meghatározzuk, tehát azt mondjuk, hogy az eloszlásfüggvény leírja a változók együttes eloszlását.

Valószínűségi változók függetlensége

Fontos speciális eset: a változók viselkedése nincs hatással egymásra, azaz függetlenek egymástól.

Definíció. Az $X, Y : \Omega \to \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett valószínűségi változókat *függetlennek* nevezzük, ha minden $s, t \in \mathbb{R}$ esetén az $\{X < s\}$ és $\{Y < t\}$ események függetlenek, azaz ha

$$F_{X,Y}(s,t) = \mathbb{P}(X <, Y < t) =$$
$$= \mathbb{P}(X < s)\mathbb{P}(Y < t) = F_X(s)F_Y(t)$$

teljesül.

Valószínűségi változók függetlensége

- A diszkrét esetben más volt a függetlenség definíciója, ott az $\{X = s\}$ és $\{X = t\}$ események függetlenségét követeltük meg.
- Belátható, hogy két *diszkrét* változó pontosan akkor független régi definíció értelmében, ha az utóbbi újabb definíció értelmében azok.
- Vagyis az új függetlenségfogalom a korábbinak az általánosítása.

Valószínűségi változók függetlensége

Példa. Az előző előadáson megelőlegeztük, most pedig belátjuk, hogy egy téglalapon egyenletesen véletlenszerűen választott pont koordinátái egymástól függetlenek.

- $\Omega = [a; b] \times [c; d]$, válasszunk itt véletlenszerűen egy pontot.
- Legyen X a pont első, míg Y a pont második koordinátája, ekkor $X : \Omega \to [a; b]$ és $Y : \Omega \to [c; d]$ valószínűségi változók.
- Ekkor minden $s \in [a; b]$ és $t \in [c; d]$ esetén

$$\mathbb{P}(X < s) = \frac{T([a; s] \times [c; d])}{T(\Omega)} = \frac{(s-a)(d-c)}{(b-a)(d-c)} = \frac{s-a}{b-a},$$

$$\mathbb{P}(Y < t) = \frac{T([a; b] \times [c; t])}{T(\Omega)} = \frac{(b-a)(t-c)}{(b-a)(d-c)} = \frac{t-c}{d-c}.$$

Valószínűségi változók függetlensége

- Minden $s \in [a; b]$ és $t \in [c; d]$ esetén

$$\mathbb{P}(X < s) = \frac{T([a; s] \times [c; d])}{T(\Omega)} = \frac{(s-a)(d-c)}{(b-a)(d-c)} = \frac{s-a}{b-a},$$

$$\mathbb{P}(Y < t) = \frac{T([a; b] \times [c; t])}{T(\Omega)} = \frac{(b-a)(t-c)}{(b-a)(d-c)} = \frac{t-c}{d-c}.$$

- Továbbá

$$\mathbb{P}(X < s, Y < t) = \frac{T([a; s] \times [c; t])}{T(\Omega)} = \frac{(s-a)(t-c)}{(b-a)(d-c)}.$$

- Tehát minden $s \in [a; b]$ és $t \in [c; d]$ esetén

$$\mathbb{P}(X < s, Y < t) = \mathbb{P}(X < s)\mathbb{P}(Y < t)$$

teljesül.

Valószínűségi változók függetlensége

- Ha $s < a$ vagy $t < c$, akkor könnyen láthatóan a $\mathbb{P}(X < s, Y < t)$ $\mathbb{P}(X < s)$, $\mathbb{P}(Y < t)$ valószínűségek mindegyike 0 lesz, így

  $$\mathbb{P}(X < s, Y < t) = \mathbb{P}(X < s)\mathbb{P}(Y < t)$$

  most is teljesül.
- Ha pedig $s > b$ és $t > d$ akkor a fenti valószínűségek értéke 1, így az egyenlet ekkor is teljesül.

Valószínűségi változók függetlensége

- Ha $s \in [a; b]$ és $t > d$, akkor

$$\begin{aligned}
\mathbb{P}(X < s, Y < t) &= \frac{T([a; s] \times [c; d])}{T(\Omega)} \\
&= \mathbb{P}(X < s) = \mathbb{P}(X < s) \cdot 1 \\
&= \mathbb{P}(X < s)\mathbb{P}(Y < t).
\end{aligned}$$

- A maradék egy eset, tehát amikor $s > b$ és $t \in [c; d]$, ugyanígy kezelhető.

