

## Statisztikai mező (4. előadás)

### Definíció

Az  $(\Omega, \mathcal{A}, \mathcal{P})$  hármast **statisztikai mezőnek** nevezzük, ha minden  $\mathbb{P} \in \mathcal{P}$ -re  $(\Omega, \mathcal{A}, \mathbb{P})$  Kolmogorov-féle valószínűségi mező.

Paraméteres statisztikai mező:  $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$ . Ekkor  $\vartheta$  az ismeretlen paraméter, mely egy  $\Theta \subseteq \mathbb{R}^q$  ismert halmaz eleme.

Például:  $\mathcal{P}$  lehet például

- a  $\lambda$  paraméterű Poisson-eloszlások halmaza;
- a normális eloszlások halmaza (ekkor  $\vartheta = (m, \sigma)$  az ismeretlen paraméter);
- az  $[a, b]$  intervallumon egyenletes eloszlások halmaza (ekkor  $\vartheta = (a, b)$  az ismeretlen paraméter).

# Minta és statisztika

## Definíció (Minta)

Legyen  $(\Omega, \mathcal{A}, \mathcal{P})$  statisztikai mező. Egy

$$\underline{X} = (X_1, X_2, \dots, X_n) : \Omega \rightarrow B \subseteq \mathbb{R}^n$$

valószínűségi vektorváltozót ( $n$  elemű) **mintának** nevezünk. Itt  $B$  a mintatér,  $n$  a minta elemszáma vagy nagysága. A minta független, ha az  $X_1, X_2, \dots, X_n$  valószínűségi változók függetlenek.

# Minta és statisztika

## Definíció (Minta)

Legyen  $(\Omega, \mathcal{A}, \mathcal{P})$  statisztikai mező. Egy

$$\underline{X} = (X_1, X_2, \dots, X_n) : \Omega \rightarrow B \subseteq \mathbb{R}^n$$

valószínűségi vektorváltozót ( $n$  elemű) **mintának** nevezünk. Itt  $B$  a mintatér,  $n$  a minta elemszáma vagy nagysága. A minta független, ha az  $X_1, X_2, \dots, X_n$  valószínűségi változók függetlenek.

## Definíció (Statisztika)

Legyen  $T : B \rightarrow \mathbb{R}^k$  függvény. Ekkor a  $T(X_1, X_2, \dots, X_n)$  valószínűségi változót statisztikának nevezzük.

Például:  $T(X_1, \dots, X_n) = \bar{X}$  a mintaátlag, vagy  $T(X_1, \dots, X_n) = s_n^*$  a korrigált tapasztalati szórás.

# Becslések és tulajdonságaik

- $(\Omega, \mathcal{A}, \mathcal{P})$  statisztikai mező;
- $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$  valamely  $\Theta$  halmazzal ( $\Theta$  a paramétertér);
- $g : \Theta \rightarrow \mathbb{R}$  függvény.
- Cél: olyan  $T$  statisztika keresése, amire a  $T(X)$  valószínűségi változó és a  $g(\vartheta)$  érték valamilyen értelemben közel esnek egymáshoz.

# Becslések és tulajdonságaik

- $(\Omega, \mathcal{A}, \mathcal{P})$  statisztikai mező;
- $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$  valamely  $\Theta$  halmazzal ( $\Theta$  a paraméterter);
- $g : \Theta \rightarrow \mathbb{R}$  függvény.
- Cél: olyan  $T$  statisztika keresése, amire a  $T(X)$  valószínűségi változó és a  $g(\vartheta)$  érték valamilyen értelemben közel esnek egymáshoz.

## Definíció (Torzítatlanság)

A  $T$  statisztika torzítatlan becslés  $g(\vartheta)$ -ra, ha minden  $\vartheta \in \Theta$ -ra

$$\mathbb{E}_\vartheta(T(X_1, \dots, X_n)) = g(\vartheta).$$

A  $T$  statisztika torzítása a  $b_T(\vartheta) = \mathbb{E}_\vartheta(T(X_1, \dots, X_n)) - g(\vartheta)$  függvény.

**Példa.**  $X_1, X_2, \dots, X_n$  független minta a  $[0, \vartheta]$  intervallumon egyenletes eloszlásból.

# Becslések és tulajdonságaik

- $(\Omega, \mathcal{A}, \mathcal{P})$  statisztikai mező;
- $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$  valamely  $\Theta$  halmazzal ( $\Theta$  a paraméterter);
- $g : \Theta \rightarrow \mathbb{R}$  függvény.
- Cél: olyan  $T$  statisztika keresése, amire a  $T(X)$  valószínűségi változó és a  $g(\vartheta)$  érték valamilyen értelemben közel esnek egymáshoz.

## Definíció (Torzítatlanság)

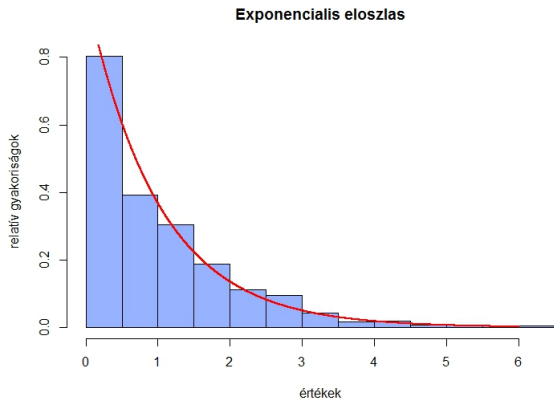
A  $T$  statisztika torzítatlan becslés  $g(\vartheta)$ -ra, ha minden  $\vartheta \in \Theta$ -ra

$$\mathbb{E}_\vartheta(T(X_1, \dots, X_n)) = g(\vartheta).$$

A  $T$  statisztika torzítása a  $b_T(\vartheta) = \mathbb{E}_\vartheta(T(X_1, \dots, X_n)) - g(\vartheta)$  függvény.

