Elégséges statisztika (5. előadás)

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

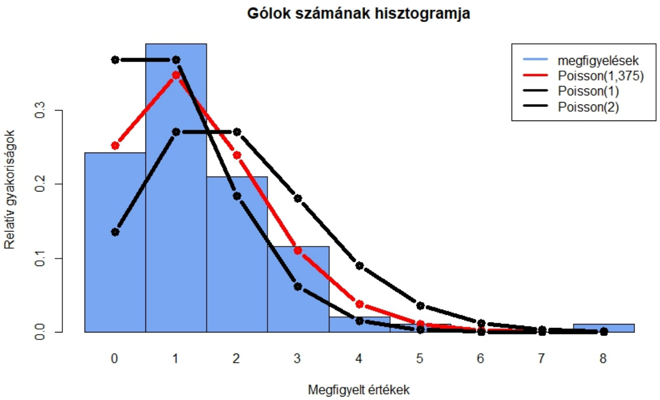
Találhatunk-e olyan általános módszereket, melyeket alkalmazva a ϑ ismeretlen paraméter becslésére egy megfelelő $T(X_1, \ldots, X_n)$ statisztika

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

Ha igen, mi az a **minimális információ a mintából**, amire szükségünk van?

Illetve, mennyire lehet kicsi a szórás, ha a mintából származó információ "mennyisége" adott?

Poisson-eloszlás paraméterének becslése



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

Minta: $0, 2, 1, 2, 4, \ldots$

Házi feladat március 7., 10:30-ig

Tegyük fel, hogy az ismerőseink által az utóbbi két hétben nézett sorozatok száma Poisson-eloszlású, ismeretlen $\lambda > 0$ paraméterrel.

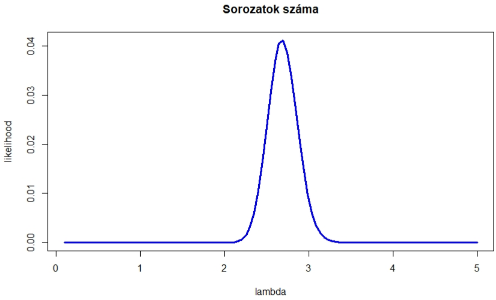
a) Határozzuk meg és ábrázoljuk a likelihood-függvényt, nézzük meg, hol lehet a maximuma.

b) Hasonlítsuk össze a $\hat{\lambda} = \overline{X}$ paraméterű Poisson-eloszlást a nézett sorozatok számának hisztogramjával.

```
sorozat=read.table("clipboard", header = FALSE, sep = "")
```

```
curve(exp(-length(sorozat$V1)*x)*xŝum(sorozat$V1), from=0.1, to=5,
lwd="3",
main="Sorozatok száma", xlab="lambda", ylab="likelihood", col="blue
```

Poisson-eloszlás paraméterének becslése



A sorozatok számának likelihood-függvénye, egy konstans szorzótól eltekintve:

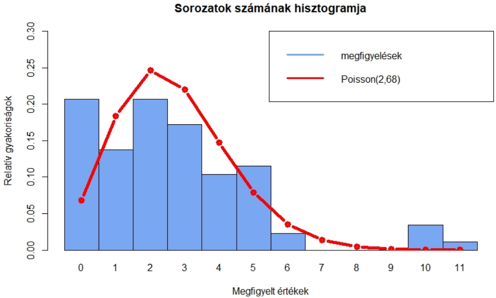
$$\frac{\lambda^{\sum_{j=1}^{n} X_j}}{\prod_{j=1}^{n} X_j!} e^{-\lambda n}$$