Valószínűségi változók függetlensége

Definíció. Az $X_1, \ldots, X_n : \Omega \to \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett valószínűségi változókat *páronként/együttesen függetlennek* nevezzük, ha minden $t_1, \ldots, t_n \in \mathbb{R}$ esetén az $\{X_1 < t_1\}, \ldots, \{X_n < t_n\}$ események páronént/együttesen függetlenek.

Abszolút folytonos változók

Példa. Válasszunk egy pontot véletlenszerűen a síkon az origó középpontú egység sugarú körlapon, és legyen Z a választott pont távolsága az origótól. Előző előadáson láttuk:

$$F_Z(t) = \mathbb{P}(Z < t) = \begin{cases} 0, & \text{ha } t \leq 0, \\ t^2, & \text{ha } 0 < t \leq 1, \\ 1, & \text{ha } t > 1. \end{cases}$$

Továbbá $s, t \in \mathbb{R}$, $s < t$ esetén

$$\mathbb{P}(s \leq Z < t) = F_Z(t) - F_Z(s).$$

Abszolút folytonos változók

Mivel $\mathbb{P}(Z = s) = 0$ minden $s \in \mathbb{R}$ esetén, valamint

$$\{s \leq Z < t\} = \{Z = s\} \cup \{s < Z < t\}$$

egy diszjunkt felbontás (azaz a jobb oldalon egymást kizáró események uniója áll), így

$$\begin{aligned}\mathbb{P}(s < Z < t) &= \mathbb{P}(s \leq Z < t) - \mathbb{P}(Z = s) \\ &= \mathbb{P}(s \leq Z < t) = F_Z(t) - F_Z(s)\end{aligned}$$

is érvényes. Hasonlóan: az $[s; t]$, $[s; t)$, $(s; t]$ és $(s; t)$ intervallumokba ugyanolyan eséllyel esik a Z változó.

Ez érvényes minden olyan valószínűségi változóra, ami az adott intervallum végpontjainak értékét 0 valószínűséggel veszi fel. Az alább definiált abszolút folytonos valószínűségi változóknál ez így lesz.

Abszolút folytonos változók

A mi példánkban:

- $\mathbb{P}(Z = t) = 0$ érvényes minden $t \in \mathbb{R}$ esetén,
- mégis van arról információnk, hogy Z milyen eséllyel lesz *közel* t-hez,
- például t egy kis $\varepsilon > 0$ sugarú környezetébe

  $$\mathbb{P}(t - \varepsilon < Z < t + \varepsilon) = F_Z(t + \varepsilon) - F_Z(t - \varepsilon)$$

  valószínűséggel esik Z értéke.
- Ez az érték kis ε-ra kicsi lesz, mert F_Z folytonossága miatt a fenti különbség mindkét tagja $F_Z(t)$-hez, különbségük pedig így 0-hoz tart, ahogy ε tart 0-hoz.
- Mindez összhangban van azzal, hogy szemléletes ténnyel, hogy a $\mathbb{P}(t - \varepsilon < Z < t + \varepsilon)$ valószínűség $\mathbb{P}(Z = t) = 0$-hoz tart, ahogy a t körüli intervallum hossza 0-hoz tart.

Abszolút folytonos változók

- Vegyük figyelembe most az intervallum hosszát is, azaz tekintsük a valószínűség és az intervallum hosszának arányát:

$$\frac{\mathbb{P}(t-\varepsilon < Z < t+\varepsilon)}{2\varepsilon} = \frac{F_Z(t+\varepsilon) - F_Z(t-\varepsilon)}{2\varepsilon}.$$

- Ha $\varepsilon \to 0$, akkor a fenti kifejezés határértékét - amennyiben létezik - tekinthetjük egyfajta súlynak, ami meghatározza, hogy milyen valószínűséggel esik Z a t közelébe.

