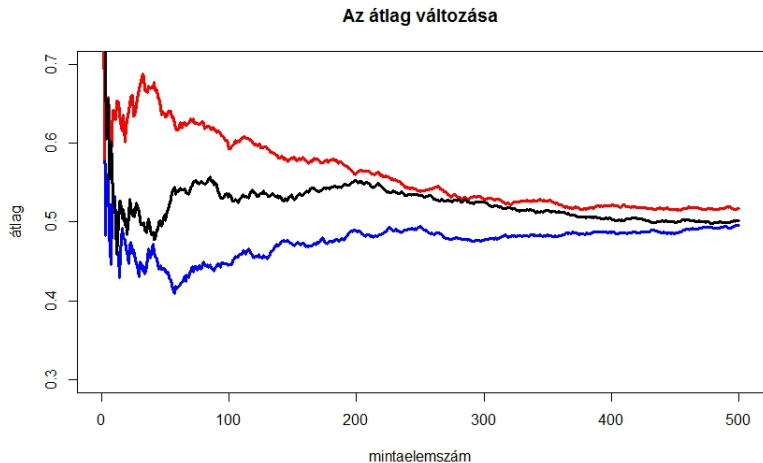


## Az átlag konvergenciája (10. előadás)



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

# Konvergenciafajták

**Valószínűségi változók sorozatának** (mint amilyen az átlagok sorozata) **többféle értelemben** definiálhatjuk a határértékét.

A  $Z_1, Z_2, \dots$ , valószínűségi változókból álló sorozat **sztochasztikusan konvergál** az  $Z$  valószínűségi változóhoz, ha minden  $\varepsilon > 0$ -ra

$$\mathbb{P}(|Z_n - Z| > \varepsilon) \rightarrow 0$$

teljesül  $n \rightarrow \infty$  esetén.

# Konvergenciafajták

**Valószínűségi változók sorozatának** (mint amilyen az átlagok sorozata) **többféle értelemben** definiálhatjuk a határértékét.

A  $Z_1, Z_2, \dots$ , valószínűségi változókból álló sorozat **sztochasztikusan konvergál** az  $Z$  valószínűségi változóhoz, ha minden  $\varepsilon > 0$ -ra

$$\mathbb{P}(|Z_n - Z| > \varepsilon) \rightarrow 0$$

teljesül  $n \rightarrow \infty$  esetén.

A  $Z_1, Z_2, \dots$ , valószínűségi változókból álló sorozat **1 valószínűséggel** konvergál az  $Z$  valószínűségi változóhoz, ha

$$\mathbb{P}(\omega \in \Omega : Z_n(\omega) \rightarrow Z(\omega) \text{ } n \rightarrow \infty \text{ esetén}) = 1.$$

# Konvergenciafajták

**Valószínűségi változók sorozatának** (mint amilyen az átlagok sorozata) **többféle értelemben** definiálhatjuk a határértékét.

A  $Z_1, Z_2, \dots$ , valószínűségi változókból álló sorozat **sztochasztikusan konvergál** az  $Z$  valószínűségi változóhoz, ha minden  $\varepsilon > 0$ -ra

$$\mathbb{P}(|Z_n - Z| > \varepsilon) \rightarrow 0$$

teljesül  $n \rightarrow \infty$  esetén.

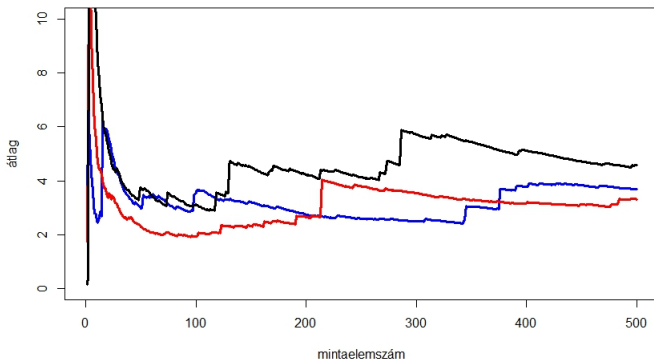
A  $Z_1, Z_2, \dots$ , valószínűségi változókból álló sorozat **1 valószínűséggel** konvergál az  $Z$  valószínűségi változóhoz, ha

$$\mathbb{P}(\omega \in \Omega : Z_n(\omega) \rightarrow Z(\omega) \text{ } n \rightarrow \infty \text{ esetén}) = 1.$$

1 valószínűséggel konvergál a sorozat  $\Rightarrow$  sztochasztikusan is, de fordítva nem feltétlenül.

# Az átlag változása

Az átlag változása nem létező várható érték esetén



Az átlag változása egy olyan esetben, amikor nem létezik a várható érték (nincs abszolút konvergencia a definícióban)

## A nagy számok gyenge törvénye: bizonyítás

Legyenek  $X_1, \dots, X_n$  független azonos eloszlású véges szórású valószínűségi változók. Legyen  $m = \mathbb{E}(X_1)$  és  $\sigma = D(X_1)$ .