$n = 87$, $\sum_{j=1}^{n} X_j = 237$, és $\overline{X} = 2,68$

Poisson-eloszlás paraméterének becslése

```
pk=c(18,12,18,15,9,10,2,0,0,0,3, 1)/87

names(pk)<-c("0","1", "2", "3", "4", "5", "6","7","8", "9", "10", "1

barplot(pk, col="#79a7f2", xlab="Megfigyelt értékek",
ylab="Relatív gyakoriságok",main="Sorozatok számának hisztogramja",
space=0, ylim=c(0, 0.3))

x=c(0,1,2,3,4,5,6,7,8, 9, 10, 11)

y=dpois(x, lambda=2.68)

x<-x+0.5

lines(y∼ x, lwd="5", col="red", type="b")

legend("topright", c("megfigyelések", "Poisson(2,68)"), col=c("#79a
"red"), lwd="3")
```

Poisson-eloszlás paraméterének becslése



A sorozatok számának hisztogramja és Poisson-eloszlás $\hat{\lambda} = \overline{X} = 2,68$ becsült paraméterrel

$n = 87$, $\sum_{j=1}^{n} X_j = 237$, és $\overline{X} = 2,68$

Maximumlikelihood-módszer

Definíció (Likelihood-függvény)

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

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

Maximumlikelihood-módszer

Definíció (Likelihood-függvény)

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

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

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

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

Maximumlikelihood-módszer

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

Definíció (Maximum-likelihood becslés)

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

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

Poisson-eloszlás paraméterének becslése

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

Poisson-eloszlásnál:

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

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

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

Annak valószínűsége λ paraméter mellett, hogy éppen ezt a sorozatot kaptuk:

$$L_{95,\lambda}(3,0,2,\ldots,2) = \frac{\lambda^3}{3!}e^{-\lambda} \cdot \frac{\lambda^0}{0!}e^{-\lambda} \cdot \frac{\lambda^2}{2!}e^{-\lambda} \cdot \ldots \cdot \frac{\lambda^2}{2!}e^{-\lambda} =$$

$$= \frac{\lambda^{3+0+2+2\ldots+2}}{3! \cdot 0! \cdot 2! \cdot \ldots \cdot 2!}e^{-95\cdot\lambda} = \frac{{\color{red}\lambda}^{131}}{3! \cdot 0! \cdot 2! \cdot \ldots \cdot 2!}e^{-95\cdot{\color{red}\lambda}}$$

Poisson-eloszlás paraméterének becslése

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

Poisson-eloszlásnál:

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

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

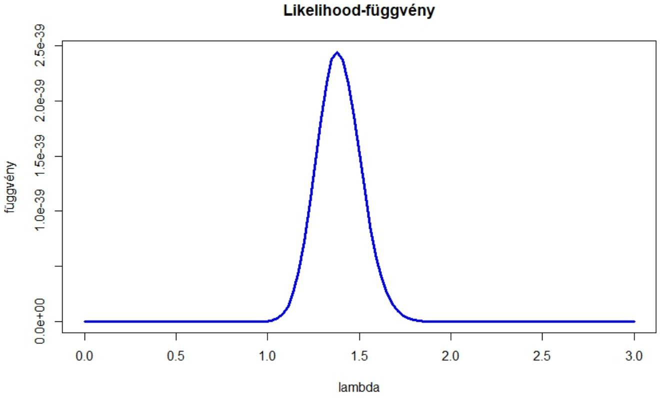
$$3, \quad 0, \quad 2, \quad 2, \quad 1, \quad 3, \dots, 2.$$

Annak valószínűsége λ paraméter mellett, hogy éppen ezt a sorozatot kaptuk:

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

A likelihood-függvényben csak n és a $\sum_{j=1}^{n} X_j$ összeg szerepel a paraméterrel együtt.

