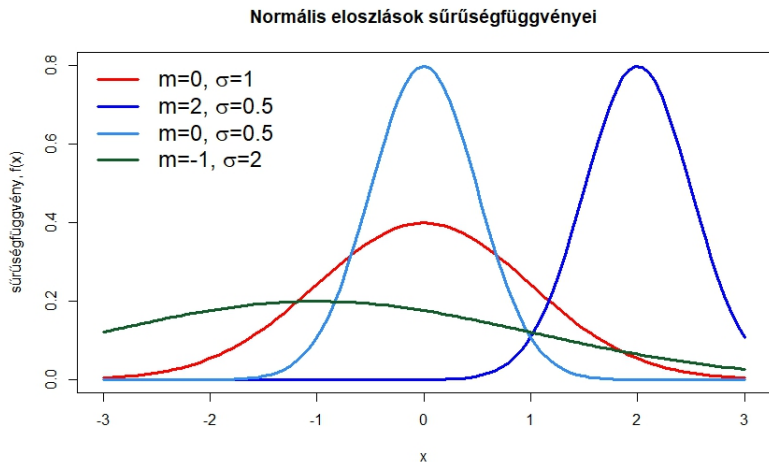
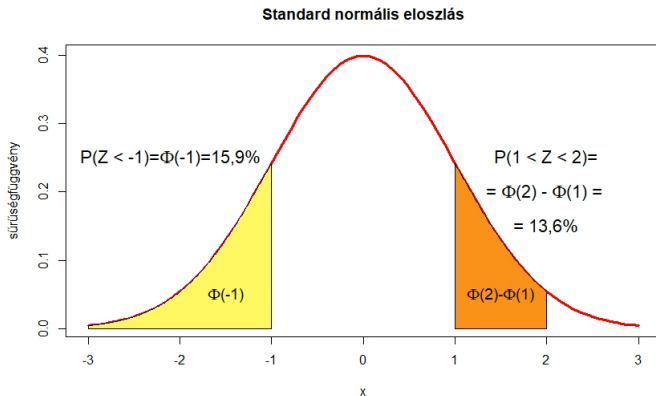


Normális eloszlás



Különböző várható értékű (m) és szórású (σ) normális eloszlások sűrűségfüggvényei

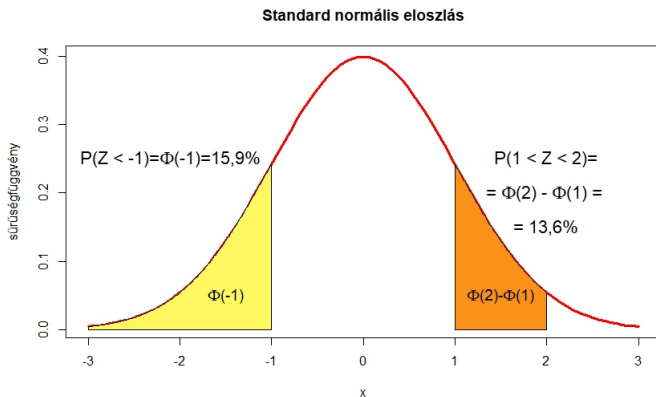
A Φ függvény



A Φ függvény a standard normális eloszlás eloszlásfüggvénye: ha $Z \sim N(0, 1)$, akkor

$$\mathbb{P}(Z \leq t) = \Phi(t) = \int_{-\infty}^t \varphi(x) dx = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

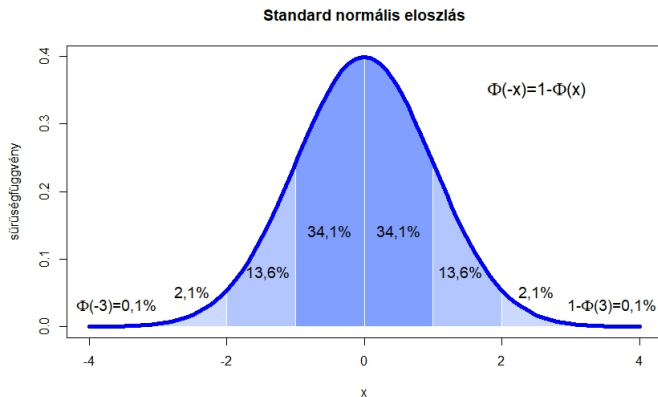
Standard normális eloszlás



A Φ függvény a standard normális eloszlás eloszlásfüggvénye: ha $Z \sim N(0, 1)$, akkor

$$\mathbb{P}(a \leq Z \leq b) = \Phi(b) - \Phi(a) = \int_a^b \varphi(x) dx = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

Standard normális eloszlás



A Φ függvény a standard normális eloszlás eloszlásfüggvénye: ha $Z \sim N(0, 1)$, akkor

$$\mathbb{P}(a \leq Z \leq b) = \Phi(b) - \Phi(a) = \int_a^b \varphi(x) dx = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

F-eloszlás

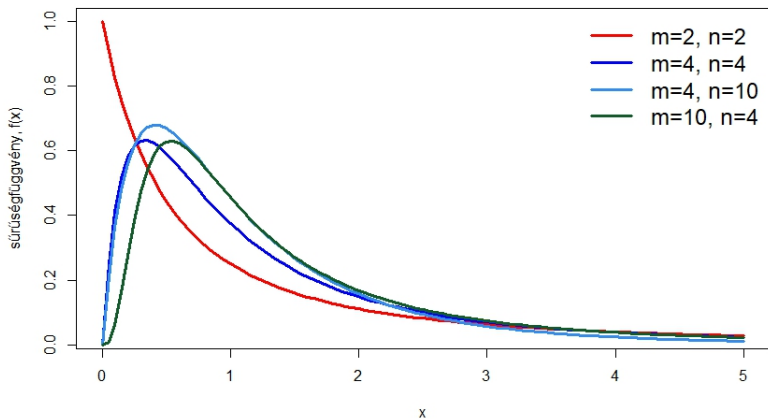
Legyenek m, n pozitív egészek, $X_1, \dots, X_m, Y_1, Y_2, \dots, Y_n$ pedig független standard normális eloszlású valószínűségi változók. Ekkor az

$$F = \frac{n(X_1^2 + X_2^2 + \dots + X_m^2)}{m(Y_1^2 + Y_2^2 + \dots + Y_n^2)}$$

valószínűségi változó eloszlását m, n szabadsági fokú **F-eloszlásnak** nevezük.