A korábbiak szerint

$$\mathbb{E}(\bar{X}) = m; \quad D^2(\bar{X}) = \frac{\sigma^2}{n}.$$

# A nagy számok gyenge törvénye: bizonyítás

Legyenek  $X_1, \dots, X_n$  független azonos eloszlású véges szórású valószínűségi változók. Legyen  $m = \mathbb{E}(X_1)$  és  $\sigma = D(X_1)$ .

A korábbiak szerint

$$\mathbb{E}(\bar{X}) = m; \quad D^2(\bar{X}) = \frac{\sigma^2}{n}.$$

A Csebisev-egyenlőtlenség szerint minden  $\varepsilon > 0$ -ra

$$0 \leq \mathbb{P}(|\bar{X} - m| > \varepsilon) \leq \frac{D^2(\bar{X})}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2 n} \rightarrow 0 \quad (n \rightarrow \infty).$$

Tehát  $\mathbb{P}(|\bar{X} - m| > \varepsilon) \rightarrow 0$  teljesül minden  $\varepsilon > 0$  esetén, azaz  $\bar{X} \rightarrow m = \mathbb{E}(X_1)$  sztochasztikusan.

# A nagy számok törvénye

## 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 törvénye

## 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)

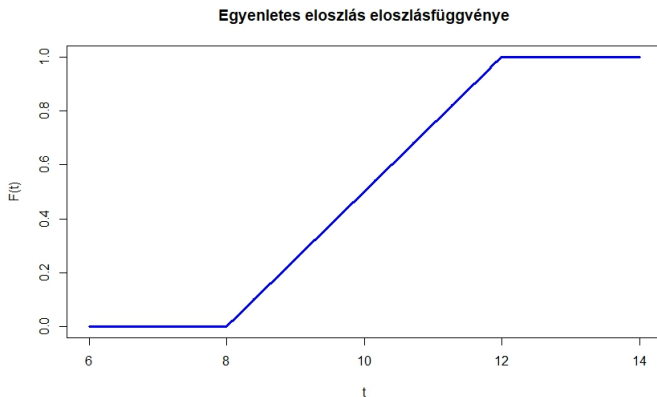
*Legyenek  $X_1, X_2, \dots$  valószínűségi változók, melyek függetlenek és azonos eloszlásúak. Tegyük fel még, hogy  $m = \mathbb{E}(X_1) < \infty$ . Ekkor*

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

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

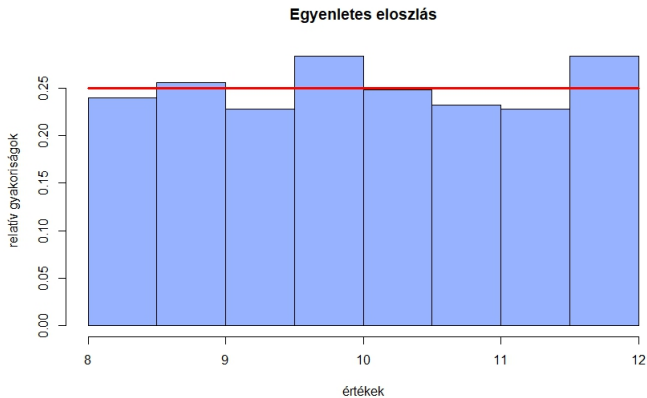
A második esetben gyengébb feltevésből erősebb állítás következik.

# Egyenletes eloszlás eloszlásfüggvénye



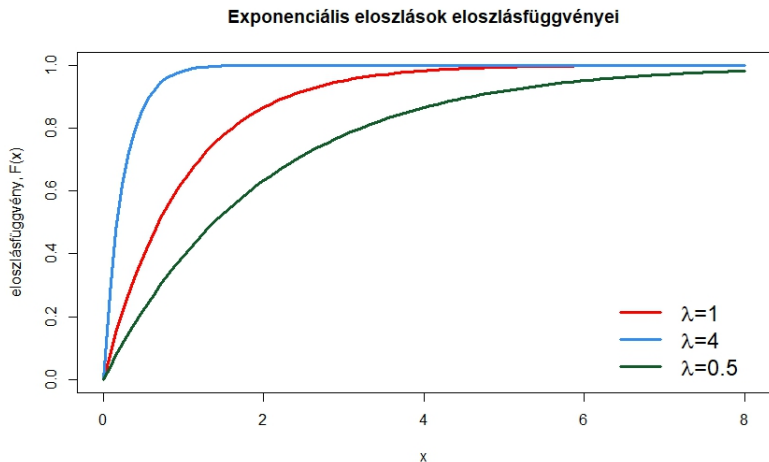
A  $[8, 12]$  intervallumon egyenletes eloszlás eloszlásfüggvénye

# Egyenletes eloszlás



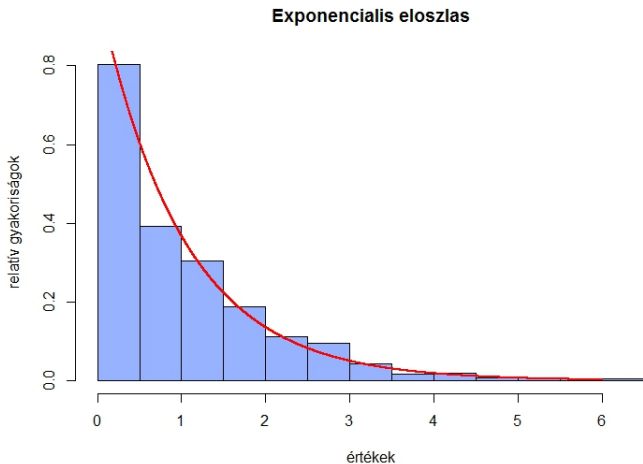
A  $[8, 12]$  intervallumon egyenletes eloszlásból vett 500 elemű minta hisztogramja