**Példa.**  $X_1, X_2, \dots, X_n$  független minta a  $[0, \vartheta]$  intervallumon egyenletes eloszlásból. Ekkor  $2\bar{X}$  torzítatlan becslés  $g(\vartheta) = \vartheta$ -ra:  $\mathbb{E}(2\bar{X}) = \vartheta$ .

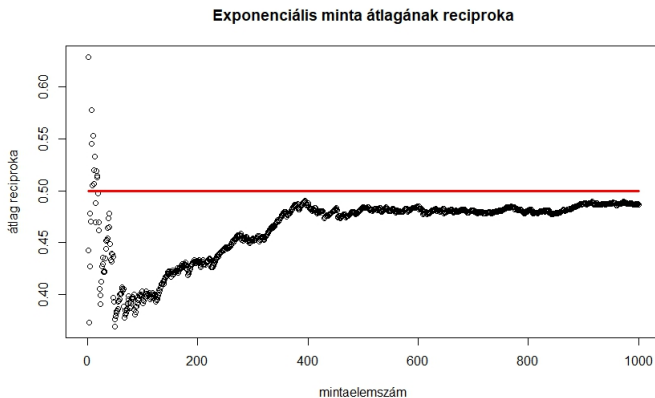
# Az exponenciális eloszlású minta hisztogramja



Exponenciális eloszlású minta hisztogramja és sűrűségfüggvénye:

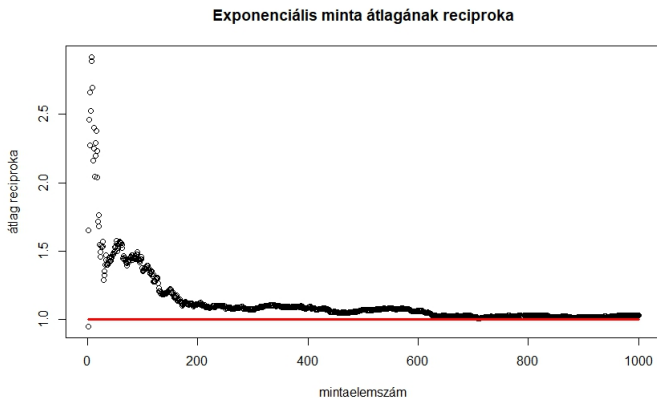
$$f(x) = \lambda \exp(-\lambda x) \mathbb{I}(x > 0); \quad \mathbb{E}(X) = D(X) = \frac{1}{\lambda}.$$

# Exponenciális eloszlás



$\lambda = 0,5$  paraméterű exponenciális eloszlást generálva a mintaátlag reciproka 0,5-höz tart, azaz **konzisztens** becslés, hiszen ez minden  $\lambda$ -ra teljesül.

# Exponenciális eloszlás



$\lambda = 1$  paraméterű exponenciális eloszlást generálva a mintaátlag reciproka 1-hez tart, azaz **konzisztens** becslés, hiszen ez minden  $\lambda$ -ra teljesül.

# Konzisztencia

## Definíció

A  $T_n = T_n(X_1, \dots, X_n)$  **konzisztens** becsléssorozat  $g(\vartheta)$ -ra, ha minden  $\vartheta \in \Theta$ -ra

$$(T_n(X_1, \dots, X_n)) \rightarrow g(\vartheta)$$

$n \rightarrow \infty$  esetén sztochasztikusan, azaz minden  $\vartheta \in \Theta$  és  $\varepsilon > 0$ -ra teljesül, hogy

$$\mathbb{P}_\vartheta(|T_n - g(\vartheta)| > \varepsilon) \rightarrow 0 \quad (n \rightarrow \infty).$$

# Konzisztencia

## Definíció

A  $T_n = T_n(X_1, \dots, X_n)$  **konzisztens** becsléssorozat  $g(\vartheta)$ -ra, ha minden  $\vartheta \in \Theta$ -ra

$$(T_n(X_1, \dots, X_n)) \rightarrow g(\vartheta)$$

$n \rightarrow \infty$  esetén sztochasztikusan, azaz minden  $\vartheta \in \Theta$  és  $\varepsilon > 0$ -ra teljesül, hogy

$$\mathbb{P}_\vartheta(|T_n - g(\vartheta)| > \varepsilon) \rightarrow 0 \quad (n \rightarrow \infty).$$

Elégséges feltétel:

$$\mathbb{E}_\vartheta(T_n(X)) \rightarrow g(\vartheta) \quad \text{és} \quad D_\vartheta(T_n(X)) \rightarrow 0$$

minden  $\vartheta \in \Theta$ -ra.

## Példák torzítatlan, konzisztens becslésekre

$X_1, X_2, \dots$  független azonos eloszlású minta. Ekkor

$$T_n = \frac{X_1 + X_2 + \dots + X_n}{n} \rightarrow \mathbb{E}_\vartheta(X_1)$$

teljesül  $n \rightarrow \infty$  esetén sztochasztikusan a nagy számok gyenge törvénye szerint, vagyis az **átlag** konzisztens becslés a **várható értékre**.

Speciális eset: a **relatív gyakoriság** konzisztens becslés a **valószínűségre**.

## Példák torzítatlan, konzisztens becslésekre

$X_1, X_2, \dots$  független azonos eloszlású minta. Ekkor

$$T_n = \frac{X_1 + X_2 + \dots + X_n}{n} \rightarrow \mathbb{E}_\vartheta(X_1)$$

teljesül  $n \rightarrow \infty$  esetén sztochasztikusan a nagy számok gyenge törvénye szerint, vagyis az **átlag** konzisztens becslés a **várható értékre**.

Speciális eset: a **relatív gyakoriság** konzisztens becslés a **valószínűségre**.