Az F -eloszlás sűrűségfüggvénye

F-eloszlások sűrűségfüggvényei



Különböző szabadsági fokú F -eloszlások sűrűségfüggvényei

χ^2 -eloszlás

Legyenek X_1, X_2, \dots, X_q független standard normális eloszlású valószínűségi változók. Az

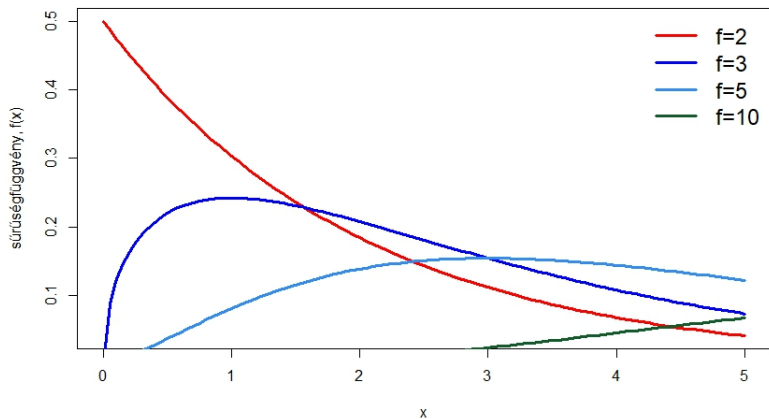
$$Y = X_1^2 + X_2^2 + \dots + X_q^2$$

valószínűségi változó eloszlását q szabadsági fokú χ^2 -eloszlásnak nevezzük. Ennek sűrűségfüggvénye:

$$f(t) = \begin{cases} \frac{t^{q/2-1}}{2^{q/2}\Gamma(q/2)} e^{-t/2}, & t \geq 0; \\ 0, & t < 0. \end{cases}$$

A χ^2 -eloszlás sűrűségfüggvénye

χ^2 -eloszlások sűrűségfüggvényei



Különböző szabadsági fokú χ^2 -eloszlások sűrűségfüggvényei

Gamma-eloszlás

Gamma-függvény. Ha $a > 0$ pozitív szám, legyen

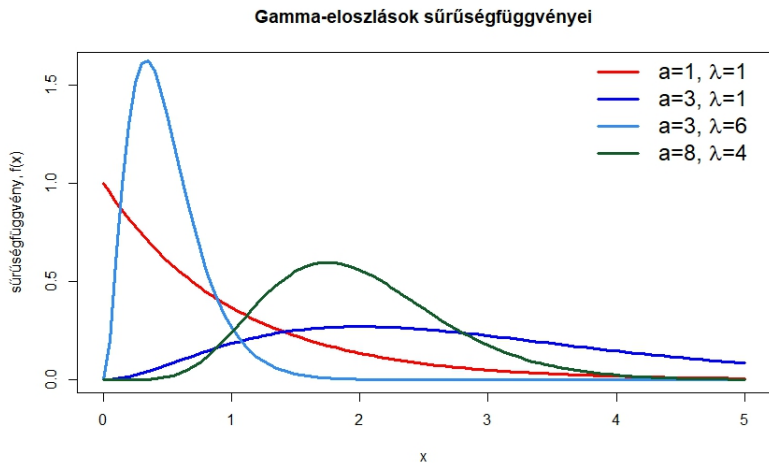
$$\Gamma(a) = \int_0^{\infty} t^{a-1} e^{-t} dt.$$

Parciális integrálással belátható, hogy $\Gamma(a) = (a-1)\Gamma(a-1)$ minden $a > 1$ -re, és így $\Gamma(n) = (n-1)!$, ha n pozitív egész.

Legyenek a és λ pozitív számok. Az X valószínűségi változó **gamma-eloszlású** a renddel és λ paraméterrel, ha sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{\lambda^a x^{a-1}}{\Gamma(a)} e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

A gamma-eloszlás sűrűségfüggvénye



Különböző szabadsági fokú gamma-eloszlások sűrűségfüggvényei

A gamma-eloszlás tulajdonságai

Az X valószínűségi változó **gamma-eloszlású** a renddel és λ paraméterrel, ha sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{\lambda^a x^{a-1}}{\Gamma(a)} e^{-\lambda x}, & x \geq 0; \\ 0, & x < 0. \end{cases}$$

- **Kapcsolat az exponenciális eloszlással:** ha $a = 1$, akkor a sűrűségfüggvény $\lambda e^{-\lambda x}$, ha $x \geq 0$, és az exponenciális eloszlást kapjuk vissza.
- **Exponenciális eloszlások összege:** ha X_1, X_2, \dots, X_n független λ paraméterű exponenciális eloszlású valószínűségi változók, akkor $X_1 + X_2 + \dots + X_n$ gamma-eloszlású $a = n$ renddel és λ paraméterrel.
- **Kapcsolat a χ^2 -eloszlással:** ha $a = q/2$ és $\lambda = 1/2$, akkor a q szabadsági fokú χ^2 -eloszlást kapjuk vissza.
- **Várható érték** és **szórás:**

$$\mathbb{E}(X) = \frac{a}{\lambda}; \quad D(X) = \frac{\sqrt{a}}{\lambda}.$$

Beta-eloszlás

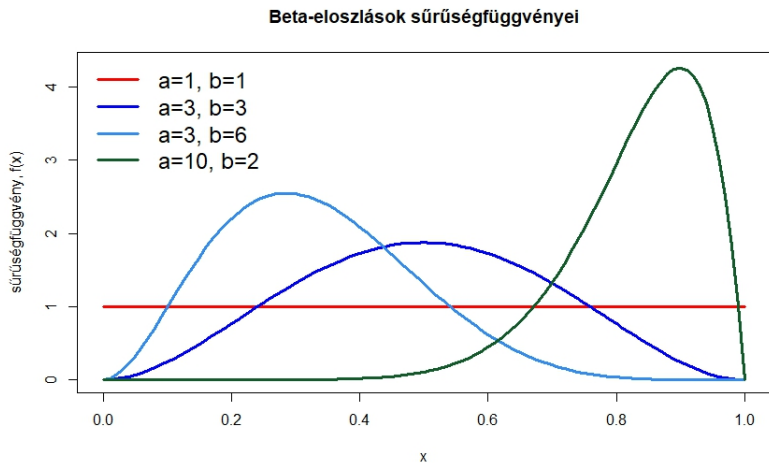
Legyenek $a, b > 1$ számok. Az X valószínűségi változó **beta-eloszlású** a és b paraméterekkel, ha sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1} (1-x)^{b-1}, & t \in [0, 1]; \\ 0, & x < 0 \text{ vagy } x > 1. \end{cases}$$

Ha X_1, X_2, \dots, X_n független, a $[0, 1]$ intervallumon egyenletes eloszlású valószínűségi változók, és X_k^* jelöli ebben a mintában a k . legnagyobb számot, akkor X_k^* eloszlása beta-eloszlás $a = k$ és $b = n - k + 1$ paraméterekkel.

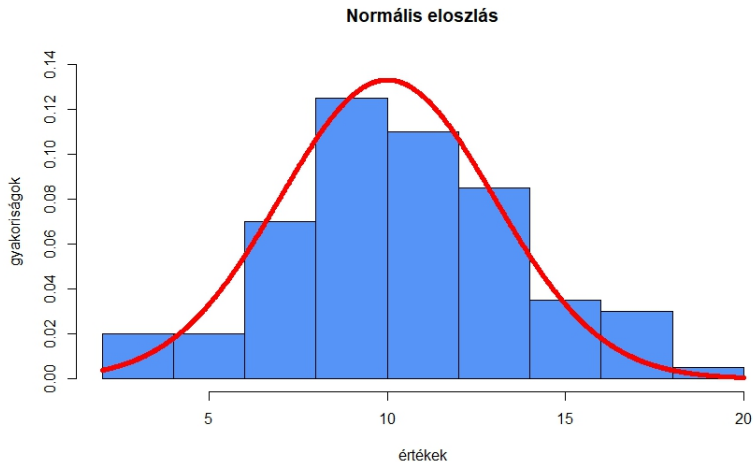
Az $a = 1$ és $b = 1$ választással az **egyenletes eloszlást** kapjuk vissza.