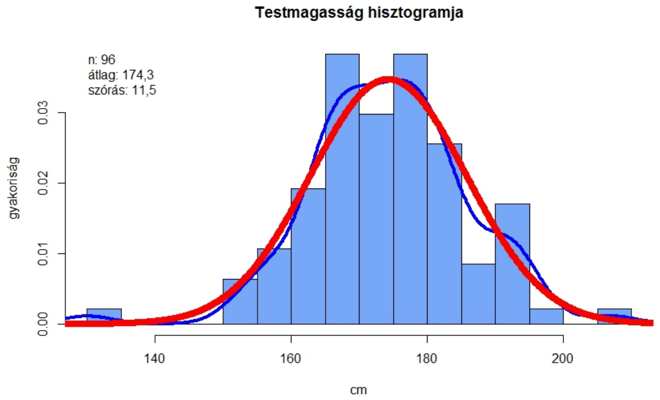
Poisson-eloszlás paraméterének becslése



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

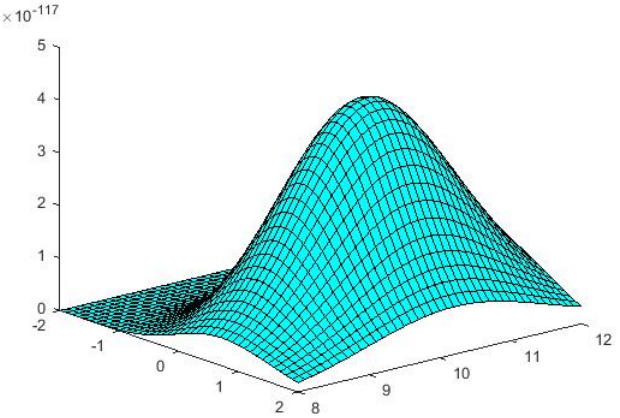
mintaátlag: $\overline{X} = \frac{131}{95} = 1,379$

Testmagasság hisztogramja



A testmagasság hisztogramja $n = 96$ elemű mintából és az $\overline{X} = 174,3$ 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: $\overline{X} = 174,8$, a tapasztalati szórás $s_n = 10,5$.

ML-becslés: normális eloszlás

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

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

ML-becslés: normális eloszlás

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

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

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

ML-becslés: normális eloszlás

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

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

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

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

ML-becslés: normális eloszlás

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

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

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

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

A likelihood-függvényben csak a $\sum_{j=1}^{n} X_j^2$ **négyzetösszeg** és a $\sum_{j=1}^{n} X_j$ **összeg** szerepel a mintából a paraméterrel együtt:

$$\sum_{j=1}^{n} (X_j - m)^2 = \sum_{j=1}^{n} X_j^2 - 2m \sum_{j=1}^{n} X_j + m^2.$$

Elégséges statisztika

Definíció

Legyen $X_1, \ldots, X_n$ *független minta* ϑ *paraméterű eloszlásból. A* T *statisztika elégséges, ha a likelihood-függvény felírható a következő alakban megfelelő* h *és* g *függvényekkel:*

$$L_{n,\vartheta}(X_1, \ldots, X_n) = h(X_1, \ldots, X_n) \cdot g_\vartheta(T(X_1, \ldots, X_n)).$$

Elégséges statisztika

Definíció

Legyen $X_1, \ldots, X_n$ független minta ϑ paraméterű eloszlásból. A T statisztika elégséges, ha a likelihood-függvény felírható a következő alakban megfelelő h és g függvényekkel:

$$L_{n,\vartheta}(X_1, \ldots, X_n) = h(X_1, \ldots, X_n) \cdot g_\vartheta(T(X_1, \ldots, X_n)).$$

- Poisson-eloszlás esetén az összeg és az átlag is elégséges statisztika.
- Normális eloszlásnál elégséges statisztika: $(\overline{X}, s_n)$, vagy $\left(\sum_{j=1}^n X_j^2, \sum_{j=1}^n X_j\right)$.
- Az elégséges statisztika nem egyértelmű.
- Ha létezik maximumlikelihood-becslés, T pedig elégséges statisztika, akkor az ML-becslés felírható $u(T(X_1, \ldots, X_n))$ alakban valamely u függvényre. Azaz, ahogy a példákban is láttuk, az elégséges statisztikából kifejezhető a maximumlikelihood-becslés.