Nevezetes eloszlások:

- Poisson-eloszlás  $\lambda$  paraméterére az átlag torzítatlan, konzisztens
- a normális eloszlás  $m$  paraméterére az átlag torzítatlan és konzisztens; a  $\sigma$  paraméterre a tapasztalati szórás és a korigált tapasztalati szórás konzisztensek, de nem torzítatlanok;  $\sigma^2$ -re  $s_n^{*2}$  torzítatlan
- exponenciális eloszlás:  $1/\bar{X}$  konzisztens  $\lambda$ -ra, de nem torzítatlan a paraméterre
- exponenciális eloszlás:  $(n+1) \cdot \min(X_1, \dots, X_n)$  torzítatlan, de nem konzisztens a várható értékre (vagyis  $1/\lambda$ -ra).

# Becslési módszerek

Legyen  $(\Omega, \mathcal{A}, \mathcal{P})$  statisztikai mező, ahol  $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$ , vagyis az ismeretlen eloszlás a  $\vartheta$  paraméterrel jellemezhető.

Találhatunk-e olyan általános módszereket, melyeket alkalmazva a  $\vartheta$  ismeretlen paraméter becslésére egy megfelelő  $T(X_1, \dots, X_n)$  statisztika

- torzítatlan: várható értéke megegyezik a becsülni kívánt  $\vartheta$  paraméterrel?
- hatásos: szórása a lehető legkisebb?
- konzisztens: a mintaelemszámmal végtelenhez tartva a becsléseink sorozata  $\vartheta$ -hoz tart:  $T_n \rightarrow \vartheta$  sztochasztikusan  $n \rightarrow \infty$  esetén?

# Maximumlikelihood-módszer

Van két telefontöltőnk, melyek ránézésre megkülönböztethetlenné. A jó minden kipróbálásnál a többletől függetlenül 90%, a rossz 20% valószínűséggel működik.

Az egyik kábelt kiválasztottuk, 10 kipróbálásból pontosan 8-szor működött. Ez melyik kábel lehetett?

## Maximumlikelihood-módszer

Van két telefontöltőnk, melyek ránézésre megkülönböztethetetlenek. A jó minden kipróbálásnál a többletől függetlenül 90%, a rossz 20% valószínűséggel működik.

Az egyik kábelt kiválasztottuk, 10 kipróbálásból pontosan 8-szor működött. Ez melyik kábel lehetett?

$$\mathbb{P}(\text{a jó kábel 8-szor működik}) = \binom{10}{8} 0,9^8 0,1^2 = 19,8\%.$$

$$\mathbb{P}(\text{a rossz kábel 8-szor működik}) = \binom{10}{8} 0,2^8 0,8^2 = 0,00007\%.$$

Inkább a jó kábel lehetett, amit kiválasztottunk.

# Maximumlikelihood-módszer

Van egy telefontöltőnk, mely minden kipróbálásnál a többlettől függetlenül  $p$  valószínűséggel működik.

Itt az ismeretlen paraméter:  $p \in [0, 1]$ .

A kábel 10 kipróbálásból pontosan 8-szor működött. Mennyi lehet  $p$  értéke?

# Maximumlikelihood-módszer

Van egy telefontöltőnk, mely minden kipróbálásnál a többletől függetlenül  $p$  valószínűséggel működik.

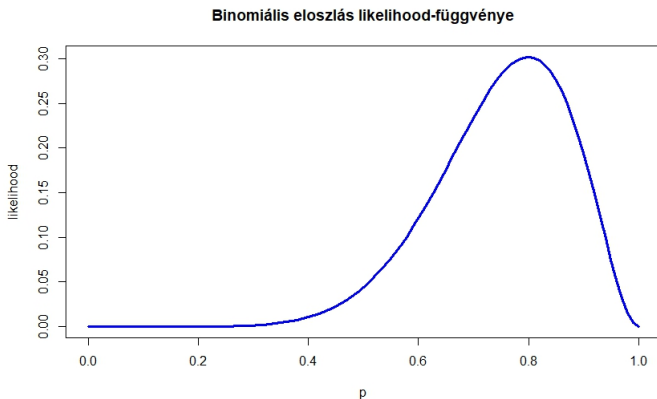
Itt az ismeretlen paraméter:  $p \in [0, 1]$ .

A kábel 10 kipróbálásból pontosan 8-szor működött. Mennyi lehet  $p$  értéke?

$$\mathbb{P}_p(\text{a kábel 8-szor működik}) = \binom{10}{8} p^8 (1-p)^2.$$

Kérdés: ez milyen  $p \in [0, 1]$ -re lesz a legnagyobb?

# Maximumlikelihood-módszer



A likelihood-függvény:  $\mathbb{P}_p(\text{a kábel 8-szor működik}) = \binom{10}{8} p^8 (1-p)^2$  a  $p$  ismeretlen paraméter függvényében. A maximumhelye:  $\hat{p} = 0,8$ .

