

## Nevezetes abszolút folytonos eloszlások

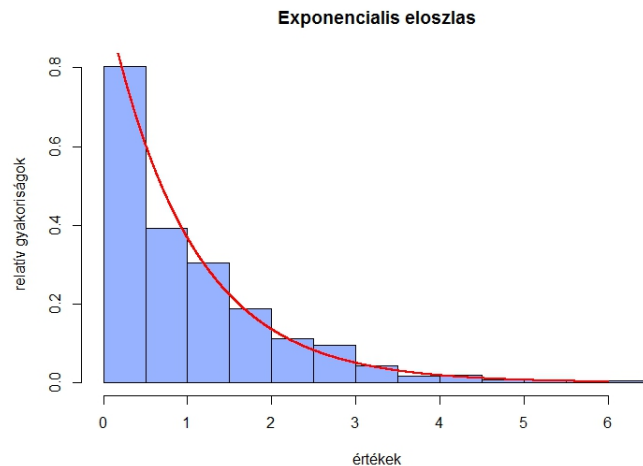
Ahogy korábban láttuk, a sűrűségfüggvény tetszőleges olyan függvény lehet, mely nemnegatív, és az integrálja 1 (pontosabban, ha egy függvény ilyen tulajdonságú, akkor van olyan valószínűségi változó, aminek ez a sűrűségfüggvénye). Vannak azonban olyan sűrűségfüggvények, melyek a statisztikában vagy matematikai modellezésben fontosabb szerepet játszanak, gyakran előjönnek. Ezeket érdemes külön megismerni. Általában jellemző lesz, hogy nem is egy-egy konkrét sűrűségfüggvényről, hanem sűrűségfüggvények egy családjáról van szó, ahol a pontos függvény egy vagy két paraméter megadása után kapható meg.

### 1. Exponenciális eloszlás

Az **exponenciális eloszlás** sokszor használható véletlen időtartamok modellezésére, például

- egy művelet elvégzésének ideje: egy ember kiszolgálása egy boltban, vagy egy számítás elvégzése egy számítógépen
- egy ember reakcióideje
- két esemény bekövetkezése között eltelt idő, például egy üzletben két ügyfél érkezése közötti idő
- járványterjedés modellezésénél: a fertőzés átadásának vagy a gyógyulásnak az ideje
- radioaktív részecske bomlási ideje

Ahogy korábban láttuk, ez az eloszlás örökifjú tulajdonságú, azaz akkor jó modellezésre, ha az, hogy  $t$  ideje várunk az esemény bekövetkezésére, nem változtat a még hátralévő várakozási idő eloszlásán.

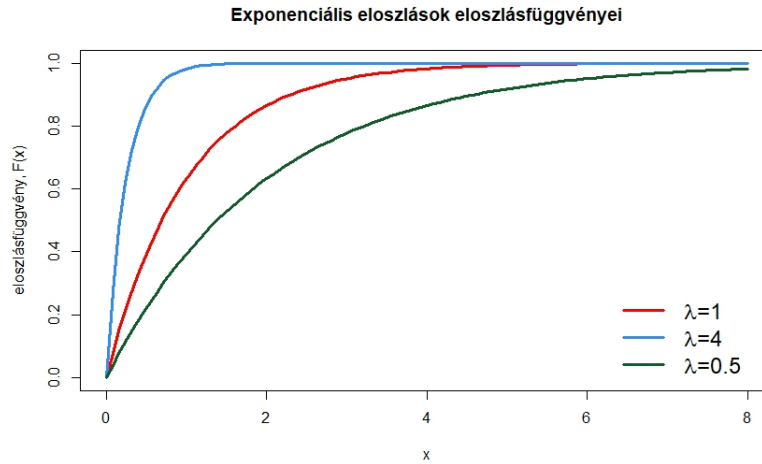


1. ábra.  $\lambda = 1$  paraméterű exponenciális eloszlás sűrűségfüggvénye és 500 darab független, 1 paraméterű exponenciális eloszlású valószínűségi változóból álló minta hisztogramja

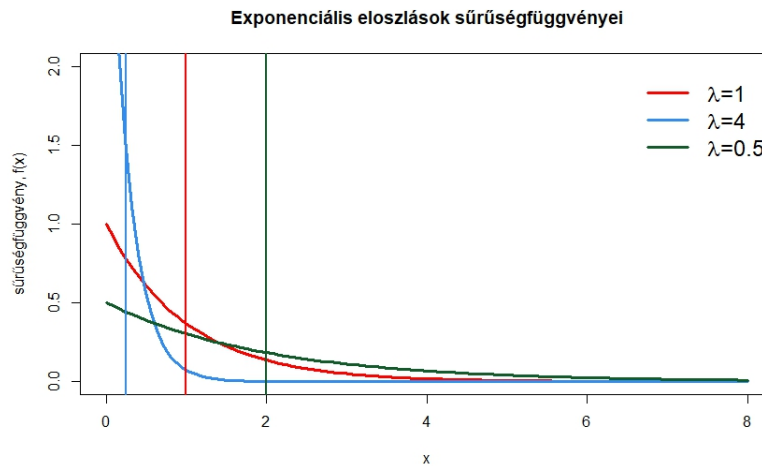
**1.1. Definíció.** Legyen  $\lambda > 0$  valós szám. Az  $X$  valószínűségi változó **exponenciális eloszlású**  $\lambda$  paraméterrel, ha eloszlásfüggvénye:

$$F(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^t f(x) dx = \begin{cases} 1 - e^{-\lambda t}, & \text{ha } t > 0; \\ 0 & \text{különben.} \end{cases}$$