Abszolút folytonos változók

$$\frac{F_Z(t+\varepsilon)-F_Z(t-\varepsilon)}{2\varepsilon} = \frac{F_Z(t+\varepsilon)-F_Z(t)+F_Z(t)-F_Z(t-\varepsilon)}{2\varepsilon}$$

$$= \frac{1}{2}\cdot\frac{F_Z(t+\varepsilon)-F_Z(t)}{(t+\varepsilon)-t} + \frac{1}{2}\cdot\frac{F_Z(t-\varepsilon)-F_Z(t)}{(t-\varepsilon)-t}$$

$$\to F_Z'(t),$$

ahogy $\varepsilon \to 0$, amennyiben a derivált létezik t-ben.

Abszolút folytonos változók

$$F_Z(t) = \mathbb{P}(Z < t) = \begin{cases} 0, & \text{ha } t \le 0, \\ t^2, & \text{ha } 0 < t \le 1, \\ 1, & \text{ha } t > 1. \end{cases}$$

- F_Z deriválható minden $t \neq 0, 1$ esetén, mert itt vagy egy konstanssal vagy egy polinomfüggvénnyel egyezik meg.
- A 0 pontban a jobb- és baloldali deriváltak léteznek, és mindkettő 0-val egyenlő, így a derivált létezik az $t = 0$ pontban is.

Abszolút folytonos változók

A $t = 1$ pontban viszont nem deriválható a függvény, hiszen, habár a bal és jobb oldali deriváltak léteznek, azok nem egyeznek meg:

$$\lim_{t \to 1-0} \frac{F_Z(t) - F_Z(1)}{t-1} = \lim_{t \to 1-0} \frac{t^2 - 1}{t-1} = \lim_{t \to 1-0} t + 1 = 2,$$

de

$$\lim_{t \to 1+0} \frac{F_Z(t) - F_Z(1)}{t-1} = \lim_{t \to 1+0} \frac{1-1}{t-1} = 0.$$

Abszolút folytonos változók

Definiáljuk az $f_Z(t)$ függvényt a következőképp: legyen $f_Z(t) = F'_Z(t)$, ha $t \neq 1$, és legyen $f_Z(1) = 0$. Tehát

$$f_Z(t) = \begin{cases} 2t, & \text{ha } 0 < t < 1 \\ 0 & \text{különben.} \end{cases}$$

A Newton–Leibniz-formula szerint ekkor

$$\mathbb{P}(s < Z < t) = F_Z(t) - F_Z(s) = \int_s^t f_Z(y)\, dy$$

érvényes minden olyan $s, t \in \mathbb{R}$, $s < t$ számpárra, ahol az F_Z deriválható az $(s; t)$ intervallumon és folytonos az $[s; t]$ zárt intervallumon, vagyis ez igaz minden $s < t \leq 1$ és $1 \leq s < t$ esetén.

Abszolút folytonos változók

Ha $s < 1 < t$, akkor kettébonthatjuk az $(s;t)$ interevallumot, és tagonként alkalmazva a Newton–Leibniz-formulát láthatjuk, hogy a fenti egyenlőség ekkor is teljesül:

$$\begin{aligned}\int_s^t f_Z(y)\,dy &= \int_s^1 f_Z(y)\,dy + \int_1^t f_Z(y)\,dy \\ &= F_Z(1) - F_Y(s) + F_Z(t) - F_Z(1) = F_Z(t) - F_Z(s).\end{aligned}$$

Abszolút folytonos változók

Ezért ha f_Z folytonos t-ben, és $\varepsilon > 0$ kicsi, akkor

$$\begin{aligned}\mathbb{P}(t-\varepsilon < Z < t+\varepsilon) &= F_Z(t+\varepsilon) - F_Z(t-\varepsilon) \\ &= \int_{t-\varepsilon}^{t+\varepsilon} f_Z(y)\,dy \approx 2\varepsilon \cdot f_Z(t),\end{aligned}$$

vagyis $f_Z(t)$-re valóban tekinthetünk egyfajta súlyként, ami leírja, hogy milyen valószínűséggel esik a Z értéke a t egy kis környezetébe.