A beta-eloszlás sűrűségfüggvénye



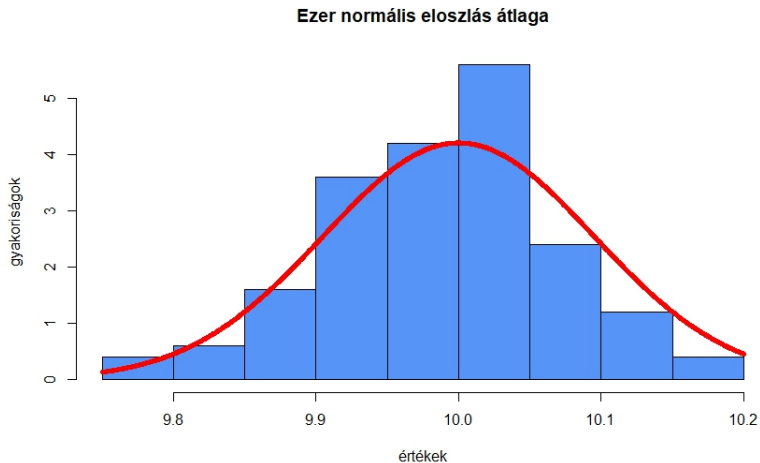
Különböző szabadsági fokú beta-eloszlások sűrűségfüggvényei

Normális eloszlás



Száz független normális eloszlású valószínűségi változó hisztogramja és a sűrűségfüggvény ($m = 10, \sigma = 3, \bar{x} = 9,88, s_n^* = 2,58$)

Normális eloszlások átlaga



Százelemű minta az alábbi eloszlásból: $n = 1000$ független normális eloszlású ($m = 10, \sigma = 3$) valószínűségi változó átlaga és az $N(10, 3/\sqrt{1000})$ normális eloszlás sűrűségfüggvénye ($\bar{x} = 9,99, s_n^* = 0,084, \sigma/\sqrt{n} = 0,095$)

Normális eloszlások átlaga

Legyenek X, Y függetlenek, normális eloszlásúak: $X \sim N(m_1, \sigma_1^2)$, $Y \sim N(m_2, \sigma_2^2)$. Ekkor a következők igazak:

- $X + b$ eloszlása normális, $m_1 + b$ várható értékkel és σ szórással;
- aX eloszlása normális am_1 várható értékkel és $|a|\sigma$ szórással;
- $X + Y$ eloszlása normális, $m_1 + m_2$ várható értékkel és $\sqrt{\sigma_1^2 + \sigma_2^2}$ szórással.

Emlékeztető: $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$, és ha X és Y függetlenek, akkor $D^2(X + Y) = D^2(X) + D^2(Y)$.

Normális eloszlások átlaga

Legyenek X, Y függetlenek, normális eloszlásúak: $X \sim N(m_1, \sigma_1^2)$, $Y \sim N(m_2, \sigma_2^2)$. Ekkor a következők igazak:

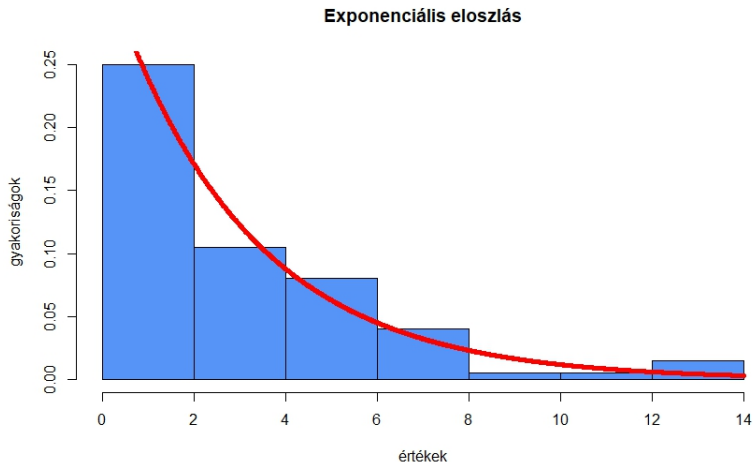
- $X + b$ eloszlása normális, $m_1 + b$ várható értékkel és σ szórással;
- aX eloszlása normális am_1 várható értékkel és $|a|\sigma$ szórással;
- $X + Y$ eloszlása normális, $m_1 + m_2$ várható értékkel és $\sqrt{\sigma_1^2 + \sigma_2^2}$ szórással.

Emlékeztető: $\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y)$, és ha X és Y függetlenek, akkor $D^2(X + Y) = D^2(X) + D^2(Y)$.

Ebből következik: ha X_1, \dots, X_n független normális eloszlásúak m várható értékkel és σ szórással, akkor

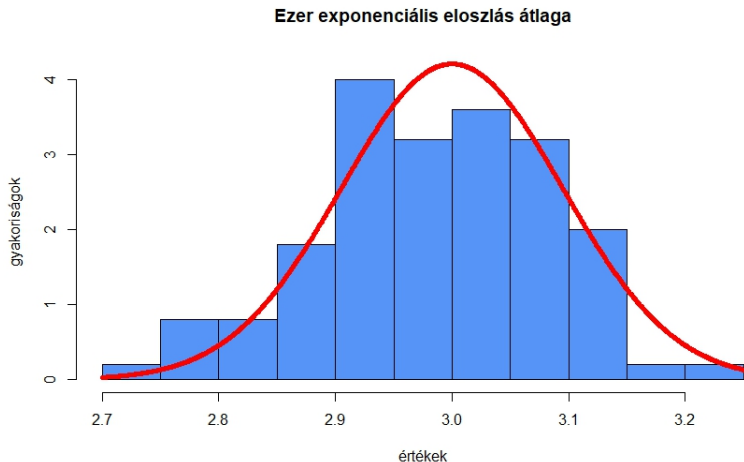
$$\frac{X_1 + \dots + X_n}{n} \sim N\left(m, \frac{\sigma^2}{n}\right)$$

Exponenciális eloszlás



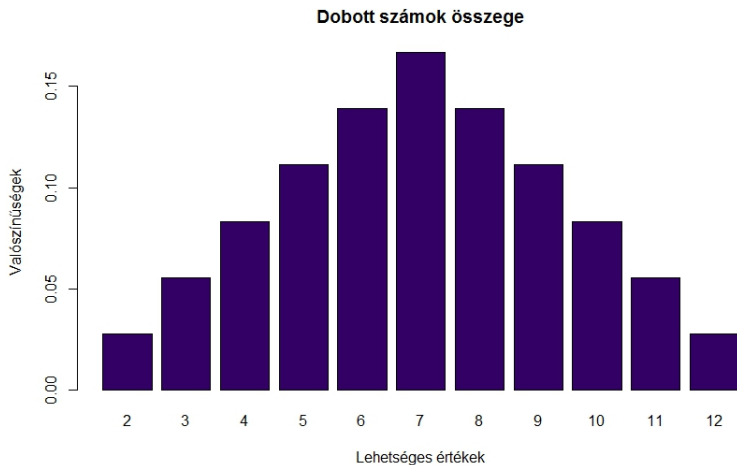
Száz független $\lambda = 1/3$ paraméterű exponenciális eloszlású valószínűségi változó hisztogramja és a sűrűségfüggvény, azaz $e^{-1/3}/3$ ($\mathbb{E}(X) = D(X) = 3, \bar{x} = 3,03, s_n^* = 2,89$)