Hatásos becslés az elégséges statisztika segítségével

A következő tétel alapján egy

- torzítatlan becslésből kiindulva
- az elégséges statisztika függvényeként, arra feltételes várható értéket véve.
- egy hatásosabb, vagyis kisebb szórású, szintén torzítatlan becslés is készíthető.

Például Poisson-eloszlás esetén, ha a λ paramétert szeretnénk becsülni:

- $\mathbb{E}_\lambda(X_1) = \lambda$, ezért X_1 torzítatlan becslése λ-nak;
- $\sum_{j=1}^{n} X_j$ elégséges statisztika

Hatásos becslés az elégséges statisztika segítségével

A következő tétel alapján egy

- torzítatlan becslésből kiindulva
- az elégséges statisztika függvényeként, arra feltételes várható értéket véve.
- egy hatásosabb, vagyis kisebb szórású, szintén torzítatlan becslés is készíthető.

Például Poisson-eloszlás esetén, ha a λ paramétert szeretnénk becsülni:

- $\mathbb{E}_\lambda(X_1) = \lambda$, ezért X_1 torzítatlan becslése λ-nak;
- $\sum_{j=1}^{n} X_j$ elégséges statisztika
- tekintsük az alábbi feltételes várható értéket (ha az összeget tudjuk, mennyi lehetett X_1:
$$\mathbb{E}(X_1|X_1 + \ldots + X_n) = \overline{X}$$
- továbbra is torzítatlan becslést kaptunk, de a szórás lényegesen kisebb lett:

$$\mathbb{E}_\lambda(\overline{X}) = \mathbb{E}_\lambda(X_1) = \lambda; \qquad D_\lambda(\overline{X}) = \frac{D_\lambda(X_1)}{\sqrt{n}} < D_\lambda(X_1).$$

Rao–Blackwell-tétel

Tétel (Rao–Blackwell)

Tegyük fel, hogy a T statisztika torzítatlan becslés a ϑ paraméterre, valamint S elégséges statisztika ϑ-ra. Ekkor megadható olyan T^* becslés, melyre

- $T^* = h(S)$ *megfelelő h függvénnyel;*
- T^* *torzítatlan ϑ-ra: $\mathbb{E}_\vartheta(T^*) = \vartheta$ minden $\vartheta \in \Theta$-ra;*
- T^* *hatásosabb, mint T: $D_\vartheta(T^*) \le D_\vartheta(T)$ minden $\vartheta \in \Theta$-ra.*

Ha az S statisztika teljes is (azaz minden olyan f függvény, melyre $\mathbb{E}_\vartheta(f(S)) = 0$ minden ϑ-ra, „majdnem mindenütt" nulla), akkor T^* hatásos, azaz minden torzítatlan becslésnél hatásosabb.

A megoldás az $\mathbb{E}(T|S)$ feltételes várható érték lesz. Például Poisson-eloszlásnál $\mathbb{E}(X_1 | \sum_{j=1}^n X_j) = \overline{X}$ hatásos.

Állítás

Poisson-eloszlásnál normális eloszlásoknál (m-re) a mintaelemek összege teljes elégséges statisztika, a mintaátlag pedig hatásos becslése a paraméternek.

Fisher-információ

Egy mintaelem **Fisher-információja**:

$$I_1(\vartheta) = \mathbb{E}_\vartheta\left(\left(\frac{\partial}{\partial\vartheta}\log f_\vartheta(X_1)\right)^2\right),$$

ahol f_ϑ a likelihoodfüggvény $\mathbb{P}_\vartheta$ mellett.

Fisher-információ

Egy mintaelem **Fisher-információja**:

$$I_1(\vartheta) = \mathbb{E}_\vartheta\left(\left(\frac{\partial}{\partial\vartheta}\log f_\vartheta(X_1)\right)^2\right),$$

ahol f_ϑ a likelihoodfüggvény $\mathbb{P}_\vartheta$ mellett.