# Maximumlikelihood-módszer

## Definíció (Likelihood-függvény)

Ha az  $(Y_1, \dots, Y_n)$  független minta diszkrét (a lehetséges értékeinek száma véges vagy megszámlálható sok), akkor a likelihood-függvénye:

$$L_{n,\vartheta}(k_1, \dots, k_n) = \prod_{j=1}^n \mathbb{P}_{j,\vartheta}(Y_j = k_j) \quad ((k_1, \dots, k_n) \in H).$$

# Maximumlikelihood-módszer

## Definíció (Likelihood-függvény)

Ha az  $(Y_1, \dots, Y_n)$  független minta diszkrét (a lehetséges értékeinek száma véges vagy megszámlálható sok), akkor a likelihood-függvénye:

$$L_{n,\vartheta}(k_1, \dots, k_n) = \prod_{j=1}^n \mathbb{P}_{j,\vartheta}(Y_j = k_j) \quad ((k_1, \dots, k_n) \in H).$$

Ha az  $(Y_1, \dots, Y_n)$  független minta abszolút folytonos, és  $Y_j$  sűrűségfüggvénye (a  $\mathbb{P}_\vartheta$  valószínűség mellett)  $f_{j,\vartheta}$ , akkor a minta likelihood-függvénye:

$$L_{n,\vartheta}(t_1, \dots, t_n) = \prod_{j=1}^n f_{j,\vartheta}(t_j) \quad (t_1, \dots, t_n \in \mathbb{R}).$$

# Maximumlikelihood-módszer

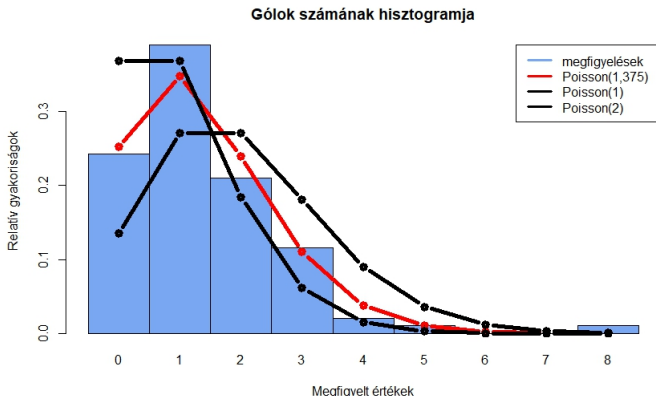
Legyen  $(\Omega, \mathcal{A}, \mathcal{P})$  statisztikai mező, ahol  $\mathcal{P} = \{\mathbb{P}_\vartheta : \vartheta \in \Theta\}$ , vagyis az ismeretlen eloszlás a  $\vartheta$  paraméterrel jellemezhető.

## Definíció (Maximum-likelihood becslés)

A  $\vartheta$  maximumlikelihood-becslése (ML-becslése) az  $X_1, \dots, X_n$  mintából  $\hat{\vartheta}$ , ha maximalizálja a  $\vartheta \mapsto L_{n,\vartheta}(X_1, \dots, X_n)$  függvényt, ahol  $L_{n,\vartheta}$  a minta likelihood-függvénye. Azaz, ha

$$L_{n,\hat{\vartheta}}(X_1, \dots, X_n) \geq L_{n,\vartheta}(X_1, \dots, X_n) \text{ minden } \vartheta \in \Theta\text{-ra.}$$

# Poisson-eloszlás paraméterének becslése



A gólok számának hisztogramja  $n = 95$  mérkőzésen, és különböző paraméterű Poisson-eloszlások ( $\mathbb{P}_\lambda(X = k) = \lambda^k / k! \cdot e^{-\lambda}$ )

## Poisson-eloszlás paraméterének becslése

Tegyük fel, hogy  $X_1, X_2, \dots, X_n$  független, azonos  $\lambda$  paraméterű Poisson-eloszlású minta, ahol  $\lambda > 0$  ismeretlen paraméter,  $n = 95$ , és  $\bar{X} = 1,379$ .

Poisson-eloszlásnál:

$$\mathbb{P}_\lambda(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda}; \quad \mathbb{E}(X_j) = \lambda; \quad D(X_j) = \sqrt{\lambda}.$$

A megfigyelések az alábbiak (a gólok száma összesen  $\sum_{j=1}^n X_j = 131$ ):

3, 0, 2, 2, 1, 3, ..., 2.

## Poisson-eloszlás paraméterének becslése

Tegyük fel, hogy  $X_1, X_2, \dots, X_n$  független, azonos  $\lambda$  paraméterű Poisson-eloszlású minta, ahol  $\lambda > 0$  ismeretlen paraméter,  $n = 95$ , és  $\bar{X} = 1,379$ .

Poisson-eloszlásnál:

$$\mathbb{P}_\lambda(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda}; \quad \mathbb{E}(X_j) = \lambda; \quad D(X_j) = \sqrt{\lambda}.$$

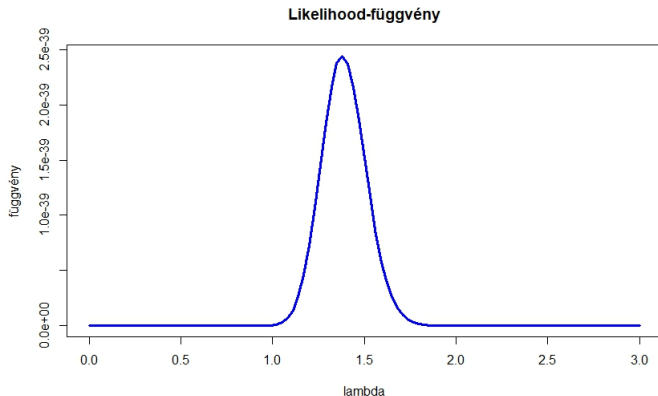
A megfigyelések az alábbiak (a gólok száma összesen  $\sum_{j=1}^n X_j = 131$ ):

$$3, \quad 0, \quad 2, \quad 2, \quad 1, \quad 3, \dots, 2.$$

Annak valószínűsége  $\lambda$  paraméter mellett, hogy éppen ezt a sorozatot kaptuk:

$$\begin{aligned} L_{95,\lambda}(3, 0, 2, \dots, 2) &= \frac{\lambda^3}{3!} e^{-\lambda} \cdot \frac{\lambda^0}{0!} e^{-\lambda} \cdot \frac{\lambda^2}{2!} e^{-\lambda} \cdot \dots \cdot \frac{\lambda^2}{2!} e^{-\lambda} = \\ &= \frac{\lambda^{3+0+2+2\dots+2}}{3! \cdot 0! \cdot 2! \cdot \dots \cdot 2!} e^{-95 \cdot \lambda} = \frac{\lambda^{131}}{3! \cdot 0! \cdot 2! \cdot \dots \cdot 2!} e^{-95 \cdot \lambda} \end{aligned}$$