**1.1. Állítás.** Tegyük fel, hogy az  $X$  valószínűségi változó exponenciális eloszlású  $\lambda > 0$  paraméterrel.



2. ábra. Különböző paraméterű ( $\lambda = \frac{1}{2}, 1$ , illetve  $4$ ) exponenciális eloszlások eloszlásfüggvényei



3. ábra. Különböző paraméterű ( $\lambda = \frac{1}{2}, 1$ , illetve  $4$ ) exponenciális eloszlások sűrűségfüggvényei és a várható értékek:  $\mathbb{E}(X) = \frac{1}{\lambda} = 2, 1$  illetve  $\frac{1}{4}$

(i)  $X$  sűrűségfüggvénye:

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

(ii)  $X$  várható értéke:  $\mathbb{E}(X) = \frac{1}{\lambda}$ , szórása:  $D(X) = \frac{1}{\lambda}$ .

(iii) **Örökifjú tulajdonság.** Legyenek  $s, t$  pozitív számok. Ekkor

$$\mathbb{P}(X \geq s + t | X \geq s) = \mathbb{P}(X \geq t).$$

*Bizonyítás.* (i) Korábban láttuk, hogy adott eloszlásfüggvény esetén a sűrűségfüggvény, ha létezik (nem mindig létezik), deriválással kapható. A definícióban megadott függvény deriváltja éppen  $f$  lesz. Mivel viszont nem mindig létezik a sűrűségfüggvény, ellenőrizni kell, hogy ez valóban sűrűségfüggvény:

$$\int_{-\infty}^t f(x) dx = \int_0^t \lambda e^{-\lambda x} dx = [-e^{-\lambda x}]_0^t = -e^{-\lambda t} - (-1) = F(t),$$

ha  $t > 0$ , és 0 különben.

(ii) A várható értékhez az alábbi integrált kell kiszámítani, amit parciális integrálással tehetünk meg:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} xf(x) dx = \int_0^{\infty} x\lambda e^{-\lambda x} dx = [-xe^{-\lambda x}]_{x=0}^{\infty} + \int_0^{\infty} e^{-\lambda x} dx = \left[ -\frac{e^{-\lambda x}}{\lambda} \right]_{x=0}^{\infty} = \frac{1}{\lambda}.$$

A szórás hasonlóképpen számítható ki a definícióból.

(iii) Ezt már korábban bizonyítottuk. □

**Példa.** Tegyük fel, hogy egy boltban egy vevő kiszolgálásának ideje (percben számolva) **3 várható értékű exponenciális eloszlású** valószínűségi változó.

Mennyi a valószínűsége, hogy a vevőt **legalább 5 percig** tart kiszolgálni? Mennyi a valószínűsége, hogy a vevő kiszolgálása **legalább 2, de legfeljebb 4 percig** tart?

Legyen  $X$  a kiszolgálás ideje. Exponenciális eloszlás esetén

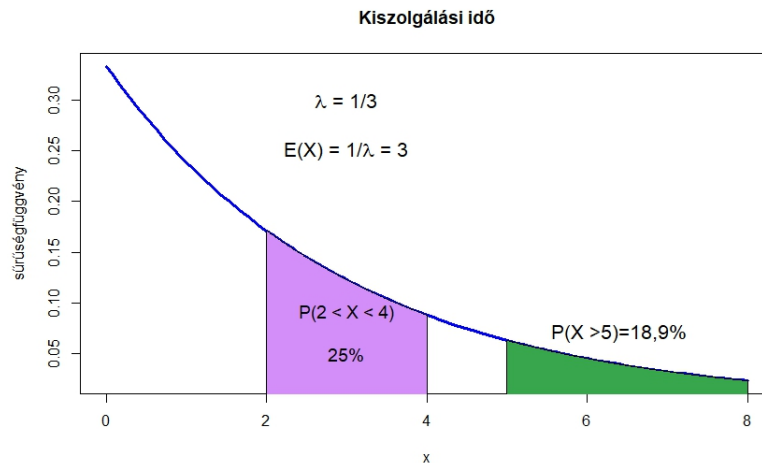
$$\mathbb{E}(X) = 1/\lambda, \text{ és most } \mathbb{E}(X) = 3 \Rightarrow \lambda = 1/3.$$

Ezért az eloszlásfüggvény alapján

$$\mathbb{P}(X \geq 5) = 1 - \mathbb{P}(X < 5) = 1 - F(5) = 1 - (1 - e^{-\lambda \cdot 5}) = e^{-\lambda \cdot 5} = e^{-5/3} = \mathbf{18,9\%}.$$

Hasonlóképpen

$$\mathbb{P}(2 \leq X \leq 4) = \mathbb{P}(X \leq 4) - \mathbb{P}(X \leq 2) = F(4) - F(2) = (1 - e^{-4/3}) - (1 - e^{-2/3}) = e^{-2/3} - e^{-4/3} = \mathbf{25\%}.$$



4. ábra. Annak valószínűsége, hogy a kiszolgálás legalább 5 percig tart, illetve hogy legalább 2, de legalább 4 percig, ha a kiszolgálási idő 3 várható értékű exponenciális eloszlású valószínűségi változó

## 2. Pareto-eloszlás

Biztosításmatematikában, vagy például a jövedelmek eloszlásának modellezésére használják gyakran az alábbi eloszlást.