# Exponenciális eloszlás



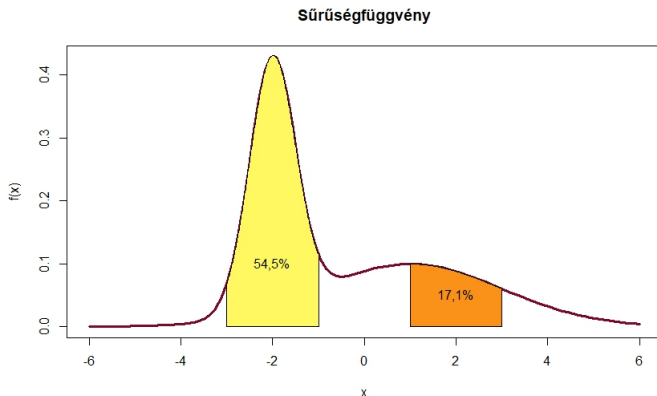
Különböző paraméterű ( $\lambda = \frac{1}{2}$ , 1, illetve 4) exponenciális eloszlások eloszlásfüggvényei

# Exponenciális eloszlás



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

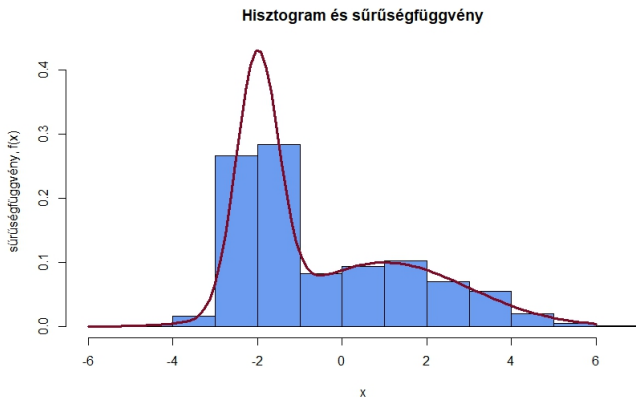
# Sűrűségfüggvény



Ha  $X$  sűrűségfüggvénye  $f$  (ami most az ábrán látható függvény):  $\mathbb{P}(-3 \leq X \leq -1) = \int_{-3}^{-1} f(x)dx = 54,5\%$ ;

$\mathbb{P}(1 \leq X \leq 3) = \int_1^3 f(x)dx = 17,1\%$ .

# Sűrűségfüggvény



Egy sűrűségfüggvény és hozzá tartozó ezer elemű független minta hisztogramja;  
nagyobb a sűrűségfüggvény  $\rightarrow$  nagyobb a gyakoriság;  
minta: független valószínűségi változók, melyek mindegyikének  $f$  a sűrűségfüggvénye

## Sűrűségfüggvény: definíció

Az  $X : \Omega \rightarrow \mathbb{R}$  valószínűségi változó **sűrűségfüggvénye** az  $f : \mathbb{R} \rightarrow \mathbb{R}$  függvény, ha

$$\mathbb{P}(X \leq t) = \int_{-\infty}^t f(x) dx$$

teljesül minden  $t \in \mathbb{R}$  számra.

**Nem minden** valószínűségi változónak van sűrűségfüggvénye, például a diszkrétnek nincs. Ha  $X$ -nek **van sűrűségfüggvénye**, akkor **abszolút folytonos** valószínűségi változónak nevezzük.

Ha az  $X$  valószínűségi változó sűrűségfüggvénye  $f$ , akkor tetszőleges  $a < b$  számokra

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

## A sűrűségfüggvény tulajdonságai

Legyen  $X$  abszolút folytonos valószínűségi változó, melynek  $F$  az eloszlásfüggvénye.

(a) Ha  $f$  az  $X$  sűrűségfüggvénye, akkor minden  $t \in \mathbb{R}$  számra

$$F(t) = \mathbb{P}(X \leq t) = \int_{-\infty}^t f(x) dx.$$

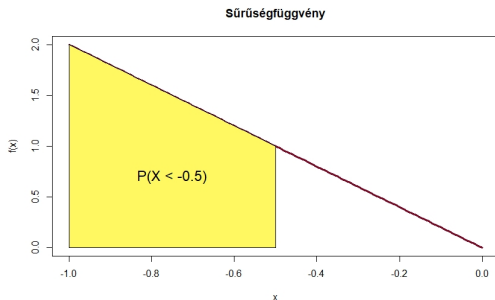
(b) Az  $f(t) = F'(t)$  függvény (azokra a  $t$ -kre, ahol  $F$  differenciálható) az  $X$  sűrűségfüggvénye.