Exponenciális eloszlások átlaga



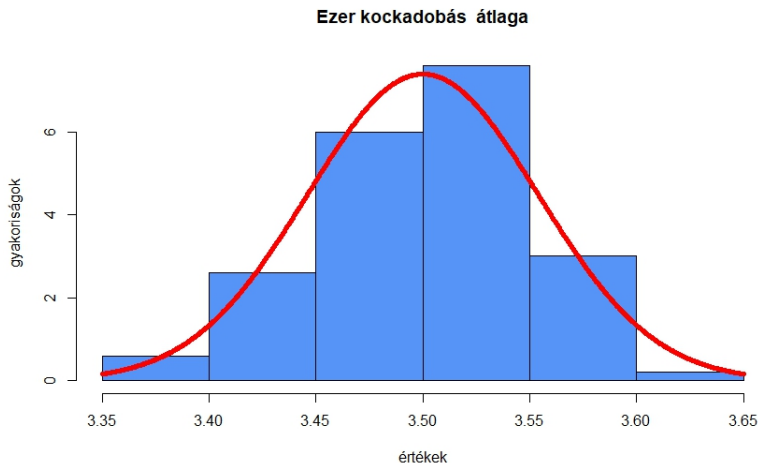
Százelemű minta az alábbi eloszlásból: $n = 1000$ független exponenciális eloszlású ($\lambda = 1/3$) valószínűségi változó átlaga, és az $N(3, 3/\sqrt{1000})$ normális eloszlás sűrűségfüggvénye ($\bar{x} = 2,98$, $s_n^* = 0,098$, $\sigma/\sqrt{n} = 0,095$)

Két kockadobás összege



Két szabályos kockadobás összegének eloszlása

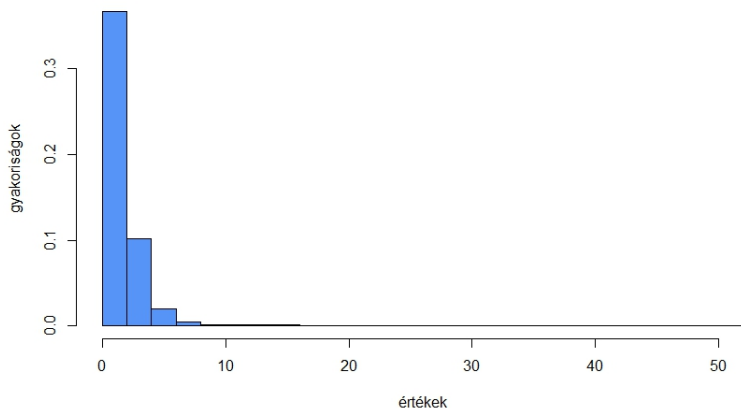
Kockadobások átlaga



Százelemű minta az alábbi eloszlásból: $n = 1000$ független szabályos kockadobás átlaga, és az $N(3, 5, D(X_1)/\sqrt{1000})$ normális eloszlás sűrűségfüggvénye ($\bar{x} = 3,501, s_n^* = 0,098, \sigma/\sqrt{n} = 0,051$)

Exponenciális eloszlás a kitevőben

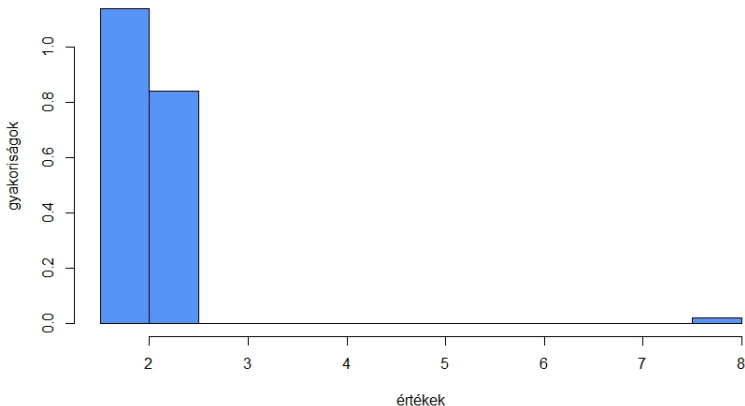
Exponenciális eloszlás a kitevőben



$e^{X_1}, e^{X_2}, \dots, e^{X_{1000}}$ hisztogramja, ahol X_i -k függetlenek, 2 paraméterű exponenciális eloszlásúak ($\mathbb{E}(e^{X_1}) = 2, D(e^{X_1}) = \infty, \bar{x} = 1,99, s_n^* = 2,33$)

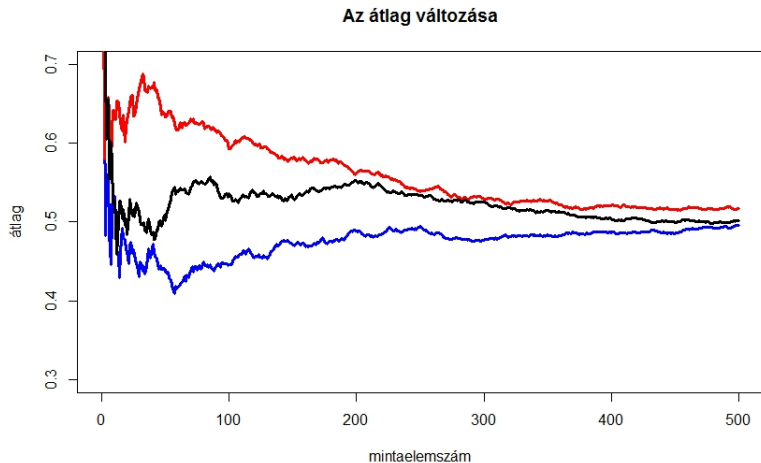
Exponenciális eloszlás a kitevőben

Ezer exponenciális eloszlás átlaga



Százelemű minta az alábbi eloszlásból: $e^{X_1}, e^{X_2}, \dots, e^{X_{1000}}$ átlaga, ahol X_i -k függetlenek, 2 paraméterű exponenciális eloszlásúak. Itt e^{X_i} várható értéke véges, de szórása végtelen.

Az átlag konvergenciája



A $[0, 1]$ intervallumon egyenletes eloszlásból vett minta átlaga $n = 500$ -ig

A nagy számok törvényei

Tétel (A nagy számok gyenge törvénye)

Legyenek X_1, X_2, \dots olyan valószínűségi változók, melyek függetlenek és azonos eloszlásúak. Tegyük fel, hogy $D(X_1) < \infty$. Ekkor minden $\varepsilon > 0$ esetén

$$\mathbb{P}(|\bar{X}_n - \mathbb{E}(X_1)| > \varepsilon) \rightarrow 0 \quad (n \rightarrow \infty),$$

azaz $\bar{X}_n \rightarrow \mathbb{E}(X_1)$ sztochasztikusan.

A nagy számok törvényei

Tétel (A nagy számok gyenge törvénye)

Legyenek X_1, X_2, \dots olyan valószínűségi változók, melyek függetlenek és azonos eloszlásúak. Tegyük fel, hogy $D(X_1) < \infty$. Ekkor minden $\varepsilon > 0$ esetén

$$\mathbb{P}(|\bar{X}_n - \mathbb{E}(X_1)| > \varepsilon) \rightarrow 0 \quad (n \rightarrow \infty),$$

azaz $\bar{X}_n \rightarrow \mathbb{E}(X_1)$ sztochasztikusan.