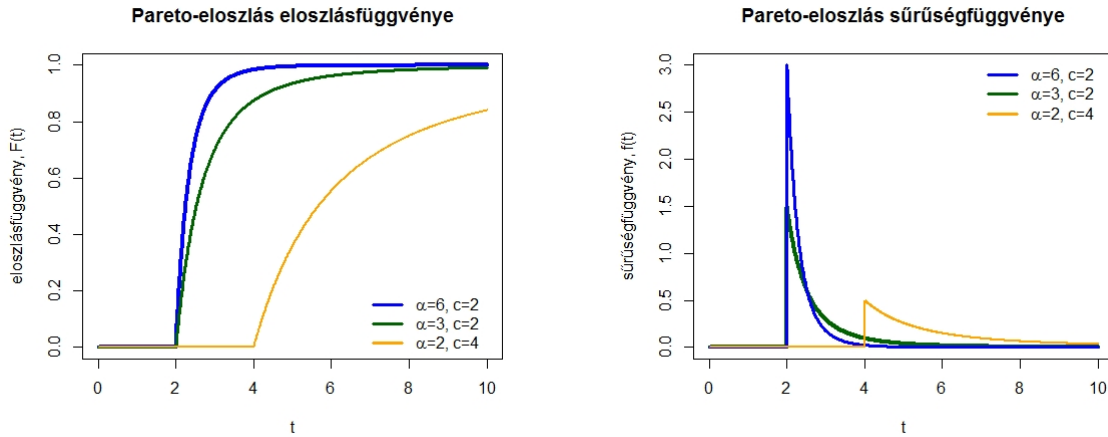
**2.1. Definíció (Pareto-eloszlás).** Az  $X$  valószínűségi változó (első típusú) Pareto-eloszlású  $\alpha > 0$  és  $c > 0$  paraméterekkel, ha eloszlásfüggvénye

$$F(t) = \mathbb{P}(X \leq t) = \begin{cases} 0; & \text{ha } t \leq c; \\ 1 - \left(\frac{c}{t}\right)^\alpha; & \text{ha } t > c. \end{cases}$$

**2.1. Állítás.** Ha  $X$  Pareto-eloszlású  $\alpha > 0$  és  $c > 0$  paraméterekkel, akkor sűrűségfüggvénye

$$f(t) = \begin{cases} 0; & \text{ha } t \leq c; \\ \frac{c^\alpha}{t^{\alpha+1}}; & \text{ha } t > c. \end{cases}$$

Itt a sűrűségfüggvényt deriválással kapjuk, és ellenőrizhető, hogy ez valóban sűrűségfüggvény lesz.



5. ábra. Különböző paraméterű Pareto-eloszlások eloszlás- és sűrűségfüggvényei

Az  $X$  valószínűségi változó  $k$ . momentuma:

$$\mathbb{E}(X^k) = \int_{-\infty}^{\infty} x^k \cdot f(x) dx = \int_c^{\infty} c^\alpha \alpha \cdot x^{k-\alpha-1} dx < \infty \Leftrightarrow k - \alpha - 1 < -1 \Leftrightarrow k < \alpha.$$

Tehát a Pareto-eloszlásnak csak  $\alpha$ -nál kisebb  $k$ -ra véges a  $k$ . momentuma. Például  $k = 2$ -vel azt kapjuk, hogy a szórás pontosan akkor véges, ha  $\alpha > 2$ .

Például ha  $\alpha = 1, 5$ , akkor a várható érték létezik és véges, de szórás nem létezik. Ilyenkor a sűrűségfüggvény, például  $c = 1$ -re:  $1,5x^{-2,5}\mathbb{I}(x > 1)$ . Ennek tehát a **várható értéke véges**, a **szórása végtelen**. Hasonlóképpen olyan Pareto-eloszlásokat is találhatunk, melyeknek a  $k$ . momentuma létezik, de a  $k + 1$ . nem.

Ez az oka annak, hogy, mivel a Pareto-eloszlás „nagy valószínűséggel vesz fel nagy értékeket” is, például tűzkár vagy árvíz kár nagyságának modellezésére használják.