Ha az  $f : \mathbb{R} \rightarrow \mathbb{R}$  függvény sűrűségfüggvény, akkor

- ❶  $f(x) \geq 0$  teljesül „majdnem minden”  $x \in \mathbb{R}$ -re (például véges vagy megszámlálható sok kivétel lehetséges).
- ❷  $\int_{-\infty}^{\infty} f(x) dx = 1$ .

*Fordítva:* ha  $f$  teljesíti ezt a két tulajdonságot, akkor van olyan valószínűségi változó, aminek  $f$  a sűrűségfüggvénye.

## Sűrűségfüggvény: példa



Ha  $X$  sűrűségfüggvénye  $f$ , akkor

$$F(-1/2) = \mathbb{P}(X \leq -1/2) = \int_{-\infty}^{-1/2} f(x) dx.$$

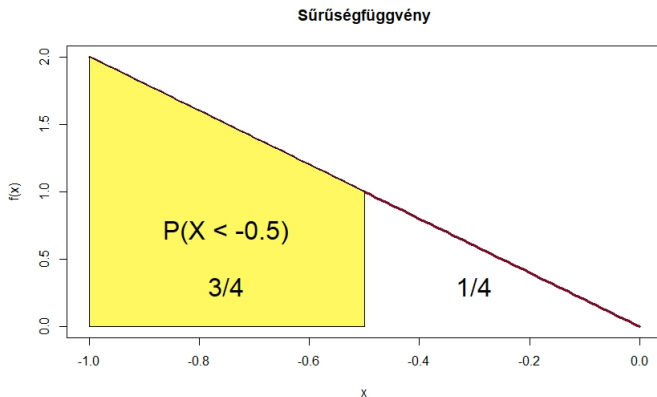
## Sűrűségfüggvény: példa

Legyen az  $X$  valószínűségi változó sűrűségfüggvénye  $f(x) = 2|x|$ , ha  $-1 < x < 0$ , és 0 különben. Mennyi  $X$  eloszlásfüggvényének értéke a  $-1/2$  helyen?

Felhasználva az **eloszlásfüggvény** és a **sűrűségfüggvény** definícióját, illetve hogy  $x \leq -1$  esetén  $f(x) = 0$ , azt kapjuk, hogy

$$\begin{aligned} F(-1/2) &= \mathbb{P}(X \leq -1/2) = \int_{-\infty}^{-1/2} f(x) dx = \int_{-1}^{-1/2} f(x) dx = \\ &= \int_{-1}^{-1/2} 2|x| dx = - \int_{-1}^{-1/2} 2x dx = - [x^2]_{x=-1}^{x=-1/2} = \\ &= - \left( \frac{1}{4} - 1 \right) = \frac{3}{4}. \end{aligned}$$

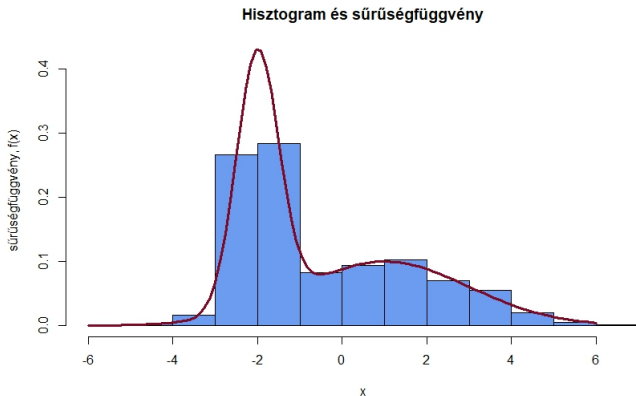
## Sűrűségfüggvény: példa



Ha  $X$  sűrűségfüggvénye  $f$ , akkor

$$F(-1/2) = \mathbb{P}(X \leq -1/2) = \int_{-\infty}^{-1/2} f(x) dx = \frac{3}{4}.$$

# Sűrűségfüggvény



Egy sűrűségfüggvény és hozzá tartozó ezer elemű független minta hisztogramja – mennyi lehet az  $f$  sűrűségfüggvényű valószínűségi változó **várható értéke** és **szórása**?

# Diszkrét és abszolút folytonos eset

A **várható értéket** olyan formában nem tudjuk definiálni abszolút folytonos eloszlásokra, mint diszkrét esetben, hiszen  $\mathbb{P}(X = x) = 0$  minden  $x$ -re. Helyette:

**diszkrét**

$X$  lehetséges értékei:  $x_1, x_2, \dots$

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j)$$

$$\mathbb{E}(X^2) = \sum_{j=1}^{\infty} x_j^2 \cdot \mathbb{P}(X = x_j)$$

**abszolút folytonos**

$X$  sűrűségfüggvénye:  $f$ .

# Diszkrét és abszolút folytonos eset

A **várható értéket** olyan formában nem tudjuk definiálni abszolút folytonos eloszlásokra, mint diszkrét esetben, hiszen  $\mathbb{P}(X = x) = 0$  minden  $x$ -re. Helyette:

## diszkrét

$X$  lehetséges értékei:  $x_1, x_2, \dots$

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j)$$

$$\mathbb{E}(X^2) = \sum_{j=1}^{\infty} x_j^2 \cdot \mathbb{P}(X = x_j)$$

## abszolút folytonos

$X$  sűrűségfüggvénye:  $f$ .