- Megfelelő (regularitási) feltételek mellett a független azonos eloszlású, n elemű minta Fisher-információja: $I_n(\vartheta) = n \cdot I_1(\vartheta)$.
- Ha T elégséges statisztika, akkor $T(X_1, \ldots, X_n)$ Fisher-információja ugyanaz, mint $(X_1, \ldots, X_n)$ Fisher-információja (például Poisson-eloszlásnál az átlag Fisher-információja ugyanaz, mint a teljes mintáé).
- **Cramér–Rao-egyenlőtlenség**: megfelelő (regularitási) feltételek mellett, ha T torzítatlan becslés ϑ-ra, akkor

$$D^2(T(X)) \geq \frac{1}{I_n(\vartheta)} = \frac{1}{nI_1(\vartheta)}.$$

(Ebben az értelemben feleakkora szóráshoz négyszer annyi mintaelem kell.)

Fisher-információ

Egy mintaelem **Fisher-információja**:

$$I_1(\vartheta) = \mathbb{E}_\vartheta\left(\left(\frac{\partial}{\partial\vartheta}\log f_\vartheta(X_1)\right)^2\right),$$

ahol f_ϑ a likelihoodfüggvény $\mathbb{P}_\vartheta$ mellett.

Néhány nevezetes eloszlás Fisher-információja egy mintaelemből:

- binomiális eloszlás n renddel és p paraméterrel:
$$\frac{n}{p(1-p)}.$$
- Poisson-eloszlás λ paraméterrel: $1/\lambda$
- normális eloszlás, ha $\vartheta = (m, \sigma)$:
$$\begin{pmatrix} \frac{1}{\sigma^2} & 0 \\ 0 & \frac{2}{\sigma^2} \end{pmatrix}.$$

Fisher-információ: Poisson-eloszlás

Egy mintaelem **Fisher-információja** (az f_λ likelihood-függvény most a valószínűség):

$$I_1(\lambda) = \mathbb{E}_\lambda\left(\left(\frac{\partial}{\partial\lambda}\log f_\lambda(X_1)\right)^2\right) = \mathbb{E}_\lambda\left(\left(\frac{\partial}{\partial\lambda}\log\left(\frac{\lambda^{X_1}}{X_1!}e^{-\lambda}\right)\right)^2\right) =$$

$$= \mathbb{E}_\lambda\left(\left(\frac{\partial}{\partial\lambda}\left(X_1\log\lambda - \log X_1! - \lambda\right)\right)^2\right) = \mathbb{E}_\lambda\left(\left(\frac{X_1}{\lambda} - 1\right)^2\right) =$$

$$= \mathbb{E}_\lambda\left(\frac{X_1^2}{\lambda} - 2\frac{X_1}{\lambda} + 1\right) = \frac{\lambda+\lambda^2}{\lambda} - 2 + 1 = \frac{1}{\lambda}.$$

Felhasználtuk, hogy Poisson-eloszlásnál $\mathbb{E}_\lambda(X_1) = D_\lambda(X_1) = \lambda$, így

$$\mathbb{E}_\lambda(X_1^2) = D^2_\lambda(X_1) + \mathbb{E}^2_\lambda(X_1) = \lambda + \lambda^2.$$

Következmény a **Cramér–Rao-egyenlőtlenség** alapján: ha T torzítatlan becslése λ-nak n elemű független mintából, akkor

$$D^2(T(X_1,\ldots,X_n)) \geq \frac{1}{n\lambda}.$$

Konfidenciaintervallum az ML-becslés alapján

Gyakran az ismeretlen paraméterre egyetlen becslés helyett konfidenciaintervallumot adhatunk meg, amire például az igaz, hogy

$$\mathbb{P}_{\vartheta}(T_1 \leq \vartheta \leq T_2) \geq 0,95 \qquad (\vartheta \in \Theta)$$

. Ez egy olyan véletlen intervallum, ami 95%-os megbízhatóságú, azaz legalább ennyi valószínűséggel tartalmazza a becsülni kívánt paramétert.