### 3. Normális eloszlás

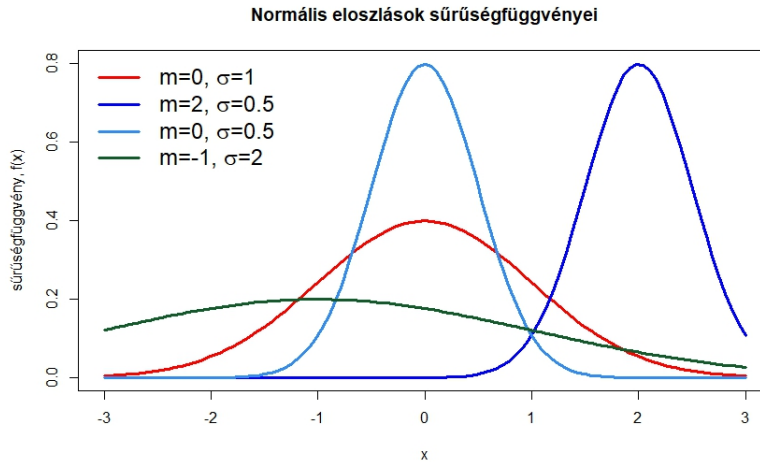
A **normális eloszlás**, melynek sűrűségfüggvénye az  $e^{-x^2}$  függvényből származtatható, gyakran előfordul például különféle *mérési eredmények* (a mérési hibák következtében), illetve élőlények *biológiai jellemzői* gyakran normális eloszlást követnek (például: testmagasság). Ahogy később látni fogjuk, ez tipikusan akkor fordul elő, ha a vizsgált mennyiség sok, egymástól független, összeadó hatás eredményeképpen áll elő. A normális eloszlás a **statisztikában** is kulcsfontosságú, ezt használjuk a  $z$ -próbában, közvetve pedig a  $t$ -próbában is, más eljárásokban is.

**3.1. Definíció.** Legyen  $m$  valós,  $\sigma$  pedig pozitív szám. Azt mondjuk, hogy az  $Y$  valószínűségi változó **normális eloszlású**  $m$  várható értékkel és  $\sigma^2$  szórnégyzettel, ha **sűrűségfüggvénye**

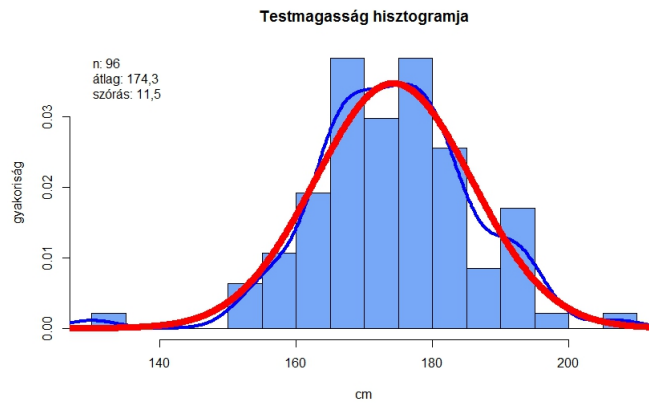
$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right) \quad (x \in \mathbb{R}).$$

Jelölése:  $Y \sim N(m, \sigma^2)$ .

Ellenőrizhető, hogy  $f$  valóban sűrűségfüggvény, azaz nemnegatív és 1 az integrálja.



6. ábra. Különböző várható értékű ( $m$ ) és szórású ( $\sigma$ ) normális eloszlások sűrűségfüggvényei



7. ábra. Testmagasság hisztogramja  $n = 96$  elemű mintából (valós adatokból), és az  $m = \bar{X} = 174,3$  várható értékű és  $\sigma = 11,5$  szórású normális eloszlás sűrűségfüggvénye (pirossal):  $f(x) = \frac{1}{\sqrt{2\pi} \cdot 11,5} \exp(-(x - 174,3)^2 / (2 \cdot 11,5^2))$

**3.1. Állítás (A normális eloszlás várható értéke és szórása).** Ha  $Y \sim N(m, \sigma^2)$ , akkor  $\mathbb{E}(Y) = m$ ,  $\mathbf{D}(Y) = \sigma$ .

Nem bizonyítjuk, de a várható értéket a szimmetria alapján, a szórást parciális integrálás segítségével lehetne levezetni a definícióból.

**3.2. Definíció. Standard normális eloszlás:** az  $m = 0$  várható értékű és  $\sigma = 1$  szórású normális eloszlás. **Eloszlásfüggvénye:**  $\Phi$ , **sűrűségfüggvénye**  $\varphi$ , ahol

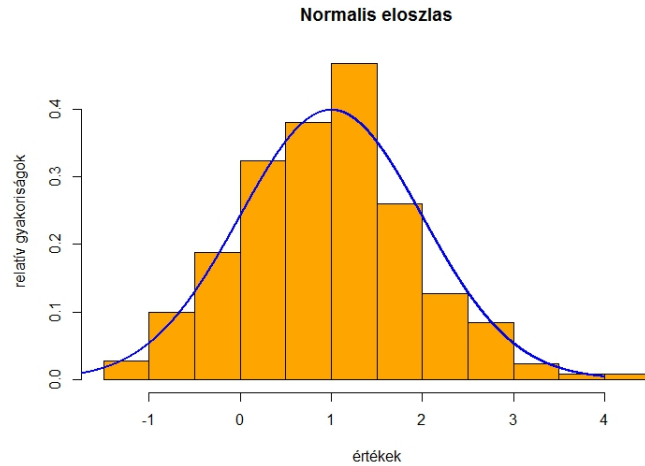
$$\Phi(t) = \int_{-\infty}^t \varphi(x) dx; \quad \varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Másképpen, ha  $Z \sim N(0, 1)$ , azaz  $Z$  standard normális eloszlású, akkor