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

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx$$

# Diszkrét és abszolút folytonos eset

A **várható értéket** olyan formában nem tudjuk definiálni abszolút folytonos eloszlásokra, mint diszkrét esetben, hiszen  $\mathbb{P}(X = x) = 0$  minden  $x$ -re. Helyette:

**diszkrét**

$X$  lehetséges értékei:  $x_1, x_2, \dots$

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j)$$

$$\mathbb{E}(X^2) = \sum_{j=1}^{\infty} x_j^2 \cdot \mathbb{P}(X = x_j)$$

**abszolút folytonos**

$X$  sűrűségfüggvénye:  $f$ .

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

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx$$

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

# Diszkrét és abszolút folytonos eset

A **várható értéket** olyan formában nem tudjuk definiálni abszolút folytonos eloszlásokra, mint diszkrét esetben, hiszen  $\mathbb{P}(X = x) = 0$  minden  $x$ -re. Helyette:

**diszkrét**

$X$  lehetséges értékei:  $x_1, x_2, \dots$

$$\mathbb{E}(X) = \sum_{j=1}^{\infty} x_j \cdot \mathbb{P}(X = x_j)$$

$$\mathbb{E}(X^2) = \sum_{j=1}^{\infty} x_j^2 \cdot \mathbb{P}(X = x_j)$$

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

$$\mathbb{E}(X^k) = \sum_{j=1}^{\infty} x_j^k \cdot \mathbb{P}(X = x_j)$$

$$\mathbb{E}(g(X)) = \sum_{j=1}^{\infty} g(x_j) \cdot \mathbb{P}(X = x_j)$$

**abszolút folytonos**

$X$  sűrűségfüggvénye:  $f$ .

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

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x) dx$$

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

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

## Várható érték és szórás abszolút folytonos esetben

Legyen  $X$  abszolút folytonos valószínűségi változó, melynek sűrűségfüggvénye  $f$ . Ekkor  $X$  **várható értéke**:

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

ha ez az integrál létezik és véges.

Tegyük fel, hogy az  $X$  valószínűségi változó abszolút folytonos, sűrűségfüggvénye  $f$ , és  $\mathbb{E}(X^2)$  létezik, azaz az  $\int_{-\infty}^{\infty} x^2 f(x) dx$  integrál véges. Ekkor  $X$  **szórásnégyzete**:

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

**szórása** pedig

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

*A szórás definíciója megegyezik a diszkrét esetben használttal.*

# Az egyenletes eloszlás várható értéke

Az  $[a, b]$  intervallumon egyenletes eloszlás sűrűségfüggvénye:

$$f(s) = \begin{cases} \frac{1}{b-a}, & \text{ha } a \leq s \leq b; \\ 0, & \text{különben.} \end{cases}$$

Ha  $X$  egyenletes eloszlású az  $[a, b]$  intervallumon, akkor várható értéke

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{\infty} s \cdot f(s) ds = \int_a^b s \cdot \frac{1}{b-a} ds = \frac{1}{b-a} \left[ \frac{s^2}{2} \right]_{s=a}^b = \\ &= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{a+b}{2}, \end{aligned}$$

hiszen az  $x$  függvény primitív függvénye  $\frac{x^2}{2}$ , és  $b^2 - a^2 = (b-a)(b+a)$ .

## Momentumok

Az  $X$  valószínűségi változók  **$k$ . momentuma** a  $k$ . hatványának várható értéke:

$$\mathbb{E}(X^k).$$

Általában igaz, hogy ha  $X$  abszolút folytonos valószínűségi változó,  $f$  a sűrűségfüggvénye, és  $\mathbb{E}(g(X))$  létezik, akkor

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

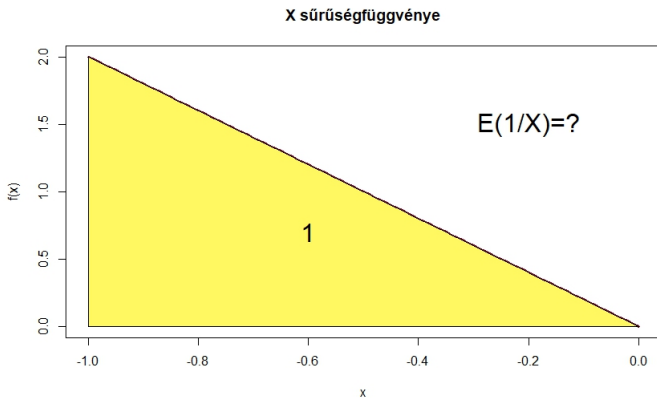
Ezért a  $k$ . momentum kiszámítása:

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

Következmény: a **szórásnégyzetet** a következőképpen számíthatjuk ki abszolút folytonos  $X$  valószínűségi változó esetén:

$$D^2(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 = \int_{-\infty}^{\infty} x^2 f(x) dx - \left[ \int_{-\infty}^{\infty} x \cdot f(x) dx \right]^2.$$

# Várható érték abszolút folytonos esetben: példa



## Várható érték abszolút folytonos esetben: példa

Legyen az  $X$  valószínűségi változó sűrűségfüggvénye  $f(x) = 2|x|$ , ha  $-1 < x < 0$ , és 0 különben. Mennyi az  $1/X$  valószínűségi változó várható értéke?