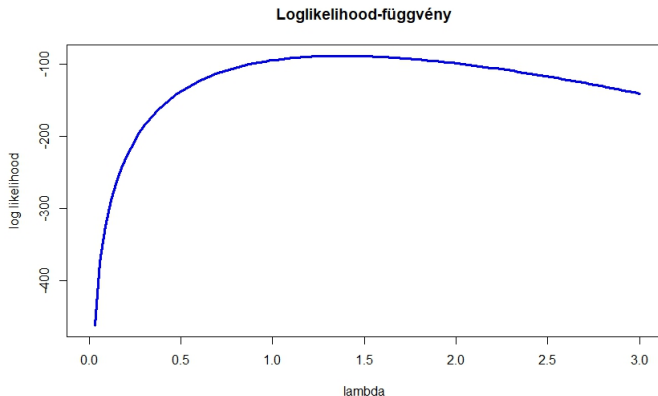
# Poisson-eloszlás paraméterének becslése



A  $\frac{\lambda^{131}}{3! \cdot 0! \cdot 2! \cdot \dots \cdot 2!} e^{-95 \cdot \lambda}$  likelihoodfüggvény a  $\lambda > 0$  paraméter függvényében;

mintaátlag:  $\bar{X} = \frac{131}{95} = 1,379$

## Poisson-eloszlás paraméterének becslése



A  $\log \frac{\lambda^{131}}{31 \cdot 0! \cdot 2! \cdot \dots \cdot 2!} e^{-95 \cdot \lambda}$  loglikelihoodfüggvény a  $\lambda > 0$  paraméter függvényében;

mintaátlag:  $\bar{X} = \frac{131}{95} = 1,379$

## ML-becslés: Poisson-eloszlás

$X_1, \dots, X_n$  függetlenek, Poisson-eloszlás  $\lambda > 0$  ismeretlen paraméterrel, azaz

$$\mathbb{P}(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k = 0, 1, 2, \dots).$$

Ekkor

$$L_{n,\lambda}(X_1, \dots, X_n) = \prod_{j=1}^n \left( \frac{\lambda^{X_j}}{X_j!} e^{-\lambda} \right) = \frac{\lambda^{X_1}}{X_1!} e^{-\lambda} \cdot \frac{\lambda^{X_2}}{X_2!} e^{-\lambda} \cdot \dots \cdot \frac{\lambda^{X_n}}{X_n!} e^{-\lambda}.$$

## ML-becslés: Poisson-eloszlás

$X_1, \dots, X_n$  függetlenek, Poisson-eloszlás  $\lambda > 0$  ismeretlen paraméterrel, azaz

$$\mathbb{P}(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k = 0, 1, 2, \dots).$$

Ekkor

$$L_{n,\lambda}(X_1, \dots, X_n) = \prod_{j=1}^n \left( \frac{\lambda^{X_j}}{X_j!} e^{-\lambda} \right) = \frac{\lambda^{X_1}}{X_1!} e^{-\lambda} \cdot \frac{\lambda^{X_2}}{X_2!} e^{-\lambda} \cdot \dots \cdot \frac{\lambda^{X_n}}{X_n!} e^{-\lambda}.$$

$$L_{n,\lambda}(X_1, \dots, X_n) = \lambda^{\sum_{j=1}^n X_j} e^{-n\lambda} \cdot \prod_{j=1}^n \frac{1}{X_j!}.$$

## ML-becslés: Poisson-eloszlás

$X_1, \dots, X_n$  függetlenek, Poisson-eloszlás  $\lambda > 0$  ismeretlen paraméterrel, azaz

$$\mathbb{P}(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k = 0, 1, 2, \dots).$$

Ekkor

$$L_{n,\lambda}(X_1, \dots, X_n) = \prod_{j=1}^n \left( \frac{\lambda^{X_j}}{X_j!} e^{-\lambda} \right) = \frac{\lambda^{X_1}}{X_1!} e^{-\lambda} \cdot \frac{\lambda^{X_2}}{X_2!} e^{-\lambda} \cdot \dots \cdot \frac{\lambda^{X_n}}{X_n!} e^{-\lambda}.$$

$$L_{n,\lambda}(X_1, \dots, X_n) = \lambda^{\sum_{j=1}^n X_j} e^{-n\lambda} \cdot \prod_{j=1}^n \frac{1}{X_j!}.$$

$$\log L_{n,\lambda}(X_1, \dots, X_n) = \log \lambda \sum_{j=1}^n X_j - n\lambda + \log \prod_{j=1}^n \frac{1}{X_j!}$$

## ML-becslés: Poisson-eloszlás

$X_1, \dots, X_n$  függetlenek, Poisson-eloszlás  $\lambda > 0$  ismeretlen paraméterrel, azaz

$$\mathbb{P}(X_j = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad (k = 0, 1, 2, \dots).$$

Ekkor

$$L_{n,\lambda}(X_1, \dots, X_n) = \prod_{j=1}^n \left( \frac{\lambda^{X_j}}{X_j!} e^{-\lambda} \right) = \frac{\lambda^{X_1}}{X_1!} e^{-\lambda} \cdot \frac{\lambda^{X_2}}{X_2!} e^{-\lambda} \cdot \dots \cdot \frac{\lambda^{X_n}}{X_n!} e^{-\lambda}.$$