Abszolút folytonos változók

Ha most az

$$F_Z(t) - F_Z(s) = \int_s^t f_Z(y)\,dy$$

formulában $s \to -\infty$, akkor $F_Z(s)$ 0-hoz tart (pontosabban már minden $s \leq 0$ esetén 0), így az integrálás segítségével az eloszlásfüggvény értéke a sűrűségfüggvényből:

$$F_Z(t) = \int_{-\infty}^t f_Z(y)\,dy.$$

Ez az a formula, aminek segítségével az abszolút folytonos valószínűségi változókat *definiáljuk*:

Abszolút folytonos változók

Definíció. Az X valószínűségi változót *abszolút folytonosnak* nevezzük, ha létezik olyan $f_X : \mathbb{R} \to [0; \infty)$ nemnegatív függény, amelyre az $\int_{-\infty}^{\infty} f_X(y)\, dy$ Riemann-integrál létezik, és minden $t \in \mathbb{R}$ esetén

$$F_X(t) = \int_{-\infty}^{t} f_X(y)\, dy,$$

ahol F_X az X eloszlásfüggvénye. Ekkor az f_X függvényt az X változó *sűrűségfüggvényének* nevezzük.

Abszolút folytonos változók

Megjegyzés.

- Ezen a kurzuson a későbbiekben az abszolút folytonos változókra gyakran csak folytonos valószínűségi változókként hivatkozunk.
- Az irodalomban azonban folytonos valószínűségi változó alatt olyan változót értenek, amelyeknek az eloszlásfüggvénye folytonos.
- Mivel az abszolút folytonos valószínűségi változók eloszlásfüggvénye egy sűrűségfüggvény integrálfüggvényeként áll elő, így folytonos.
- Tehát az abszolút folytonos valószínűségi változók a fenti értelemben folytonosak, azonban általában ez fordítva nem igaz.
- Mivel mi a továbbiakban kizárólag abszolút folytonos változókkal foglalkozunk, a két elnevezést szinonimaként használjuk.

Abszolút folytonos változók

Megjegyzés.

- A fenti definícióban lévő f_X sűrűségfüggvény nem egyértelmű.
- Ha például véges sok ponton megváltoztatjuk az értékét, akkor ez az f_X integrálját nem változtatja meg, a fenti definiáló tulajdonság továbbra is érvényben marad, tehát az így kapott függvény szintén sűrűrségfüggvény lesz.

Abszolút folytonos változók

A fentiek szerint a Z változó (abszolút) folytonos. Az ezen a példán illusztrált módszer általában is alkalmazható a folytonosság igazolására:

Tétel. Legyen X olyan valószínűségi változó, amelynek F_X eloszlásfüggvénye folytonos és véges sok pont kivételével mindenhol deriválható. Ekkor az X valószínűségi változó abszolút folytonos, és az

$$f_X(t) = \begin{cases} F_X'(t), & \text{ahol a derivált létezik,} \\ 0 & \text{különben} \end{cases}$$

függvény sűrűségfüggvénye X-nek.

Abszolút folytonos változók

Állítás. Legyen X egy folytonos valószínűségi változó, amelynek sűrűségfüggvénye f_X. Ekkor tetszőleges $s, t \in \mathbb{R}$, $s < t$ esetén

$$\mathbb{P}(s \leq X < t) = \int_s^t f_X(y)\, dy.$$

Bizonyítás. Az előző előadáson láttuk, hogy

$$\mathbb{P}(s \leq X < t) = F_X(t) - F_X(s).$$

Az eloszlásfüggvény értékei felírhatók az X sűrűségfüggvény segítségével:

$$F_X(t) - F_X(s) = \int_{-\infty}^t f_X(y)\, dy - \int_{-\infty}^s f_X(y)\, dy = \int_s^t f_X(y)\, dy.$$

Abszolút folytonos változók

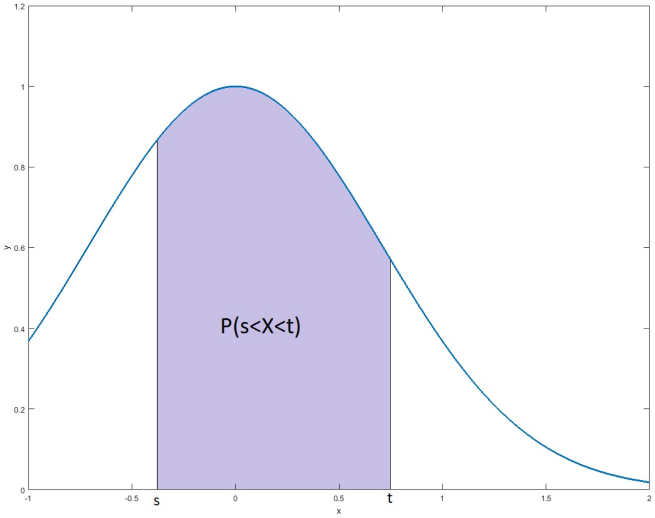
Belátható minden X folytonos valószínűségi változóra, hogy tetszőleges $t \in \mathbb{R}$ esetén $\mathbb{P}(X = t) = 0$, és így a fenti állításban szereplő eseményben mindkét egyenlőtlenségnél elhagyhatjuk vagy megengedhetjük az egyenlőséget, tehát