Konfidenciaintervallum az ML-becslés alapján

Gyakran az ismeretlen paraméterre egyetlen becslés helyett konfidenciaintervallumot adhatunk meg, amire például az igaz, hogy

$$\mathbb{P}_\vartheta(T_1 \leq \vartheta \leq T_2) \geq 0,95 \qquad (\vartheta \in \Theta)$$

. Ez egy olyan véletlen intervallum, ami 95%-os megbízhatóságú, azaz legalább ennyi valószínűséggel tartalmazza a becsülni kívánt paramétert.

Állítás

Ha likelihoodfüggvény teljesít bizonyos regularitási feltételeket, akkor a ϑ paraméternek az $X_1, X_2, \ldots, X_n$ mintából számolt $\hat{\vartheta}_n$ maximumlikelihood-becslése

- *létezik;*
- *aszimptotikusan torzítatlan:* $\lim_{n\to\infty} \mathbb{E}_\vartheta(\hat{\vartheta}_n) = \vartheta$ *minden* $\vartheta \in \Theta$*-ra;*
- *aszimptotikusan hatásos:* $\lim_{n\to\infty} \sqrt{nI_1(\vartheta)}D_\vartheta(\hat{\vartheta}_n) = 1$ *minden* $\vartheta \in \Theta$*-ra;*
- *aszimptotikusan normális eloszlású:* $\sqrt{nI_1(\vartheta)}(\hat{\vartheta}_n - \vartheta)$ *eloszlásban tart a standard normális eloszláshoz minden* $\vartheta \in \Theta$*-ra* $n \to \infty$ *esetén.*

Itt $I_1(\vartheta)$ *az egyelemű mintából számolt Fisher-információt jelöli.*

Konfidenciaintervallum az ML-becslés alapján

Ez alapján **aszimptotikus konfidenciaintervallum**, ami $n \to \infty$ esetén $1-\alpha$-hoz tartó valószínűséggel tartalmazza ϑ-t:

$$\left(\hat{\vartheta}_n - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{1}{\sqrt{\hat{I}_n(\vartheta)}};\ \hat{\vartheta}_n + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{1}{\sqrt{\hat{I}_n(\vartheta)}}\right),$$

ahol $\hat{I}_n(\vartheta) = n\cdot I_1(\hat{\vartheta}_n)$, vagyis a Fisher-információ kifejezésébe a maximumlikelihood-becslést írjuk be.

Másképpen felírva:

$$\lim_{n\to\infty}\mathbb{P}_\vartheta\left(\hat{\vartheta}_n - \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{1}{\sqrt{\hat{I}_n(\vartheta)}} \le \vartheta \le \hat{\vartheta}_n + \Phi^{-1}\left(1-\frac{\alpha}{2}\right)\cdot\frac{1}{\sqrt{\hat{I}_n(\vartheta)}}\right) = 1-\alpha.$$

Itt is vegyük észre, hogy ϑ ismeretlen, de rögzített paraméter, és az intervallum alsó és felső végpontjai véletlenek, az eloszlásuk is ϑ-tól függ (ezt jelöli a ϑ az alsó indexben).

Bayes-becslés: bevezetés

Eddig a statisztikai elemzésekben a **frekventista hozzáállást** követtük. A $\vartheta \in \Theta$ paraméter számunkra ismeretlen, és csak azt tesszük fel róla, hogy egy ismert Θ halmaznak, a paramétertérnek az eleme. A Θ paramétertér rögzített halmaz saját struktúra nélkül, a tulajdonságoknak (például konzisztencia, torzítatlanság) minden $\vartheta \in \Theta$-ra teljesülniük kell. Ezért előzetes információt nem tudunk jól felhasználni.

Bayes-becslés: bevezetés

Eddig a statisztikai elemzésekben a **frekventista hozzáállást** követtük. A $\vartheta \in \Theta$ paraméter számunkra ismeretlen, és csak azt tesszük fel róla, hogy egy ismert Θ halmaznak, a paramétertérnek az eleme. A Θ paramétertér rögzített halmaz saját struktúra nélkül, a tulajdonságoknak (például konzisztencia, torzítatlanság) minden $\vartheta \in \Theta$-ra teljesülniük kell. Ezért előzetes információt nem tudunk jól felhasználni.