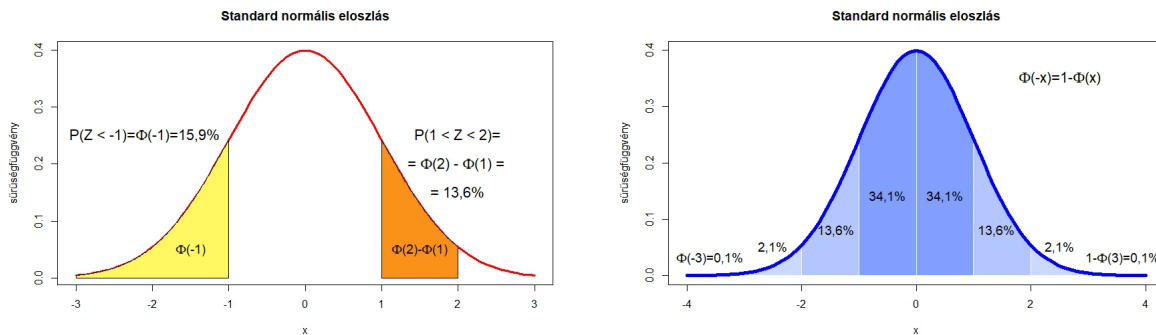
$$\mathbb{P}(Z \leq t) = \Phi(t) = \int_{-\infty}^t \varphi(x) dx = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx.$$

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

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



8. ábra. Normális eloszlás ( $m = 1, \sigma = 1$ ) sűrűségfüggvénye és 500 darab független,  $N(1, 1)$  eloszlású valószínűségi változóból álló minta hisztogramja



9. ábra. A standard normális eloszlás sűrűségfüggvénye

Továbbá

$$\begin{aligned} \mathbb{P}(a < X < b) &= \mathbb{P}(a \leq X \leq b) = \mathbb{P}(X \leq b) - \mathbb{P}(X < a) = \\ &= \int_a^b \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) dx = \Phi(b) - \Phi(a). \end{aligned}$$

Más normális eloszlások esetén ezeket a valószínűségeket a  $\Phi$  függvényre vezethetjük vissza.

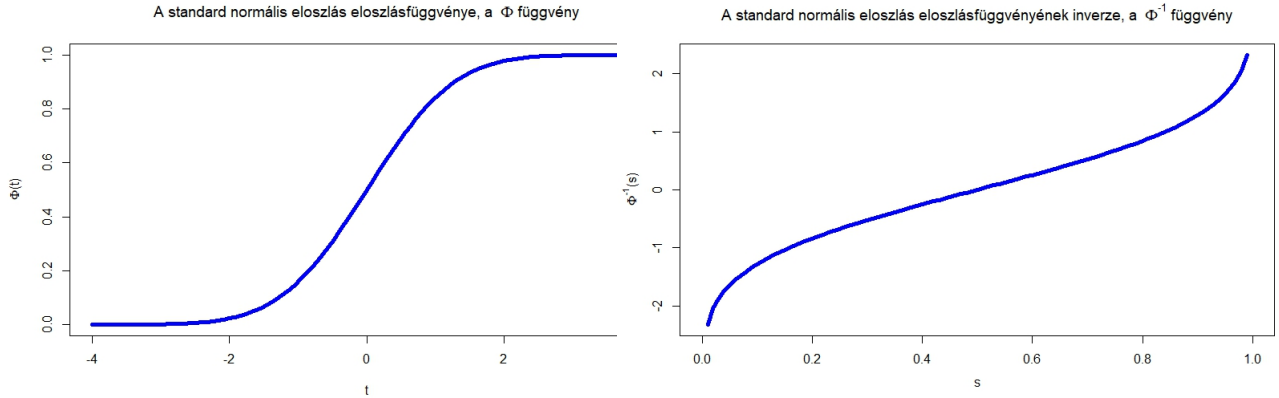
### 3.1. A normális eloszlás tulajdonságai

Tegyük fel, hogy  $Y$  normális eloszlású  $m$  várható értékkel és  $\sigma^2$  szórásnégyzettel, azaz  $Y \sim N(m, \sigma^2)$ . Ekkor tetszőleges  $a \leq b$  valós számokra

$$\mathbb{P}(Y < b) = \mathbb{P}(Y \leq b) = \mathbb{P}\left(\frac{Y - m}{\sigma} \leq \frac{b - m}{\sigma}\right) = \Phi\left(\frac{b - m}{\sigma}\right)$$

$$\mathbb{P}(Y > a) = \mathbb{P}(Y \geq a) = 1 - \Phi\left(\frac{a - m}{\sigma}\right)$$

$$\mathbb{P}(a < Y < b) = \mathbb{P}(a \leq Y \leq b) = \Phi\left(\frac{b - m}{\sigma}\right) - \Phi\left(\frac{a - m}{\sigma}\right)$$



10. ábra. A standard normális eloszlás eloszlásfüggvénye és annak inverze

$$\mathbb{P}(|Y - m| \leq b) = 2\Phi\left(\frac{b}{\sigma}\right) - 1$$

**Lineáris transzformáció.** Legyen  $Y$  normális eloszlású valószínűségi változó  $m$  várható értékkel és  $\sigma$  szórással, és  $a, b$  valós számok. Ekkor az  $aY + b$  valószínűségi változó normális eloszlású  $am + b$  várható értékkel és  $a^2\sigma^2$  szórásnégyzettel, azaz

$$Y \sim N(m, \sigma^2) \quad \Rightarrow \quad aY + b \sim N(am + b, a^2\sigma^2).$$