$$\mathbb{P}(s < X < t) = \mathbb{P}(s \leq X < t) = \mathbb{P}(s < X \leq t) = \mathbb{P}(s \leq X \leq t),$$

és ezeket a fenti integrál segítségével írhatjuk fel.

Továbbá a fenti állításban megengedhetjük az $s = -\infty$ és $t = \infty$ értéket is, ekkor a fenti egyenlőtlenségek úgy értendők, hogy X-re nem szabunk alsó ill. felső korlátot, az integrálásnál pedig a megfelelő improprius integrált tekintjük.

Abszolút folytonos változók



Abszolút folytonos változók

Ahhoz hasonlóan, ahogy az eloszlásfüggvények esetén történt, a sűrűségfüggvényeket is karakterizálhatjuk:

Tétel. Egy $f : \mathbb{R} \to [0; \infty)$ nemnegatív függvény pontosan akkor sűrűségfüggvénye egy X folytonos valószínűségi változónak, ha Riemann-integrálható a valós egyenesen, és

$$\int_{-\infty}^{\infty} f(y)\,dy = 1.$$

Bizonyítás vázlat. Az világos, hogy ha f egy X változó sűrűségfüggvénye, akkor a fenti egyenletnek teljesülie kell, hiszen

$$\int_{-\infty}^{\infty} f(y)\,dy = \lim_{t\to\infty} \int_{-\infty}^{t} f(y)\,dy = \lim_{t\to\infty} F_X(t) = 1$$

az eloszlásfüggvény tulajdonságai alapján. A tétel másik irányát nem bizonyítjuk.

Várható érték a folytonos esetben

Definíció. Legyen X abszolút folytonos valószínűségi változó, aminek sűrűségfüggvénye f_X. Ekkor X *várható értéke*

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t f_X(t)\, dt,$$

amennyiben a fenti integrál abszolút konvergens, azaz az

$$\int_{-\infty}^{\infty} |t|\, f_X(t)\, dt$$

integrál véges.

Várható érték a folytonos esetben

Példa. Számoljuk ki a fenti Z valószínűségi változó várható értékét.

$$f_Z(t) = \begin{cases} 2t, & \text{ha } 0 < t < 1 \\ 0 & \text{különben.} \end{cases}$$

A $(0;1)$ intervallumon kívül ez a függvény 0, így a várható érték kiszámításánál valójában csak ezen a véges intervallumon kell integrálnunk (és így persze abszolút konvergens lesz az integrál).
A fenti definícióba behelyettesítve, majd a Newton–Leibniz-formulát alkalmazva:

$$\mathbb{E}(Z) = \int_{-\infty}^{\infty} t f_Z(t)\, dt = \int_0^1 2t^2\, dt = \left[\frac{2t^3}{3}\right]_0^1 = \frac{2}{3}.$$

Várható érték a folytonos esetben

Tétel. Legyenek $X, Y : \Omega \to \mathbb{R}$ ugyanazon valószínűségi mezőn értelmezett abszolút folytonos valószínűségi változók, melyekre $\mathbb{E}(X)$ és $\mathbb{E}(Y)$ létezik. Legyen továbbá $c \in \mathbb{R}$ egy tetszőleges valós szám, ekkor

$$\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y), \qquad \mathbb{E}(cX) = c \cdot \mathbb{E}(X).$$

Az tételben szereplő egyenlőségek úgy értendők, hogy amenynyiben az X ill. Y várható értéke létezik, akkor az egyenlőségek mindkét oldala létezik, és azok megegyeznek.