Tétel (A nagy számok erős törvénye)

Legyenek X_1, X_2, \dots valószínűségi változók, melyek függetlenek és azonos eloszlásúak. Tegyük fel még, hogy $m = \mathbb{E}(X_1) < \infty$. Ekkor

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n} \rightarrow \mathbb{E}(X_1) = m$$

teljesül 1 valószínűséggel $n \rightarrow \infty$ esetén.

A második esetben gyengébb feltevésből erősebb állítás következik.

Centrális határeloszlástétel

Tétel (Centrális határeloszlástétel)

Legyenek X_1, X_2, \dots **független azonos eloszlású** valószínűségi változók, melyekre $\mathbb{E}(X_1) = m$ és $D(X_1) = \sigma < \infty$, azaz **szórásuk véges**. Ekkor tetszőleges t valós számra

$$\mathbb{P}\left(\frac{X_1 + X_2 + \dots + X_n - n \cdot m}{\sigma\sqrt{n}} \leq t\right) \rightarrow \mathbb{P}(Z \leq t) \quad (n \rightarrow \infty),$$

ahol Z standard normális eloszlású, azaz

$$\mathbb{P}(Z \leq t) = \Phi(t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx.$$

Ezt úgy is fogalmazhatjuk, hogy

$$\frac{X_1 + X_2 + \dots + X_n - n \cdot m}{\sigma\sqrt{n}} \rightarrow N(0, 1)$$

teljesül $n \rightarrow \infty$ esetén eloszlásban. Azonos eloszlású: $\mathbb{P}(X_i \leq t) = P(X_j \leq t)$ minden i, j párra és t valós számra

Centrális határeloszlástétel

Legyenek X_1, X_2, \dots független azonos eloszlású valószínűségi változók, melyekre $\mathbb{E}(X_1) = m$ és $D(X_1) = \sigma < \infty$. Ekkor

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(a \leq \frac{X_1 + X_2 + \dots + X_n - n \cdot m}{\sigma \sqrt{n}} < b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

A határértéket $\Phi(b) - \Phi(a) = \mathbb{P}(a \leq Y \leq b)$ alakban is írhatjuk, ahol $Y \sim N(0, 1)$.

Centrális határeloszlástétel

Legyenek X_1, X_2, \dots független azonos eloszlású valószínűségi változók, melyekre $\mathbb{E}(X_1) = m$ és $D(X_1) = \sigma < \infty$. Ekkor

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(a \leq \frac{X_1 + X_2 + \dots + X_n - n \cdot m}{\sigma \sqrt{n}} < b \right) = \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx.$$

A határértéket $\Phi(b) - \Phi(a) = \mathbb{P}(a \leq Y \leq b)$ alakban is írhatjuk, ahol $Y \sim N(0, 1)$.

Így is átfogalmazható a tétel állítása:

$$\mathbb{P}(nm + a\sigma\sqrt{n} \leq X_1 + X_2 + \dots + X_n < nm + b\sigma\sqrt{n}) \rightarrow \Phi(b) - \Phi(a).$$

Ez azt jelenti, hogy az \bar{X}_n átlag eloszlása közel van egy m várható értékű, σ/\sqrt{n} szórású normális eloszláshoz.

Centrális határeloszlástétel

Legyenek X_1, X_2, \dots **független** azonos eloszlású valószínűségi változók, melyekre $\mathbb{E}(X_1) = m$ és $D(X_1) = \sigma < \infty$, azaz **szórásuk véges**. Ekkor

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(a \leq \frac{X_1 + X_2 + \dots + X_n - n \cdot m}{\sigma \sqrt{n}} < b \right) &= \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx = \\ &= \Phi(b) - \Phi(a) = \mathbb{P}(a \leq Z \leq b), \end{aligned}$$

ahol $Z \sim N(0, 1)$ **standard normális eloszlású**. Tovább alakítva:

Centrális határeloszlástétel

Legyenek X_1, X_2, \dots **független** azonos eloszlású valószínűségi változók, melyekre $\mathbb{E}(X_1) = m$ és $D(X_1) = \sigma < \infty$, azaz **szórásuk véges**. Ekkor

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(a \leq \frac{X_1 + X_2 + \dots + X_n - n \cdot m}{\sigma \sqrt{n}} < b \right) &= \frac{1}{\sqrt{2\pi}} \int_a^b e^{-x^2/2} dx = \\ &= \Phi(b) - \Phi(a) = \mathbb{P}(a \leq Z \leq b), \end{aligned}$$

ahol $Z \sim N(0, 1)$ **standard normális eloszlású**. Tovább alakítva:

$$\mathbb{P}(nm + a\sigma\sqrt{n} \leq X_1 + X_2 + \dots + X_n < nm + b\sigma\sqrt{n}) \rightarrow \mathbb{P}(a \leq Z \leq b).$$

Ha n -nel osztunk, hogy az átlag jelenjen meg:

$$\mathbb{P} \left(m + a \frac{\sigma}{\sqrt{n}} \leq \frac{X_1 + X_2 + \dots + X_n}{n} < m + b \frac{\sigma}{\sqrt{n}} \right) \rightarrow \mathbb{P}(a \leq Z \leq b).$$

Vagyis az **átlag eloszlása** „közel van” egy m várható értékű, $\frac{\sigma}{\sqrt{n}}$ szórású **normális eloszláshoz**.

Centrális határeloszlástétel: példa

Legyenek X_1, X_2, \dots független, 2 várható értékű, exponenciális eloszlású valószínűségi változók. Mi a limesze a $\mathbb{P}(X_1 + \dots + X_n - 2n < 2\sqrt{n})$ mennyiségnek $n \rightarrow \infty$ esetén?

Centrális határeloszlástétel: példa

Legyenek X_1, X_2, \dots független, 2 várható értékű, exponenciális eloszlású valószínűségi változók. Mi a limesze a $\mathbb{P}(X_1 + \dots + X_n - 2n < 2\sqrt{n})$ mennyiségnek $n \rightarrow \infty$ esetén?

Mivel a valószínűségi változók **függetlenek**, **azonos eloszlásúak és véges szórásúak**, teljesülnek a centrális határeloszlástétel feltételei. Ezért

$$\mathbb{P}(X_1 + \dots + X_n - 2n < 2\sqrt{n}) = \mathbb{P}\left(\frac{X_1 + \dots + X_n - 2n}{2\sqrt{n}} < 1\right) \rightarrow \Phi(1),$$

ha $n \rightarrow \infty$, hiszen $m = 2$ a várható érték, és mivel az eloszlás exponenciális, a várható érték egyenlő a szórással, így $\sigma = 2$ a szórás.

Konvergenciafajták

Definíció

A Z_1, Z_2, \dots , valószínűségi változókból álló sorozat **eloszlásban konvergál** az Z valószínűségi változóhoz, ha minden olyan t számra, melyre Z eloszlásfüggvénye folytonos t -ben, teljesül, hogy

$$\mathbb{P}(Z_n \leq t) \rightarrow \mathbb{P}(Z \leq t) \quad (n \rightarrow \infty).$$

Konvergenciafajták

Definíció

A Z_1, Z_2, \dots , valószínűségi változókból álló sorozat **eloszlásban konvergál** az Z valószínűségi változóhoz, ha minden olyan t számra, melyre Z eloszlásfüggvénye folytonos t -ben, teljesül, hogy

$$\mathbb{P}(Z_n \leq t) \rightarrow \mathbb{P}(Z \leq t) \quad (n \rightarrow \infty).$$

Ha $Z_n \rightarrow Z$ teljesül 1 valószínűséggel, akkor $Z_n \rightarrow Z$ sztochasztikusan és eloszlásban is.