Ennek a módszernek egy hátránya például:

- egy gyógyszert n-szer kipróbálva 0 esetben lépett fel mellékhatás;
- legyen p a mellékhatás kialakulásának valószínűsége;
- ezt a relatív gyakorisággal becsülhetjük: $\hat{p} = 0$. Ez torzítatlan, konzisztens becslés, a frekventista hozzáállás követelményeinek megfelel.
- Ugyanakkor, ha $n = 20$ kipróbálásból nem alakult ki mellékhatás, az egész más, mintha $n = 2000$ kipróbálásnál sem találkoztunk ezzel. Ezt a becslés nem tükrözi.

Bayes-i hozzáállás: a paramétert magát is valószínűségi változónak tekintjük, ebbe beépítve valamilyen előzetes információt.

A priori és a posteriori eloszlás: példa

Példa. Egy biztosító háromféle ügyfelet különböztet meg:

- az óvatos, átlagos és merész sofőröket;
- az, hogy egy ügyfél hány balesetet okoz egy év alatt, Poisson-eloszlású, $0,1$, $0,2$ és $0,3$ paraméterekkel az egyes csoportoknál
- egy új ügyfél mindhárom csoportba azonos, $1/3$ valószínűséggel tartozik
- az éveket függetlennek tekintjük
- Péter az utóbbi tíz évben egyáltalán nem okozott balesetet
- adjunk becslést a Péter által egy év alatt okozott balesetek számát leíró Poisson-eloszlás λ paraméterére

A priori és a posteriori eloszlás: példa

Itt a λ ismeretlen paraméter valószínűségi változónak tekinthető, aminek az eloszlásáról az előzetes feltételezésünk:

$$\mathbb{P}(\lambda = 0,1) = \mathbb{P}(\lambda = 0,2) = \mathbb{P}(\lambda = 0,3) = \frac{1}{3}.$$

A paraméter eloszlásáról alkotott előzetes, a minta megismerése előtt feltételezett eloszlás az **a priori eloszlás**. Ez tükrözi a paraméterre vonatkozó, a megfigyelésnél korábbi információkat.

Legyen A az az esemény, hogy Péter tíz év alatt nem okozott balesetet. Ennek valószínűsége, ha s a rá jellemző eloszlás igazi paramétere:

$$\mathbb{P}(A|\lambda = s) = (e^{-s})^{10} = e^{-10s},$$

a Poisson-eloszlás definícióját $k = 0$-val használva.

A priori és a posteriori eloszlás: példa

Hogyan módosul az egyes lehetséges értékek valószínűsége, ha felhasználjuk, hogy *A* bekövetkezett?

A Bayes-tétel alapján

$$\mathbb{P}(\lambda = 0,1|A) = \frac{\mathbb{P}(A|\lambda = 0,1)\mathbb{P}(\lambda = 0,1)}{\sum_{j=1}^{3} \mathbb{P}(A|\lambda = j \cdot 0,1)\mathbb{P}(\lambda = j \cdot 0,1)}$$

A priori és a posteriori eloszlás: példa

Hogyan módosul az egyes lehetséges értékek valószínűsége, ha felhasználjuk, hogy A bekövetkezett?

A Bayes-tétel alapján

$$\mathbb{P}(\lambda = 0,1|A) = \frac{\mathbb{P}(A|\lambda = 0,1)\mathbb{P}(\lambda = 0,1)}{\sum_{j=1}^{3} \mathbb{P}(A|\lambda = j \cdot 0,1)\mathbb{P}(\lambda = j \cdot 0,1)}$$

$$\mathbb{P}(\lambda = 0,1|A) = \frac{e^{-10\cdot 0,1} \cdot \frac{1}{3}}{\left(e^{-10\cdot 0,1} + e^{-10\cdot 0,2} + e^{-10\cdot 0,3}\right) \cdot \frac{1}{3}} = \frac{e^{-1}}{e^{-1} + e^{-2} + e^{-3}} = 66,5\%.$$

A priori és a posteriori eloszlás: példa

Hogyan módosul az egyes lehetséges értékek valószínűsége, ha felhasználjuk, hogy A bekövetkezett?