Mivel  $X$  sűrűségfüggvénye azonosan 0, ha  $x > 0$ , ezért  $X < 0$  és  $1/X < 0$  biztosan teljesül. Így  $\mathbb{E}(X) < 0$  teljesülni fog.

Pontosabban, mivel

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

a  $g(x) = 1/x$  függvénnyel:

$$\mathbb{E}(1/X) = \int_{-\infty}^{\infty} \frac{1}{x} \cdot f(x) dx = \int_{-1}^0 \frac{2|x|}{x} dx = \int_{-1}^0 (-2) dx = -2.$$

## Várható érték abszolút folytonos esetben: példa

Tegyük fel, hogy a holnap hulló csapadék mennyiségének sűrűségfüggvénye az alábbi:

$$f(x) = \begin{cases} 0,2, & \text{ha } 0 \leq x \leq 1; \\ 0,4, & \text{ha } 1 < x \leq 3; \\ 0 & \text{ha } x < 0 \text{ vagy } x > 3. \end{cases}$$

Jelölje a csapadékmennyiséget  $X$ . A csapadékmennyiség **várható értéke**:

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_0^1 x \cdot 0,2 dx + \int_1^3 x \cdot 0,4 dx = \\ &= 0,2 \cdot \left( \frac{1^2}{2} - \frac{0^2}{2} \right) + 0,4 \cdot \left( \frac{3^2}{2} - \frac{1^2}{2} \right) = 1,7. \end{aligned}$$

## Szórás abszolút folytonos esetben: példa

Tegyük fel, hogy a holnap hulló csapadék mennyiségének sűrűségfüggvénye az alábbi:

$$f(x) = \begin{cases} 0,2, & \text{ha } 0 \leq x \leq 1; \\ 0,4, & \text{ha } 1 < x \leq 3; \\ 0 & \text{ha } x < 0 \text{ vagy } x > 3. \end{cases}$$

Már láttuk, hogy  $\mathbb{E}(X) = 1,7$ .

A csapadékmennyiség négyzetének várható értéke:

$$\begin{aligned} \mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 \cdot f(x) \, ds = \int_0^1 x^2 \cdot 0,2 \, dx + \int_1^3 x^2 \cdot 0,4 \, dx = \\ &= 0,2 \cdot \left( \frac{1^3}{3} - \frac{0^3}{3} \right) + 0,4 \cdot \left( \frac{3^3}{3} - \frac{1^3}{3} \right) = 3,53. \end{aligned}$$

Ez alapján a csapadékmennyiség szórása:

$$D(X) = \sqrt{\mathbb{E}(X^2) - (\mathbb{E}(X))^2} = \sqrt{3,53 - 1,7^2} = 0,8.$$

# Az egyenletes eloszlás várható értéke

Az  $[a, b]$  intervallumon egyenletes eloszlás sűrűségfüggvénye:

$$f(s) = \begin{cases} \frac{1}{b-a}, & \text{ha } a \leq s \leq b; \\ 0, & \text{különben.} \end{cases}$$

Ha  $X$  egyenletes eloszlású az  $[a, b]$  intervallumon, akkor várható értéke

$$\begin{aligned} \mathbb{E}(X) &= \int_{-\infty}^{\infty} s \cdot f(s) ds = \int_a^b s \cdot \frac{1}{b-a} ds = \frac{1}{b-a} \left[ \frac{s^2}{2} \right]_{s=a}^b = \\ &= \frac{1}{b-a} \cdot \frac{b^2 - a^2}{2} = \frac{a+b}{2}, \end{aligned}$$

hiszen az  $x$  függvény primitív függvénye  $\frac{x^2}{2}$ , és  $b^2 - a^2 = (b-a)(b+a)$ .

## Az egyenletes eloszlás szórása

Az  $[a, b]$  intervallumon egyenletes eloszlás sűrűségfüggvénye:

$$f(s) = \begin{cases} \frac{1}{b-a}, & \text{ha } a \leq s \leq b; \\ 0, & \text{különben.} \end{cases}$$

Ha  $X$  egyenletes eloszlású az  $[a, b]$  intervallumon, akkor várható értéke

$$\begin{aligned} \mathbb{E}(X^2) &= \int_{-\infty}^{\infty} s^2 \cdot f(s) ds = \int_a^b s^2 \cdot \frac{1}{b-a} ds = \frac{1}{b-a} \left[ \frac{s^3}{3} \right]_{s=a}^b \\ &= \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3}, \end{aligned}$$

hiszen az  $x^2$  függvény primitív függvénye  $\frac{x^3}{3}$ , és  $b^3 - a^3 = (b-a)(a^2 + ab + b^2)$ .