Várható érték a folytonos esetben

Megjegyzés. A várható érték tetszőleges - azaz nem csak diszkrét vagy folytonos - esetben is definiálható, és a linearitás általában is érvényes, ha az X és Y várható értéke létezik.

Transzformált várható értéke a folytonos esetben

Tétel. Legyen X egy folytonos valószínűségi változó f_X sűrűségfüggvénnyel, $g : \mathbb{R} \to \mathbb{R}$ pedig egy olyan függvény, melyre $g(X)$ szintén valószínűségi változó. Ha az $\mathbb{E}(g(X))$ várható érték létezik, akkor

$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(t) f_X(t)\, dt.$$

Fontos speciális eset: az X^2 változó várható értéke, ahol a tételben a $g(t) = t^2$ választással élünk. Ebben az esetben a tétel állítása: ha X egy folytonos valószínűségi változó, amelyre $\mathbb{E}(X^2)$ létezik, akkor

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} t^2 f_X(t)\, dt.$$

Transzformált várható értéke a folytonos esetben

Példa. Az előző példa folytatásaként kiszámoljuk a fenti Z valószínűségi változó négyzetének a várható értékét.

$$f_Z(t) = \begin{cases} 2t, & \text{ha } 0 < t < 1 \\ 0 & \text{különben,} \end{cases}$$

tehát

$$\mathbb{E}(Z^2) = \int_{-\infty}^{\infty} t^2 f_Z(t)\, dt = \int_0^1 2t^3\, dt = \left[\frac{t^4}{2}\right]_0^1 = \frac{1}{2}.$$

Valószínűségi változók szórása

- A szórásnégyzetet és a szórást a diszkrét esethez hasonlóan definiálhatjuk a folytonos változókra, sőt, valójában teljesen általános változókra is a várható érték segítségével.
- Mivel a szórásra vonatkozó állítások igazolásánál csak a várható érték tulajdonságait használtuk, amelyek pedig a folytonos esetben ill. teljesen általánosan is érvényesek, így a szórás tulajdonásgai itt is érvényben maradnak.
- Ezek igazolása lényegében szóról szóra megegyezik a diszkrét esetben látottakkal, ezért a bizonyítások részletezését mellőzzük, és megelégszünk a definíció és a főbb állítások összefoglalásával.

Valószínűségi változók szórása

Definíció. Legyen X egy valószínűségi változó, melynek létezik a várható értéke. Ha az $\mathbb{E}((X - \mathbb{E}(X))^2)$ várható érték is létezik, akkor ezt az X *szórásnégyzetének* (vagy *varianciájának*) nevezzük, és $\mathbb{D}^2(X)$-el jelöljük. A szórásnégyzet nemnegatív gyökét az X *szórásának* nevezzük és $\mathbb{D}(X)$-el jelöljük.

Valószínűségi változók szórása

Tulajdonságok:

- A szórásnégyzet általában is egy nemnegatív valószínűségi változó várható értékeke, és így nemnegatív.
- Minden $c \in \mathbb{R}$ esetén érvényesek a

$$\mathbb{D}(X + c) = \mathbb{D}(X), \qquad \mathbb{D}(cX) = |c|\mathbb{D}(X)$$

  ez úgy értendő, hogy ha a szórás létzik, akkor a formulák mindkét oldala is, és azok egyenlők.
- A szórásnégyzet esetén az ezekkel ekvivalens formulák $\mathbb{D}^2(X + c) = \mathbb{D}^2(X)$ ill. $\mathbb{D}^2(cX) = c^2\mathbb{D}(X)$.
- Egy X egy valószínűségi változónak pontosan akkor létezik a szórásnégyzete, ha $\mathbb{E}(X^2)$ létezik, és ekkor

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

  is érvényes.