Lehetséges, hogy $Z_n \rightarrow Z$ eloszlásban, de Z_n nem tart Z -hez sztochasztikusan (és ezért 1 valószínűséggel sem).

A centrális határeloszlástétel alkalmazása

Tegyük fel, hogy egy véletlenszerűen választott ember p valószínűséggel szavaz egy adott pártra. Legalább hány embert kell megkérdeznünk (feltételezve, hogy mindenki a többiektől függetlenül válaszol és igazat mond), hogy annak valószínűsége, hogy a pártra szavazók aránya legfeljebb $0,01$ -gyel tér el p -től, tetszőleges p esetén legalább 95% legyen?

A centrális határeloszlástétel alkalmazása

Tegyük fel, hogy egy véletlenszerűen választott ember p valószínűséggel szavaz egy adott pártra. Legalább hány embert kell megkérdeznünk (feltételezve, hogy mindenki a többiektől függetlenül válaszol és igazat mond), hogy annak valószínűsége, hogy a pártra szavazók aránya legfeljebb 0,01-gyel tér el p -től, tetszőleges p esetén legalább 95% legyen?

n megkérdezett, mindenki p valószínűséggel támogatja a pártot

X : a pártot támogatók száma a megkérdezettek között

Kell:

$$\mathbb{P}\left(\left|\frac{X}{n} - p\right| \leq 0,01\right) \geq 0,95$$

teljesüljön minden $0 \leq p \leq 1$ -re.

A centrális határeloszlástétel alkalmazása

n megkérdezett, mindenki p valószínűséggel támogatja a pártot

X : a pártot támogatók száma a megkérdezettek között,

Kell:

$$\mathbb{P}\left(\left|\frac{X}{n} - p\right| \leq 0,01\right) \geq 0,95$$

teljesüljön minden $0 \leq p \leq 1$ -re, ahol $X = \sum_{j=1}^n X_j$, az X_j -k függetlenek, és

$$\mathbb{P}(X_j = 1) = 1 - \mathbb{P}(X_j = 0) = p; \quad \mathbb{E}(X_j) = p; \quad D(X_j) = \sqrt{p(1-p)}.$$

$$\begin{aligned} \mathbb{P}\left(\left|\frac{X}{n} - p\right| \leq 0,01\right) &= \mathbb{P}\left(\left|\frac{\sum_{j=1}^n X_j - np}{\sqrt{np(1-p)}}\right| \leq \frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right) \\ &\approx 2\Phi\left(\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right) - 1. \end{aligned}$$

A centrális határeloszlástétel alkalmazása

n megkérdezett, mindenki p valószínűséggel támogatja a pártot

Kell:

$$\mathbb{P}\left(\left|\frac{X}{n} - p\right| \leq 0,01\right) \approx 2\Phi\left(\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right) - 1 \geq 0,95$$

teljesüljön minden $0 \leq p \leq 1$ -re. Vagyis mivel $p(1-p) \leq 1/4$:

$$\Phi\left(\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}}\right) \geq 0,975;$$

$$\frac{0,01\sqrt{n}}{\sqrt{p(1-p)}} \geq \Phi^{-1}(0,975) = \text{qnorm}(0,975) = 1,96;$$

$$n \geq p(1-p) \cdot 1,96^2 \cdot \frac{1}{0,01^2};$$

$$n \geq \frac{1}{4} \cdot 1,96^2 \cdot \frac{1}{0,01^2} = 9607.$$

A centrális határeloszlástétel alkalmazása

Tegyük fel, hogy egy véletlenszerűen választott ember p valószínűséggel szavaz egy adott pártra. Legalább hány embert kell megkérdeznünk (feltételezve, hogy mindenki a többiektől függetlenül válaszol és igazat mond), hogy annak valószínűsége, hogy a pártra szavazók aránya legfeljebb $0,01$ -gyel tér el p -től, tetszőleges p esetén legalább 95% legyen?

- Csebisev-egyenlőtlenséggel: $n \geq 50000$ biztosan elég
- centrális határeloszlástétellel közelítve: $n \geq 9607$ elég

A centrális határeloszlástétel alkalmazása

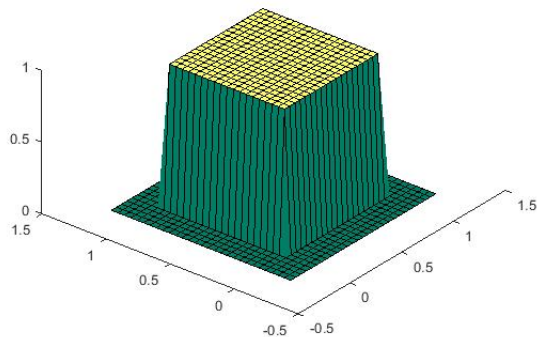
Tegyük fel, hogy egy véletlenszerűen választott ember p valószínűséggel szavaz egy adott pártra. Legalább hány embert kell megkérdeznünk (feltételezve, hogy mindenki a többiektől függetlenül válaszol és igazat mond), hogy annak valószínűsége, hogy a pártra szavazók aránya legfeljebb 0,01-gyel tér el p -től, tetszőleges p esetén legalább 95% legyen?

- Csebisev-egyenlőtlenséggel: $n \geq 50000$ biztosan elég
- centrális határeloszlástétellel közelítve: $n \geq 9607$ elég
- valójában: $n = 9607$, $p = 1/2$ esetén 0,94987 adódik a 0,95 helyett
- valójában $n \geq 9650$ kell (pontos számolással)

Házi feladat december 5., hétfő, 10:15-ig

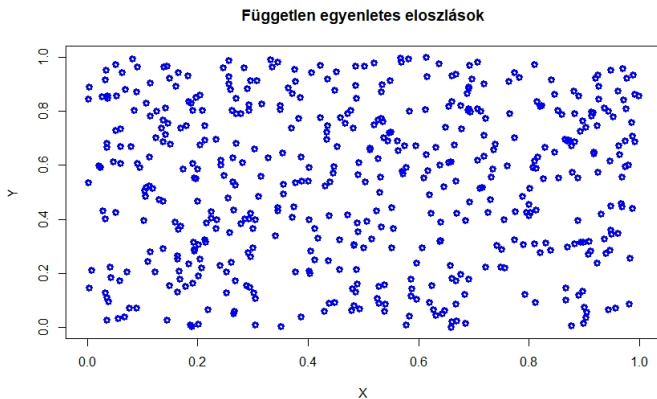
Tegyük fel, hogy egy véletlenszerűen kiválasztott ember jövedelme (ezer forintban számolva) 250 várható értékű és 50 szórású normális eloszlású valószínűségi változó. Függetlenül kiválasztva 200 embert, mennyi a valószínűsége, hogy az átlagos jövedelmük több 240 ezer forintnál?