## Az egyenletes eloszlás szórása

Az  $[a, b]$  intervallumon egyenletes eloszlás sűrűségfüggvénye:

$$f(s) = \begin{cases} \frac{1}{b-a}, & \text{ha } a \leq s \leq b; \\ 0, & \text{különben.} \end{cases}$$

Ha  $X$  egyenletes eloszlású az  $[a, b]$  intervallumon, akkor várható értéke

$$\begin{aligned} \mathbb{E}(X^2) &= \int_{-\infty}^{\infty} s^2 \cdot f(s) ds = \int_a^b s^2 \cdot \frac{1}{b-a} ds = \frac{1}{b-a} \left[ \frac{s^3}{3} \right]_{s=a}^b \\ &= \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3} = \frac{a^2 + ab + b^2}{3}, \end{aligned}$$

hiszen az  $x^2$  függvény primitív függvénye  $\frac{x^3}{3}$ , és  $b^3 - a^3 = (b-a)(a^2 + ab + b^2)$ .

$$\begin{aligned} D^2(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{a^2 + ab + b^2}{3} - \left( \frac{a+b}{2} \right)^2 \\ &= \frac{a^2 + ab + b^2}{3} - \frac{a^2 + 2ab + b^2}{4} = \frac{a^2 - 2ab + b^2}{12} = \frac{(b-a)^2}{12}. \end{aligned}$$

## Egyenletes eloszlás (uniform distribution)

Az  $X$  valószínűségi változó **egyenletes eloszlású** az  $[a, b]$  intervallumon, ha sűrűségfüggvénye

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{ha } a \leq x \leq b; \\ 0, & \text{különben.} \end{cases}$$

Ekkor

i)  $X$  eloszlásfüggvénye:

$$F(t) = \mathbb{P}(X \leq t) = \begin{cases} 0, & \text{ha } t \leq a; \\ \frac{t-a}{b-a}, & \text{ha } a < t < b; \\ 1, & \text{ha } t \geq b. \end{cases}$$

ii) Ha  $a \leq c \leq d \leq b$ , akkor

$$\mathbb{P}(c \leq X \leq d) = \int_c^d f(x) dx = \int_c^d \frac{1}{b-a} dx = \frac{d-c}{b-a}.$$

iii) Az  $X$  valószínűségi változó **várható értéke** és **szórása**:

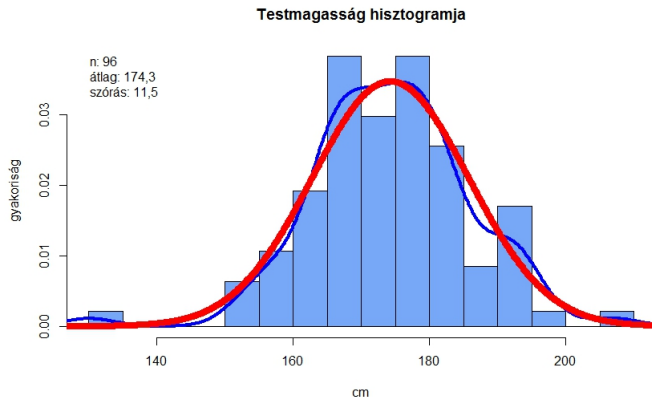
$$\mathbb{E}(X) = \frac{a+b}{2}; \quad D(X) = \frac{b-a}{\sqrt{12}}.$$

# Normális eloszlás: bevezetés

Ha egy, a valóságban megfigyelhető valószínűségi változó eloszlását, pontosabban a **sűrűségfüggvényét** szeretnénk meghatározni:

- az adatokból készíthetünk hisztogramot;
- a hisztogram és a sűrűségfüggvény alakja sok független megfigyelés esetén „közel” van egymáshoz;
- megfigyelhetjük, hogy különféle mennyiségek esetén a hisztogramok gyakran **hasonló alakúak** → a gyakran előforduló sűrűségfüggvény-típusokat érdemes külön megérteni;
- az egyik ilyen a **normális eloszlás**, melynek sűrűségfüggvénye az  $e^{-x^2}$  függvényből származtatható
- például különféle *mérési eredmények* (a mérési hibák következtében), illetve élőlények *biológiai jellemzői* gyakran normális eloszlást követnek (például: testmagasság)
- a normális eloszlás a **statisztikában** is kulcsfontosságú

# Testmagasság



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

## Normális eloszlás: definíció

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

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

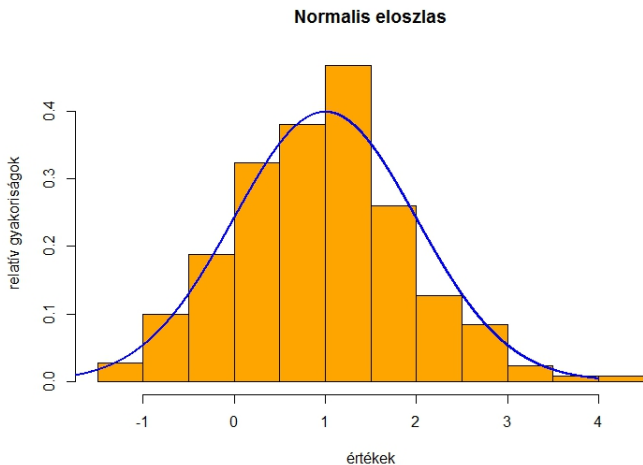
Jelölése:  $Y \sim N(m, \sigma^2)$ .

Ha  $Y \sim N(m, \sigma^2)$ , akkor  $\mathbb{E}(Y) = m$ ,  $D(Y) = \sigma$ .