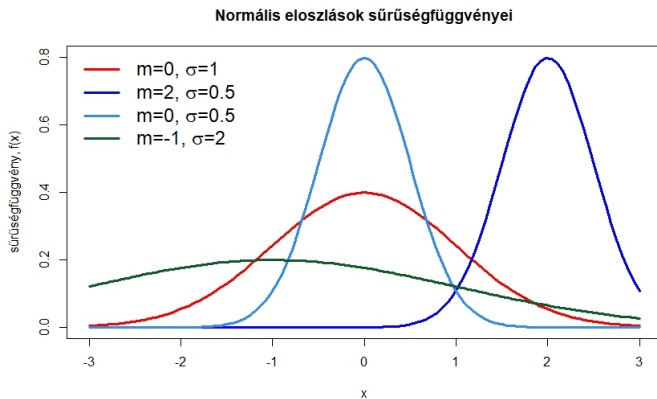
$$L_{n,\lambda}(X_1, \dots, X_n) = \lambda^{\sum_{j=1}^n X_j} e^{-n\lambda} \cdot \prod_{j=1}^n \frac{1}{X_j!}.$$

$$\log L_{n,\lambda}(X_1, \dots, X_n) = \log \lambda \sum_{j=1}^n X_j - n\lambda + \log \prod_{j=1}^n \frac{1}{X_j!}$$

$$\frac{\partial}{\partial \lambda} \log L_{n,\lambda}(X_1, \dots, X_n) = \frac{\sum_{j=1}^n X_j}{\lambda} - n > 0 \Leftrightarrow \lambda < \bar{X}.$$

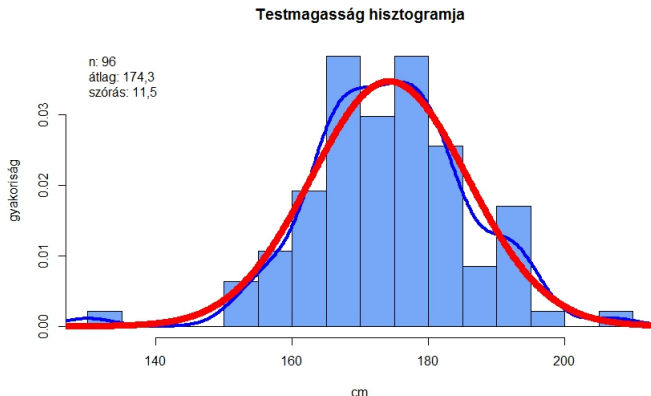
Ezért az ML-becslés:  $\hat{\lambda} = \bar{X}$ . A log monoton növekedését használtuk.

# Normális eloszlás sűrűségfüggvénye



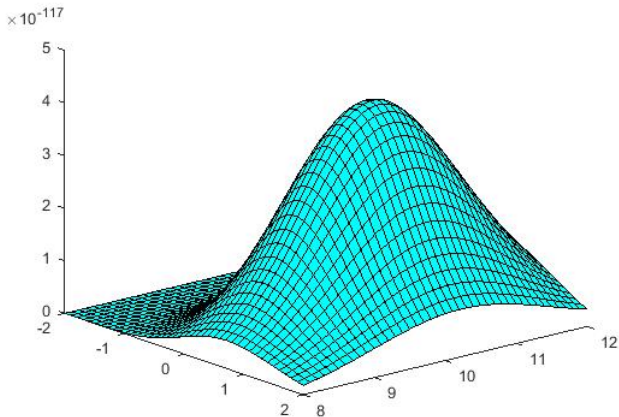
Különböző várható értékű és szórású normális eloszlások sűrűségfüggvénye

# Testmagasság hisztogramja



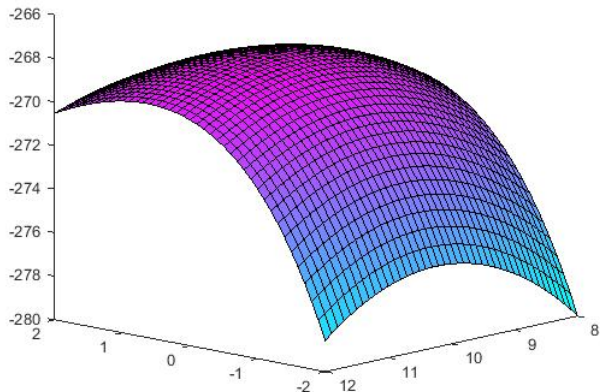
A testmagasság hisztogramja  $n = 96$  elemű mintából és az  $\bar{X} = 174,3$  várható értékű és  $s_n^* = 11,5$  szórású normális eloszlás sűrűségfüggvénye (pirossal).

# Likelihoodfüggvény



$n = 94$  elemű minta testmagasság-adatok alapján, normális eloszlást feltételezve.  
Az átlag:  $\bar{X} = 174,8$ , a tapasztalati szórás  $s_n = 10,5$ .

# Log-likelihoodfüggvény



$n = 94$  elemű minta testmagasság-adatok alapján, normális eloszlást feltételezve.  
Az átlag:  $\bar{X} = 174,8$ , a tapasztalati szórás  $s_n = 10,5$ .