**Független összeg.** Ha  $Y_1, Y_2$  **független, normális eloszlású** valószínűségi változók, akkor  $Y_1 + Y_2$  is **normális eloszlású**, várható értéke  $m_1 + m_2$ , szórásnégyzete  $\sigma_1^2 + \sigma_2^2$ , ahol  $Y_1 \sim N(m_1, \sigma_1^2)$  és  $Y_2 \sim N(m_2, \sigma_2^2)$

**Példa.** Ha  $Y$  és  $Z$  függetlenek, normális eloszlásúak,  $Y \sim N(2, 3^2)$  és  $Z \sim N(1, 4^2)$ , akkor

$$Y + Z \sim N(3, 5^2); \quad Y - Z \sim N(1, 5^2); \quad Y + 3Z \sim N(5, 57).$$

Legyenek  $Y_1, Y_2, \dots, Y_n$  **független normális eloszlású** valószínűségi változók, melyek várható értéke  $m$ , szóráruk  $\sigma$ . Ekkor az **összegük** és **az átlaguk is normális eloszlású**, és

$$Y_1 + Y_2 + \dots + Y_n \sim N(nm, n\sigma^2);$$

$$\frac{Y_1 + Y_2 + \dots + Y_n}{n} \sim N\left(m, \frac{\sigma^2}{n}\right).$$

**Példa.** Tegyük fel, hogy az emberek testmagassága 176 cm várható értékű és 7 szórással valószínűségi változó. Ekkor

- 100 ember testmagasságának átlaga szintén normális eloszlású, 176 várható értékkel és  $7/\sqrt{100} = 0,7$  szórással;
- 10000 ember testmagasságának átlaga normális eloszlású, 176 várható értékkel és  $7/\sqrt{10000} = 0,07$  szórással.

**Példa.** Tegyük fel, hogy a holnap várható középhőmérséklet,  $Y$  valószínűségi változó normális eloszlású  $m = 4$  várható értékkel és  $\sigma = 3$  szórással. Mennyi a valószínűsége, hogy a holnapi középhőmérséklet értéke legfeljebb 7 fok?

$$\mathbb{P}(Y \leq 7) = \Phi\left(\frac{7 - m}{\sigma}\right) = \Phi\left(\frac{7 - 4}{3}\right) = \Phi(1) = \mathbf{84,1\%}.$$

Az R-ben ugyanezt a `pnorm(7, mean=4, sd=3)` paranccsal érhetjük el. A  $\Phi$  függvény az R-ben a `pnorm`.

Mennyi a valószínűsége, hogy a középhőmérséklet 1 és 7 fok közé esik?

$$\begin{aligned}\mathbb{P}(1 < Y \leq 7) &= \mathbb{P}(Y \leq 7) - \mathbb{P}(Y \leq 1) = \Phi\left(\frac{7-m}{\sigma}\right) - \Phi\left(\frac{1-m}{\sigma}\right) \\ &= \Phi\left(\frac{7-4}{3}\right) - \Phi\left(\frac{1-4}{3}\right) = \Phi(1) - \Phi(-1) = \\ &= 2\Phi(1) - 1 = \mathbf{68,2\%},\end{aligned}$$

mert

$$\Phi(-x) = 1 - \Phi(x)$$

minden valós  $x$ -re érvényes a sűrűségfüggvény 0-ra való szimmetriája miatt.

**Kvantilisek.** Emlékeztetőül: egy (folytonos) eloszlás  $z$ -kvantilise az a  $t$  szám, melyre  $\mathbb{P}(X \leq t) = z$ , vagyis az a szám, melynél a valószínűségi változó értéke pontosan  $z$  valószínűséggel kisebb.

A fentiek alapján a normális eloszlásnál ezt is kiszámíthatjuk, és ennek van a statisztikában gyakran fontos szerepe. Ha tehát  $X \sim N(m, \sigma^2)$ , akkor

$$\mathbb{P}(X \leq t) = \Phi\left(\frac{t-m}{\sigma}\right) = z \Rightarrow \frac{t-m}{\sigma} = \Phi^{-1}(z) \Rightarrow t = \sigma\Phi^{-1}(z) + m.$$

Speciálisan a standard normális eloszlásnál  $m = 0$ ,  $\sigma = 1$ , a kvantilis  $\Phi^{-1}(z)$  lesz, ami nem más, mint a  $z$  megbízhatósági szintű, vagyis  $1 - z$  szignifikanciaszintű egyoldali  $z$ -próba kritikus értéke.

Kiszámítás az R-ben: `qnorm(0.9, mean=4, sd=3)` a 90%-os kvantilis adott várható érték és szórás esetén. Ennek értéke: 7,84.

A fenti példában tehát ez azt jelenti, hogy a középhőmérséklet értéke 90% valószínűséggel lesz 7,84 foknál kisebb.