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

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

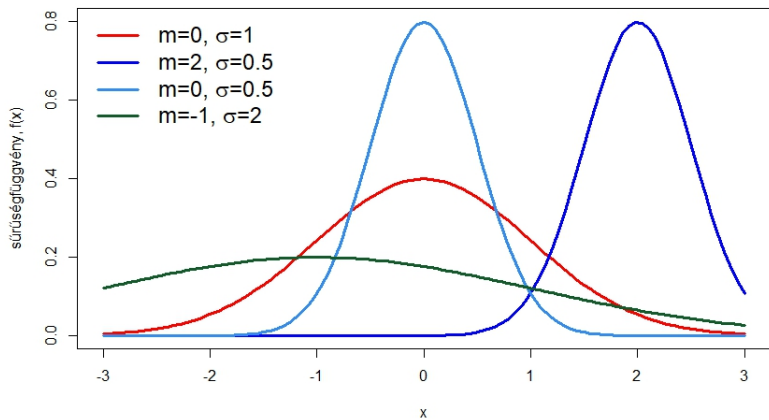
# Normális eloszlás



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

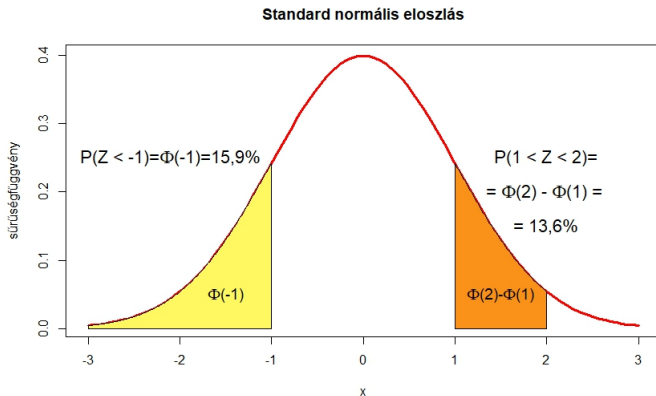
# Normális eloszlás

Normális eloszlások sűrűségfüggvényei



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

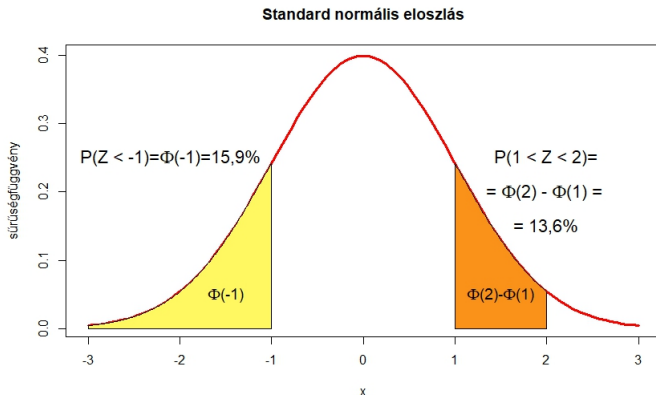
# A $\Phi$ függvény



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

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

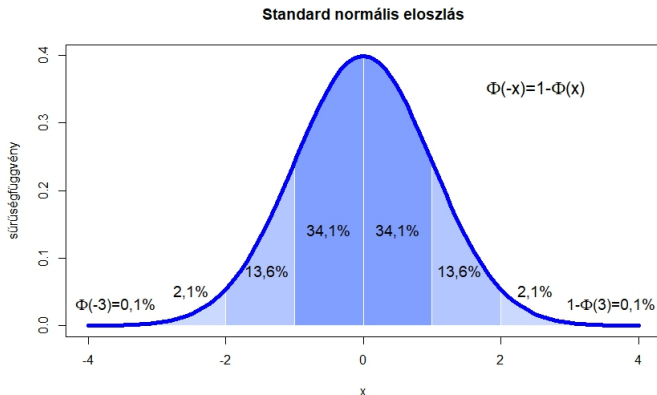
# Standard normális eloszlás



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

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

# Standard normális eloszlás



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

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

## Standard normális eloszlás

A  $Z$  valószínűségi változó **standard normális eloszlású**, azaz  $Z \sim N(0,1)$ , ha sűrűségfüggvénye

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right).$$

Ekkor **eloszlásfüggvénye**  $\Phi$ , azaz

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

Továbbá

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

*Más normális eloszlások esetén ezeket a valószínűségeket a  $\Phi$  függvényre vezetjük vissza. A  $\Phi$  függvény az R-ben: pnorm*

## 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

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

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

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

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

## Normális eloszlás: példa

Tegyük fel, hogy az  $Y$  valószínűségi változó normális eloszlású  $m = 4$  várható értékkel és  $\sigma = 3$  szórással. Ekkor

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

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

mert

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

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

# A normális eloszlás tulajdonságai

**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

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

## A normális eloszlás tulajdonságai

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

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

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

## A normális eloszlás tulajdonságai

Legyenek  $Y_1, Y_2, \dots, Y_n$  **független normális eloszlású** valószínűségi változók, melyek várható értéke  $m$ , szórásuk  $\sigma$ . 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 szórású 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.