## ML-becslés: normális eloszlás

$X_1, \dots, X_n$  függetlenek, eloszlásuk normális eloszlás  $m, \sigma > 0$  paraméterekkel. Ekkor

$$L_{n,m,\sigma}(X_1, \dots, X_n) = \prod_{j=1}^n f_{j,\vartheta}(X_j) = \prod_{j=1}^n \left[ \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X_j - m)^2}{2\sigma^2}\right) \right].$$

## ML-becslés: normális eloszlás

$X_1, \dots, X_n$  függetlenek, eloszlásuk normális eloszlás  $m, \sigma > 0$  paraméterekkel. Ekkor

$$L_{n,m,\sigma}(X_1, \dots, X_n) = \prod_{j=1}^n f_{j,\vartheta}(X_j) = \prod_{j=1}^n \left[ \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X_j - m)^2}{2\sigma^2}\right) \right].$$

$$L_{n,m,\sigma}(X_1, \dots, X_n) = \left( \frac{1}{\sqrt{2\pi}\sigma} \right)^n \exp\left(-\sum_{j=1}^n \frac{(X_j - m)^2}{2\sigma^2}\right).$$

## ML-becslés: normális eloszlás

$X_1, \dots, X_n$  függetlenek, eloszlásuk normális eloszlás  $m, \sigma > 0$  paraméterekkel. Ekkor

$$L_{n,m,\sigma}(X_1, \dots, X_n) = \prod_{j=1}^n f_{j,\vartheta}(X_j) = \prod_{j=1}^n \left[ \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(X_j - m)^2}{2\sigma^2}\right) \right].$$

$$L_{n,m,\sigma}(X_1, \dots, X_n) = \left(\frac{1}{\sqrt{2\pi}\sigma}\right)^n \exp\left(-\sum_{j=1}^n \frac{(X_j - m)^2}{2\sigma^2}\right).$$

$$\log L_{n,m,\sigma}(X_1, \dots, X_n) = -n \log(\sqrt{2\pi}) - n \log \sigma - \sum_{j=1}^n \frac{(X_j - m)^2}{2\sigma^2}.$$

Rögzített  $\sigma$  mellett ez akkor maximális, ha a harmadik tagban  $\sum_{j=1}^n (X_j - m)^2 = \sum_{j=1}^n X_j^2 - 2 \sum_{j=1}^n X_j m + nm^2$  minimális  $\Rightarrow \hat{m} = \bar{X}$ .

## ML-becslés: normális eloszlás

$$\log L_{n,\sigma}(X_1, \dots, X_n) = -n \log(\sqrt{2\pi}) - n \log \sigma - \sum_{j=1}^n \frac{(X_j - \bar{X})^2}{2\sigma^2}.$$

A  $\sigma$  szerinti parciális derivált:

$$\frac{\partial}{\partial \sigma} \log L_{n,\sigma}(X_1, \dots, X_n) = -\frac{n}{\sigma} + \sum_{j=1}^n \frac{(X_j - \bar{X})^2}{\sigma^3}.$$

Ez pontosan akkor pozitív, ha  $\sigma^2 < \frac{1}{n} \sum_{j=1}^n (X_j - \bar{X})^2 = s_n^2$ .

Tehát az ML-becslés:

$$\hat{m} = \bar{X}; \quad \hat{\sigma} = s_n = \sqrt{\left( \frac{1}{n} \sum_{j=1}^n X_j^2 \right) - \bar{X}^2}.$$

Tehát normális eloszlásnál az  $m$  paraméter becslése a mintaátlag, a szórásé a tapasztalati szórás.

## Az ML-becslés tulajdonságai

- Nem minden statisztikai mezőn létezik ML-becslés.
- Az ML-becslés nem feltétlenül egyértelmű.
- Az ML-becslés nem feltétlenül torzítatlan.

## Az ML-becslés tulajdonságai

- Nem minden statisztikai mezőn létezik ML-becslés.
- Az ML-becslés nem feltétlenül egyértelmű.
- Az ML-becslés nem feltétlenül torzítatlan.
- A  $\psi(\vartheta)$  függvény ML-becslése  $\psi(\hat{\vartheta})$ , ahol  $\hat{\vartheta}$  ML-becslés  $\vartheta$ -ra.

## Az ML-becslés tulajdonságai

- Nem minden statisztikai mezőn létezik ML-becslés.
- Az ML-becslés nem feltétlenül egyértelmű.
- Az ML-becslés nem feltétlenül torzítatlan.
- A  $\psi(\vartheta)$  függvény ML-becslése  $\psi(\hat{\vartheta})$ , ahol  $\hat{\vartheta}$  ML-becslés  $\vartheta$ -ra.
- **Megfelelő feltételek** (erős regularitási feltételek mellett) az ML-becslés **aszimptotikusan torzítatlan**: a várható értéke  $\vartheta$ -hoz tart  $n \rightarrow \infty$  esetén, **aszimptotikusan hatásos** (vagyis minimális szórású) és **aszimptotikusan normális eloszlású**, azaz  $\sqrt{n}(\hat{\vartheta}_n - \vartheta)$  normális eloszláshoz konvergál eloszlásban  $n \rightarrow \infty$  esetén (a  $\mathbb{P}_\vartheta$  valószínűségekre vonatkozóan).

# Az ML-becslés tulajdonságai

- Nem minden statisztikai mezőn létezik ML-becslés.
- Az ML-becslés nem feltétlenül egyértelmű.
- Az ML-becslés nem feltétlenül torzítatlan.
- A  $\psi(\vartheta)$  függvény ML-becslése  $\psi(\hat{\vartheta})$ , ahol  $\hat{\vartheta}$  ML-becslés  $\vartheta$ -ra.
- **Megfelelő feltételek** (erős regularitási feltételek mellett) az ML-becslés **aszimptotikusan torzítatlan**: a várható értéke  $\vartheta$ -hoz tart  $n \rightarrow \infty$  esetén, **aszimptotikusan hatásos** (vagyis minimális szórású) és **aszimptotikusan normális eloszlású**, azaz  $\sqrt{n}(\hat{\vartheta}_n - \vartheta)$  normális eloszláshoz konvergál eloszlásban  $n \rightarrow \infty$  esetén (a  $\mathbb{P}_\vartheta$  valószínűségekre vonatkozóan).
- Az alábbi egyenlet a maximumlikelihood-egyenlet:

$$\frac{\partial}{\partial \vartheta} \log L_{n,\vartheta}(X_1, \dots, X_n) = 0.$$

Megfelelő feltételek mellett az ML-becslés a maximumlikelihood-egyenlet megoldása (ha az ML-becslés nem számítható ki, de az egyenlet megoldható, gyakran az egyenlet megoldásával helyettesítik az ML-becslést).