## 4. További nevezetes abszolút folytonos eloszlások

Az alábbi eloszlások többek között **statisztikai alkalmazásokban** fordulnak elő:

- $t$ -eloszlás:  $t$ -próba, például két eloszlás **várható értékének** összehasonlítására
- $F$ -eloszlás:  $F$ -próba, például két eloszlás **szórásának** összehasonlítására
- $\chi^2$ -eloszlás:  $\chi^2$ -próba, például annak eldöntésére, hogy két tulajdonság között van-e **összefüggés**
- gamma-eloszlás: nemnegatív valószínűségi változók modellezésére
- beta-eloszlás:  $[0, 1]$ -értékű valószínűségi változók modellezésére

### 4.1. $t$ -eloszlás

Legyenek  $X_1, X_2, \dots, X_f$  és  $Y$  független standard normális eloszlású valószínűségi változók. Ekkor a

$$Z = \frac{Y}{\sqrt{(X_1^2 + X_2^2 + \dots + X_f^2)/f}}$$

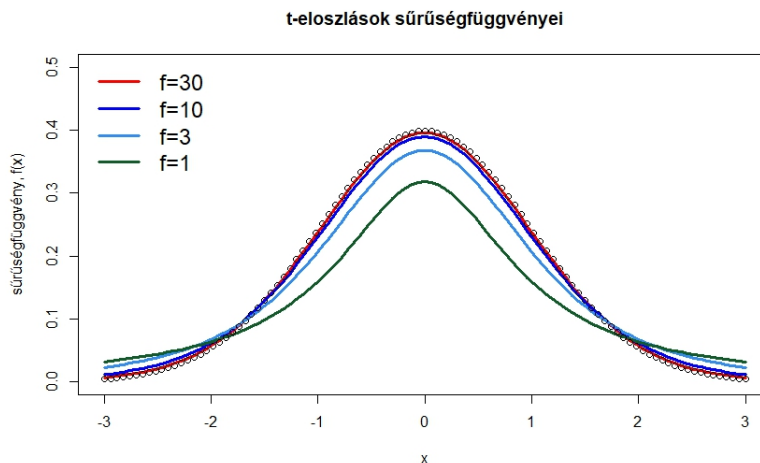
valószínűségi változó eloszlását  $f$  szabadsági fokú  **$t$ -eloszlásnak** (vagy Student-eloszlásnak) nevezzük.

Az  $f = 1$  szabadsági fokú  $t$ -eloszlás, vagyis  $Y/X$  eloszlása a **Cauchy-eloszlás**. Ennek sűrűségfüggvénye:

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

A Cauchy-eloszlásnak sem várható értéke, sem szórása nem létezik:  $\int_{-\infty}^{\infty} x \cdot f(x) dx$  nem értelmezhető, mert  $\int_0^{\infty} \frac{1}{x} dx$  integrál nem véges. Hasonlóképpen a többi  $t$ -eloszlásnak sem létezik a várható értéke.

Ha  $f$  nagy, akkor az  $f$  szabadsági fokú  $t$ -eloszlás közel esik a standard normális eloszláshoz,  $f > 30$  esetén a  $t$ -eloszlást gyakran a normális eloszlással helyettesítik a számolásokban.



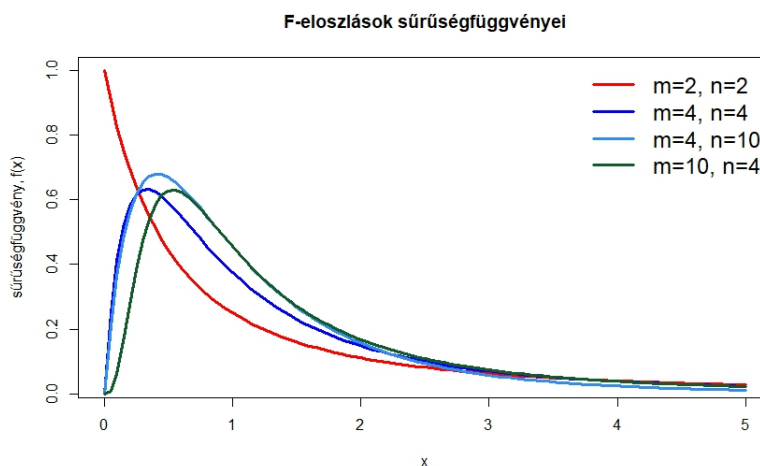
11. ábra. Különböző szabadsági fokú  $t$ -eloszlások sűrűségfüggvényei. A pöttyözött vonal a standard normális eloszlás sűrűségfüggvényét jelöli, ez közel van a  $t$ -eloszlás sűrűségfüggvényéhez, ha  $f$  nagy.

## 4.2. $F$ -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** nevezzük.



12. ábra. Különböző szabadsági fokú  $F$ -eloszlások sűrűségfüggvényei

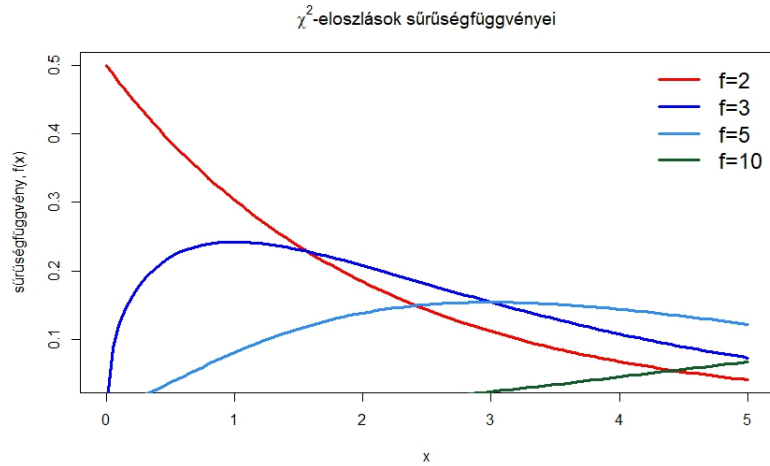
## 4.3. $\chi^2$ -eloszlás

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  **$\chi^2$ -eloszlásnak** nevezzük. Ennek sűrűségfüggvénye:

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



13. ábra. Különböző szabadsági fokú  $\chi^2$ -eloszlások sűrűségfüggvényei

#### 4.4. Gamma-eloszlás

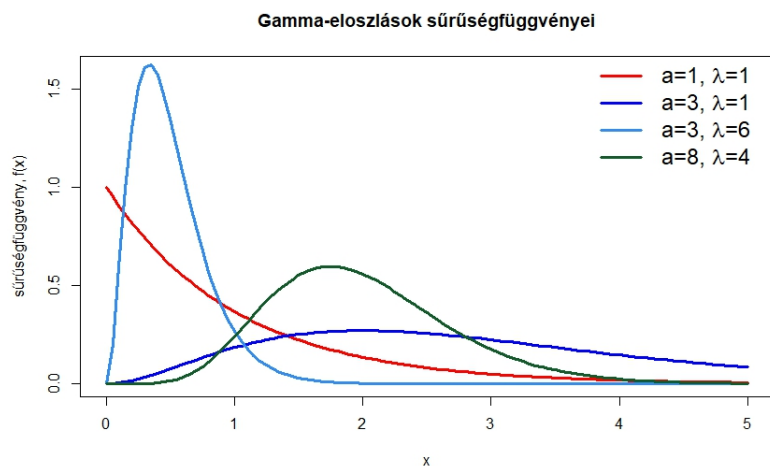
**Gamma-függvény.** Ha  $a > 0$  pozitív szám, legyen

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

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.

**4.1. Definíció.** Legyenek  $a$  és  $\lambda$  pozitív számok. Az  $X$  valószínűségi változó **gamma-eloszlású** a  $a$  renddel és  $\lambda$  paraméterrel, ha sűrűségfüggvénye

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



14. ábra. Különböző paraméterű gamma-eloszlások sűrűségfüggvényei

A gamma-eloszlás tulajdonságai:

- **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  $\lambda$  paraméterű exponenciális eloszlású valószínűségi változók, akkor  $X_1 + X_2 + \dots + X_n$  gamma-eloszlású  $a = n$  renddel és  $\lambda$  paraméterrel.
- **Kapcsolat a  $\chi^2$ -eloszlással:** ha  $a = q/2$  és  $\lambda = 1/2$ , akkor a  $q$  szabadsági fokú  $\chi^2$ -eloszlást kapjuk vissza.
- **Várható érték és szórás:**

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

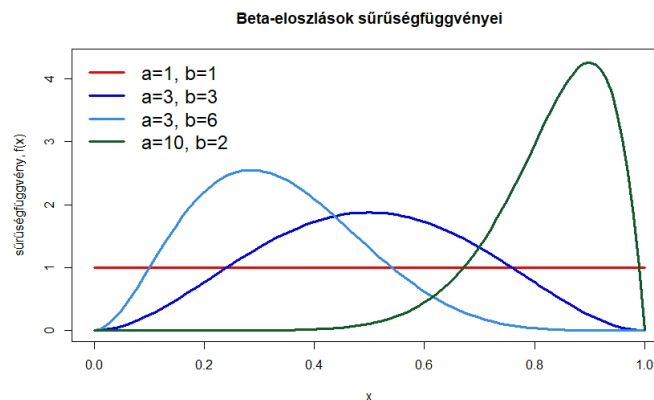
## 4.5. Beta-eloszlás

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

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

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  $k$ . legkisebb 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.



15. ábra. Különböző paraméterű beta-eloszlások sűrűségfüggvényei

**Házi feladat november 26., péntek, 12:05-ig** Sorsoljunk (például R-ben) 100 darab független standard normális eloszlású valószínűségi változót, legyenek ezek  $X_1, X_2, \dots, X_{100}$ , majd ezt ismételjük meg, az előzőtől függetlenül, ezek legyenek  $Y_1, Y_2, \dots, Y_{100}$ . Készítsünk hisztogramot az  $X_1/Y_1, X_2/Y_2, \dots, X_{100}/Y_{100}$  megfigyelésekből, úgy, hogy a hisztogram alatti terület 1 legyen. Erre a hisztogramra tegyük rá a Cauchy-eloszlás sűrűségfüggvényét, és nézzük meg, hogy mennyire jó az illeszkedés, illetve mennyivel jobban illeszkedik a görbe a hisztogramra, ha 100 helyett 1000 sorsolást végzünk.

Az 1000 elemű minta esetén számítsuk ki az  $X_1/Y_1, X_2/Y_2, \dots, X_{1000}/Y_{1000}$  minta átlagát és korrigált tapasztalati szórását (**sd**) is.

A megoldáshoz az R kód is hozzátartozik, nem csak az ábra és a végeredmények.