Kétdimenziós egyenletes eloszlás



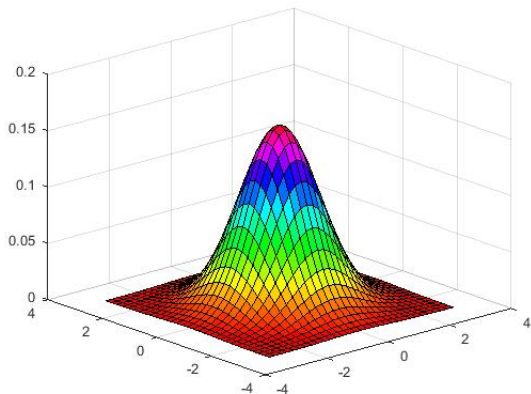
(X, Y) együttes sűrűségfüggvénye, ahol X és Y függetlenek és a $[0, 1]$ intervallumon egyenletes eloszlásúak

Kétdimenziós egyenletes eloszlás



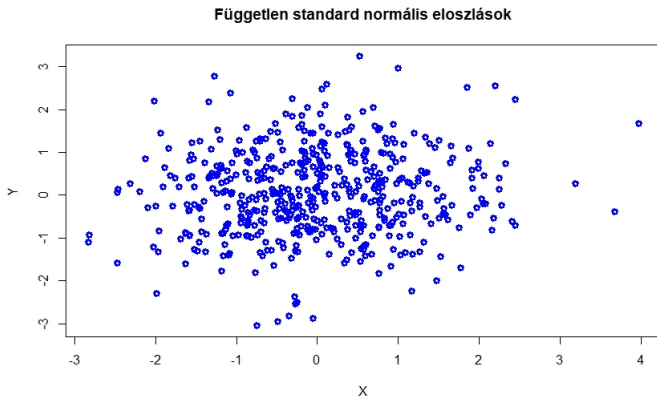
500 darab véletlen pont a síkon, melyek koordinátái függetlenek és a $[0, 1]$ intervallumon egyenletes eloszlásúak (az együttes sűrűségfüggvény az előző ábrán látható).

Kétdimenziós normális eloszlás



Két független standard normális eloszlás együttes sűrűségfüggvénye. Azaz: (X, Z) együttes sűrűségfüggvénye, ahol X, Z függetlenek, $N(0, 1)$ eloszlásúak.

Kétdimenziós normális eloszlás



500 darab véletlen pont a síkon, melyek koordinátái független standard normális eloszlásúak. Ahol nagyobb az együttes sűrűségfüggvény (előző ábra), oda több pont esik.

Együttes eloszlás: példa

Kétszer dobunk szabályos kockával. Legyen X az első dobás, Y pedig a dobott számok közül a nagyobb. Ekkor az (X, Y) valószínűségi vektorváltozó együttes eloszlása:

X/Y	1	2	3	4	5	6	összesen
1	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
2	0	$1/18$	$1/36$	$1/36$	$1/36$	$1/36$	$1/6$
3	0	0	$1/12$	$1/36$	$1/36$	$1/36$	$1/6$
4	0	0	0	$1/9$	$1/36$	$1/36$	$1/6$
5	0	0	0	0	$5/36$	$1/36$	$1/6$
6	0	0	0	0	0	$1/6$	$1/6$
összesen	$1/36$	$1/12$	$5/36$	$7/36$	$1/4$	$11/36$	1

A peremeloszlások:

X : $(1, 1/6), (2, 1/6), (3, 1/6), \dots, (6, 1/6)$

Y : $(1, 1/36), (2, 1/12), \dots, (6, 11/36)$

Peremeloszlások sűrűségfüggvénye

Tegyük fel, hogy az (X_1, X_2, \dots, X_n) valószínűségi vektorváltozó **együttes sűrűségfüggvénye** f . Hogyan kapható meg például az **első peremeloszlás**, azaz X_1 sűrűségfüggvénye?

Peremeloszlások sűrűségfüggvénye

Tegyük fel, hogy az (X_1, X_2, \dots, X_n) valószínűségi vektorváltozó **együttes sűrűségfüggvénye** f . Hogyan kapható meg például az **első peremeloszlás, azaz X_1 sűrűségfüggvénye**?

Az X_j valószínűségi változó sűrűségfüggvénye (melyet f_j -vel jelölünk), azaz a j . peremsűrűségfüggvény így kapható meg f -ből:

$$f_j(t) = \int_{\mathbb{R}^{n-1}} f(s_1, \dots, s_{j-1}, t, s_{j+1}, \dots, s_n) ds_1 \dots ds_{j-1} ds_{j+1} \dots ds_n.$$

Speciálisan $n = 2$ -re:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy; \quad f_2(y) = \int_{-\infty}^{\infty} f(x, y) dx.$$

Feltételes eloszlás

Legyen X diszkrét valószínűségi változó, lehetséges értékei x_1, x_2, \dots , a hozzájuk tartozó valószínűségek: $\mathbb{P}(X = x_k) = p_k$. Legyen A pozitív valószínűségű esemény. Ekkor az X -nek az A eseményre vonatkozó feltételes eloszlása:

$$q_k = \mathbb{P}(X = x_k | A) = \frac{\mathbb{P}(\{X = x_k\} \cap A)}{\mathbb{P}(A)}.$$

Ekkor

$$\sum_{k=1}^{\infty} q_k = \sum_{k=1}^{\infty} \frac{\mathbb{P}(\{X = x_k\} \cap A)}{\mathbb{P}(A)} = 1,$$

azaz a (q_k) sorozat is valószínűségeloszlás.

Másrészt az $\{X = x_k\}$, $k = 1, 2, \dots$ események teljes eseményrendszert alkotnak.

Feltételes eloszlás

Legyen X diszkrét valószínűségi változó, lehetséges értékei x_1, x_2, \dots , a hozzájuk tartozó valószínűségek: $\mathbb{P}(X = x_k) = p_k$. Legyen A pozitív valószínűségű esemény. Ekkor az X -nek az A eseményre vonatkozó feltételes eloszlása:

$$q_k = \mathbb{P}(X = x_k | A) = \frac{\mathbb{P}(\{X = x_k\} \cap A)}{\mathbb{P}(A)}.$$

Ekkor

$$\sum_{k=1}^{\infty} q_k = \sum_{k=1}^{\infty} \frac{\mathbb{P}(\{X = x_k\} \cap A)}{\mathbb{P}(A)} = 1,$$

azaz a (q_k) sorozat is valószínűségeloszlás.

Másrészt az $\{X = x_k\}$, $k = 1, 2, \dots$ események teljes eseményrendszert alkotnak.

Feltételes várható érték:

$$\mathbb{E}(X|A) = \sum_{k=1}^{\infty} k \cdot q_k = \sum_{k=1}^{\infty} k \cdot \mathbb{P}(X = x_k | A).$$

Együttes eloszlás: példa

Kétszer dobunk szabályos kockával. Legyen X az első dobás, Y pedig a dobott számok közül a nagyobb. Ekkor az (X, Y) valószínűségi vektorváltozó együttes eloszlása:

X/Y	1	2	3	4	5	6	összesen
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	0	1/18	1/36	1/36	1/36	1/36	1/6
3	0	0	1/12	1/36	1/36	1/36	1/6
4	0	0	0	1/9	1/36	1/36	1/6
5	0	0	0	0	5/36	1/36	1/6
6	0	0	0	0	0	1/6	1/6
összesen	1/36	1/12	5/36	7/36	1/4	11/36	1

Legyen $A = \{Y = 3\}$ a feltétel.