Valószínűségi változók szórása

Példa. A Z változó szórásnégyzete és szórása:

$$\mathbb{D}^2(Z) = \mathbb{E}(Z^2) - \mathbb{E}(Z)^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18},$$

illetve

$$\mathbb{D}(Z) = \frac{1}{3\sqrt{2}} = \frac{\sqrt{2}}{6}.$$

Függetlenség és szórás

Tétel. Ha X és Y *véges szórású* és *független* valószínűségi változók, akkor

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

n változó esetén: ha $X_1, \ldots, X_n$ páronként független valószínűségi változók, melyek szórása véges, akkor

$$\mathbb{D}^2(X_1 + \cdots + X_n) = \mathbb{D}^2(X_1) + \cdots + \mathbb{D}^2(X_n).$$

Eloszlások megadása

- Általában egy X valószínűségi változó eloszlását az eloszlásfüggvényével adhatjuk meg.
- Láttuk, hogy azon esemény valószínűsége, hogy X értéke egy adott intervallumba esik, ennek segítségével könnyen felírható, valójában pedig igaz az is, hogy bármely "hasonlóképp felírt" esemény valószínűsége kiszámolható az eloszlásfüggvény értékeiből.
- Azaz: ha ismerjük az eloszlásfüggvényt, akkor ismerjük az eloszlást.

Eloszlások megadása

- A diszkrét esetben a a $\mathbb{P}(X = t)$ valószínűségek megadása az eloszlásfüggvény megadásával ekvivalens.
- Általában, pl. az abszolút folytonos esetben ez nincs így.
- Az abszolút folytonos esetben azoban a sűrűségfüggvény kiváltja a súlyfüggvény szerepét, azaz az eloszlás megadása történhet a sűrűségfüggvény megadásával is.
- A sűrűségfüggvényből az eloszlásfüggvény integrálás segítségével adódik, a különböző események valószínűségei - szintén integrálás segítségével - a sűrűségfüggvényből közvetlenül megkaphatók.
- Például tetszőleges $s, t \in \mathbb{R}$, $s < t$ esetén

$$\mathbb{P}(s < X < t) = \int_s^t f_X(y)\, dy.$$

Egyenletes eloszlás

Legyegyszerűbb folytonos példa: az egyenletes eloszlás.

Definíció. Az X valószínűségi változót *egyenletes eloszlásúnak*nevezzük az $(a; b)$ intervallumon, ha sűrűségfüggvénye

$$f_X(t) = \begin{cases} \dfrac{1}{b-a}, & \text{ha } t \in (a; b), \\ 0 & \text{különben.} \end{cases}$$

Jelölése: $X \sim U(a; b)$.

Egyenletes eloszlás

Az egyenletes eloszlás eloszlásfüggvénye:

Ha $X \sim U(a;b)$, akkor (definíció szerint)

$$F_X(t) = \int_{-\infty}^{t} f_X(y)\, dy.$$

- Ez 0, ha $t \leq a$, mert az a-tól balra az f_X függvény értéke 0.
- Ha $t > b$, akkor valójában a és b között integrálunk, mert b-től jobbra f_X ismét 0. Az $(a;b)$ intervallumon integrálva a konstans $1/(b-a)$ függvényt az eredmény 1 lesz.
- Végezetül, ha $t \in (a;b]$, akkor

$$F_X(t) = \int_{a}^{t} \frac{1}{b-a}\, dy = \frac{t-a}{b-a}.$$

Egyenletes eloszlás

Összefoglalva:

$$F_X(t) = \begin{cases} 0, & \text{ha } t \leq a, \\ \dfrac{t-a}{b-a}, & \text{ha } a \leq t < b, \\ 1, & \text{ha } t > b. \end{cases}$$

Egyenletes eloszlás

Az egyenletes eloszlás várható értéke:

A várható érték definíciója szerint

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} t f_X(t)\,dt = \int_a^b \frac{t}{b-a}\,dt = \left[\frac{t^2}{2(b-a)}\right]_a^b$$

$$= \frac{b^2-a^2}{2(b-a)} = \frac{(b-a)(b+a)}{2(a-b)} = \frac{a+b}{2}.$$

Egyenletes eloszlás

Az egyenletes eloszlás szórásnégyzete:

A transzformált várható értékére vonatkozó formula szerint

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} t^2 f_X(t)\,dt = \int_a^b \frac{t^2}{b-a}\,dt = \left[\frac{t^3}{3(b-a)}\right]_a^b$$

$$= \frac{b^3-a^3}{3(b-a)} = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{a^2+ab+b^2}{3},$$

és így

$$\mathbb{D}^2(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{a^2+ab+b^2}{3} - \frac{(a+b)^2}{4}$$

$$= \frac{4a^2+4ab+4b^2-3a^2-6ab-3b^2}{12} = \frac{a^2-2ab+b^2}{12}$$

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