## ML-becslés: egyenletes eloszlás

Ebben az esetben az ML-becslés nem számítható ki az ML-egyenlet gyökeként, vagyis nem kapható meg deriválással.

$X_1, \dots, X_n$  függetlenek, eloszlásuk egyenletes eloszlás az  $[a, b]$  intervallumon. Ekkor

$$L_{n,a,b}(X_1, \dots, X_n) = \prod_{j=1}^n f_{j,\vartheta}(X_j) = \prod_{j=1}^n \mathbb{I}(a \leq X_j \leq b) \cdot \frac{1}{b-a}.$$

$$L_{n,a,b}(X_1, \dots, X_n) = \left(\frac{1}{b-a}\right)^n \mathbb{I}(a \leq \min_j X_j \text{ és } \max_j X_j \leq b).$$

## ML-becslés: egyenletes eloszlás

Ebben az esetben az ML-becslés nem számítható ki az ML-egyenlet gyökeként, vagyis nem kapható meg deriválással.

$X_1, \dots, X_n$  függetlenek, eloszlásuk egyenletes eloszlás az  $[a, b]$  intervallumon. Ekkor

$$L_{n,a,b}(X_1, \dots, X_n) = \prod_{j=1}^n f_{j,\vartheta}(X_j) = \prod_{j=1}^n \mathbb{I}(a \leq X_j \leq b) \cdot \frac{1}{b-a}.$$

$$L_{n,a,b}(X_1, \dots, X_n) = \left(\frac{1}{b-a}\right)^n \mathbb{I}(a \leq \min_j X_j \text{ és } \max_j X_j \leq b).$$

Az első tényező legyen minél nagyobb (vagyis  $b - a$  minél kisebb) úgy, hogy a második tényező nem nulla. Ebből:

$$\hat{a} = \min_j X_j; \quad \hat{b} = \max_j X_j.$$

## Maximum-likelihood becslések

- binomiális eloszlás ismert  $k$  renddel:  $\hat{p} = \bar{X}/k$
- Poisson-eloszlás:  $\hat{\lambda} = \bar{X}$
- geometriai eloszlás:  $\hat{p} = 1/\bar{X}$
- normális eloszlás:  $\hat{m} = \bar{X}, \hat{\sigma} = s_n$
- exponenciális eloszlás:  $\hat{\lambda} = 1/\bar{X}$
- egyenletes eloszlás:  $\hat{a} = \min_j X_j; \quad \hat{b} = \max_j X_j$

## ML-becslés: exponenciális eloszlás

$X_1, \dots, X_n$  függetlenek, eloszlásuk exponenciális eloszlás  $\lambda > 0$  paraméterrel. Ekkor

$$L_{n,\lambda}(X_1, \dots, X_n) = \prod_{j=1}^n f_{j,\lambda}(X_j) = \prod_{j=1}^n \left[ \lambda \exp(-\lambda X_j) \mathbb{I}(X_j > 0) \right].$$

$$L_{n,\lambda}(X_1, \dots, X_n) = \lambda^n \exp\left(-\lambda \sum_{j=1}^n X_j\right).$$

## ML-becslés: exponenciális eloszlás

$X_1, \dots, X_n$  függetlenek, eloszlásuk exponenciális eloszlás  $\lambda > 0$  paraméterrel. Ekkor

$$L_{n,\lambda}(X_1, \dots, X_n) = \prod_{j=1}^n f_{j,\lambda}(X_j) = \prod_{j=1}^n \left[ \lambda \exp(-\lambda X_j) \mathbb{I}(X_j > 0) \right].$$

$$L_{n,\lambda}(X_1, \dots, X_n) = \lambda^n \exp \left( -\lambda \sum_{j=1}^n X_j \right).$$

$$\log L_{n,\lambda}(X_1, \dots, X_n) = n \log \lambda - \lambda \sum_{j=1}^n X_j$$

## ML-becslés: exponenciális eloszlás

$X_1, \dots, X_n$  függetlenek, eloszlásuk exponenciális eloszlás  $\lambda > 0$  paraméterrel. Ekkor

$$L_{n,\lambda}(X_1, \dots, X_n) = \prod_{j=1}^n f_{j,\lambda}(X_j) = \prod_{j=1}^n \left[ \lambda \exp(-\lambda X_j) \mathbb{I}(X_j > 0) \right].$$

$$L_{n,\lambda}(X_1, \dots, X_n) = \lambda^n \exp \left( -\lambda \sum_{j=1}^n X_j \right).$$

$$\log L_{n,\lambda}(X_1, \dots, X_n) = n \log \lambda - \lambda \sum_{j=1}^n X_j$$

$$\frac{\partial}{\partial \lambda} \log L_{n,\lambda}(X_1, \dots, X_n) = \frac{n}{\lambda} - n\bar{X} > 0 \Leftrightarrow \lambda < 1/\bar{X} \Rightarrow \hat{\lambda} = 1/\bar{X}.$$

## Házi feladat március 11., kedd, 10:15-ig

Ismételjük meg ezerszer az alábbi kísérletet: sorsolunk 100 független,  $m = 5$  várható értékű és  $\sigma = 2$  szórású normális eloszlású valószínűségi változót, és kiszámítjuk az  $s_n$  tapasztalati szórást.

Készítsünk hisztogramot az így kapott ezer megfigyelésből, számítsuk ki a minta-átlagot és a szórást is.

Ismételjük meg a kísérletet egy másik, tetszőlegesen választott  $\sigma$  szórású normális eloszlással is, és készítsük el így is az ábrát.

A maximum likelihood becslés általános tulajdonságai közül melyikeket figyelhetjük meg az ábrákon, azok alapján?