Ekkor X feltételes eloszlása:

$$q_1 = \mathbb{P}(X = 1|Y = 3) = \frac{1/36}{5/36} = \frac{1}{5} = q_2; \quad q_3 = \frac{1/12}{5/36} = \frac{3}{5}.$$

Együttes eloszlás: példa

Kétszer dobunk szabályos kockával. Legyen X az első dobás, Y pedig a dobott számok közül a nagyobb. Ekkor az (X, Y) valószínűségi vektorváltozó együttes eloszlása:

X/Y	1	2	3	4	5	6	összesen
1	1/36	1/36	1/36	1/36	1/36	1/36	1/6
2	0	1/18	1/36	1/36	1/36	1/36	1/6
3	0	0	1/12	1/36	1/36	1/36	1/6
4	0	0	0	1/9	1/36	1/36	1/6
5	0	0	0	0	5/36	1/36	1/6
6	0	0	0	0	0	1/6	1/6
összesen	1/36	1/12	5/36	7/36	1/4	11/36	1

Legyen $A = \{Y = 3\}$ a feltétel. Ekkor X feltételes várható értéke:

$$\mathbb{E}(X|Y = 3) = \sum_{k=1}^3 \mathbb{P}(X = k|Y = 3) \cdot k = \frac{1}{5} \cdot 1 + \frac{1}{5} \cdot 2 + \frac{3}{5} \cdot 3 = \frac{12}{5} = 2,4.$$

Együttes sűrűségfüggvény: példa

Tegyük fel, hogy az (X, Y) valószínűségi vektorváltozó együttes sűrűségfüggvénye

$$f(x, y) = \begin{cases} x + y, & \text{ha } 0 \leq x \leq 1 \text{ és } 0 \leq y \leq 1; \\ 0 & \text{különben.} \end{cases}$$

Számítsuk ki X és Y korrelációs együtthatóját:

$$R(X, Y) = \frac{\text{cov}(X, Y)}{D(X)D(Y)} = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{D(X)D(Y)}.$$

Együttes sűrűségfüggvény: példa

$$f(x, y) = \begin{cases} x + y, & \text{ha } 0 \leq x \leq 1 \text{ és } 0 \leq y \leq 1; \\ 0 & \text{különben.} \end{cases}$$

Állítás

*Legyen az (X, Y) valószínűségi vektorváltozó együttes sűrűségfüggvénye f .
Ekkor*

$$\mathbb{E}(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dy dx.$$

Együttes sűrűségfüggvény: példa

$$f(x, y) = \begin{cases} x + y, & \text{ha } 0 \leq x \leq 1 \text{ és } 0 \leq y \leq 1; \\ 0 & \text{különben.} \end{cases}$$

Állítás

Legyen az (X, Y) valószínűségi vektorváltozó együttes sűrűségfüggvénye f .
Ekkor

$$\mathbb{E}(g(X, Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dy dx.$$

$$\begin{aligned} \mathbb{E}(XY) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy \cdot f(x, y) dy dx = \int_0^1 \int_0^1 xy(x + y) dy dx = \\ &= \int_0^1 \int_0^1 x^2 y dy dx + \int_0^1 \int_0^1 xy^2 dy dx = \int_0^1 \frac{x^2}{2} dx + \int_0^1 \frac{x}{3} dx = \frac{1}{6} + \end{aligned}$$

felhasználva, hogy $\int_0^1 x^k dx = [x^{k+1}/(k+1)]_{x=0}^1 = 1/(k+1)$.

Együttes sűrűségfüggvény: példa

Az (X, Y) együttes sűrűségfüggvénye:

$$f(x, y) = \begin{cases} x + y, & \text{ha } 0 \leq x \leq 1 \text{ és } 0 \leq y \leq 1; \\ 0 & \text{különben.} \end{cases}$$

Az X valószínűségi változó sűrűségfüggvénye:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 x + y dy = x + \frac{1}{2},$$

ha $0 \leq x \leq 1$, és 0 különben. Ezért

Együttes sűrűségfüggvény: példa

Az (X, Y) együttes sűrűségfüggvénye:

$$f(x, y) = \begin{cases} x + y, & \text{ha } 0 \leq x \leq 1 \text{ és } 0 \leq y \leq 1; \\ 0 & \text{különben.} \end{cases}$$

Az X valószínűségi változó sűrűségfüggvénye:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 x + y dy = x + \frac{1}{2},$$

ha $0 \leq x \leq 1$, és 0 különben. Ezért

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_1(x) dx = \int_0^1 x \cdot \left(x + \frac{1}{2}\right) dx = \int_0^1 x^2 + \frac{x}{2} dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}.$$

Együttes sűrűségfüggvény: példa

Az X valószínűségi változó sűrűségfüggvénye:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 x + y dy = x + \frac{1}{2},$$

ha $0 \leq x \leq 1$, és 0 különben. Ezért

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_1(x) dx = \int_0^1 x \cdot \left(x + \frac{1}{2}\right) dx = \int_0^1 x^2 + \frac{x}{2} dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}.$$

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_1(x) dx = \int_0^1 x^2 \cdot \left(x + \frac{1}{2}\right) dx = \int_0^1 x^3 + \frac{x^2}{2} dx = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}.$$

Együttes sűrűségfüggvény: példa

Az X valószínűségi változó sűrűségfüggvénye:

$$f_1(x) = \int_{-\infty}^{\infty} f(x, y) dy = \int_0^1 x + y dy = x + \frac{1}{2},$$

ha $0 \leq x \leq 1$, és 0 különben. Ezért

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f_1(x) dx = \int_0^1 x \cdot \left(x + \frac{1}{2}\right) dx = \int_0^1 x^2 + \frac{x}{2} dx = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}.$$

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 f_1(x) dx = \int_0^1 x^2 \cdot \left(x + \frac{1}{2}\right) dx = \int_0^1 x^3 + \frac{x^2}{2} dx = \frac{1}{4} + \frac{1}{6} = \frac{5}{12}.$$

A szimmetria miatt hasonlóképpen:

$$\mathbb{E}(Y) = \frac{7}{12}; \quad \mathbb{E}(Y^2) = \frac{5}{12}.$$

Együttes sűrűségfüggvény: példa

$$f(x, y) = \begin{cases} x + y, & \text{ha } 0 \leq x \leq 1 \text{ és } 0 \leq y \leq 1; \\ 0 & \text{különben.} \end{cases}$$

$$\mathbb{E}(XY) = \frac{1}{3}; \quad \mathbb{E}(X) = \mathbb{E}(Y) = \frac{7}{12}; \quad \mathbb{E}(X^2) = \mathbb{E}(Y^2) = \frac{5}{12}.$$

Az X és Y korrelációs együtthatója:

$$\begin{aligned} R(X, Y) &= \frac{\text{cov}(X, Y)}{D(X)D(Y)} = \frac{\mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)}{D(X)D(Y)} = \\ &= \frac{1/3 - (7/12)^2}{\sqrt{5/12 - (7/12)^2} \cdot \sqrt{5/12 - (7/12)^2}} = \frac{1/3 - (7/12)^2}{5/12 - (7/12)^2} = -0,0 \end{aligned}$$

Nagyon gyenge negatív korreláció van a két valószínűségi változó között.