A Bayes-tétel alapján

$$\mathbb{P}(\lambda = 0,1 | A) = \frac{\mathbb{P}(A|\lambda = 0,1)\mathbb{P}(\lambda = 0,1)}{\sum_{j=1}^{3} \mathbb{P}(A|\lambda = j \cdot 0,1)\mathbb{P}(\lambda = j \cdot 0,1)}$$

$$\mathbb{P}(\lambda = 0,1|A) = \frac{e^{-10\cdot 0,1} \cdot \frac{1}{3}}{\left(e^{-10\cdot 0,1} + e^{-10\cdot 0,2} + e^{-10\cdot 0,3}\right) \cdot \frac{1}{3}} = \frac{e^{-1}}{e^{-1} + e^{-2} + e^{-3}} = 66,5\%.$$

Hasonlóképpen:

$$\mathbb{P}(\lambda = 0,2|A) = \frac{e^{-2}}{e^{-1} + e^{-2} + e^{-3}} = 24,5\%$$

$$\mathbb{P}(\lambda = 0,3|A) = \frac{e^{-3}}{e^{-1} + e^{-2} + e^{-3}} = 9\%.$$

A feltételes valószínűségek tehát eltérnek az a priori eloszlástól, ott mindegyik $1/3$ volt, most az óvatos a legvalószínűbb.

A priori és a posteriori eloszlás: példa

A priori eloszlás: a paraméterre vonatkozó előzetes információt tükrözi, a mintát nem használja. A példában:

$$\mathbb{P}(\lambda = 0,1) = \mathbb{P}(\lambda = 0,2) = \mathbb{P}(\lambda = 0,3) = \frac{1}{3}.$$

A posteriori eloszlás: a paraméter feltételes eloszlása a mintára, megfigyelésekre nézve. A példában:

$$\mathbb{P}(\lambda = 0,1|A) = 66,5\%; \quad \mathbb{P}(\lambda = 0,2|A) = 24,5\%; \quad \mathbb{P}(\lambda = 0,3|A) = 9\%.$$

Ez még nem ad meg egyetlen számot becslésként, ebben a példában ilyet nem is fogunk meghatározni.

Bayes-becslés

Példa. Tegyük fel, hogy egy biztosító minden ügyfeléről azt feltételezi, hogy az általa egy év alatt okozott balesetek száma Poisson-eloszlású, paramétere pedig az ügyfélre jellemző λ, mely egy ügyfelet tekintve egyenletes eloszlású a $[0, 1]$ intervallumon. Adjunk becslést a Péterre jellemző λ paraméterre, ha Péter az utóbbi tíz évben egyetlen balesetet sem okozott.

Bayes-becslés

Példa. Tegyük fel, hogy egy biztosító minden ügyfeléről azt feltételezi, hogy az általa egy év alatt okozott balesetek száma Poisson-eloszlású, paramétere pedig az ügyfélre jellemző λ, mely egy ügyfelet tekintve egyenletes eloszlású a $[0,1]$ intervallumon. Adjunk becslést a Péterre jellemző λ paraméterre, ha Péter az utóbbi tíz évben egyetlen balesetet sem okozott.

Az alábbi egyenlet ugyanúgy érvényes, mint eddig, minden $s \in [0,1]$-re:

$$\mathbb{P}(A|\lambda = s) = (e^{-s})^{10} = e^{-10s}.$$

A Bayes-tétel (folytonos változata) azonban most ezt adja:

$$p^*(s|A) = \frac{\mathbb{P}(A|\lambda = s)p(s)}{\int_{-\infty}^{\infty} \mathbb{P}(A|\lambda = t)p(t)dt},$$

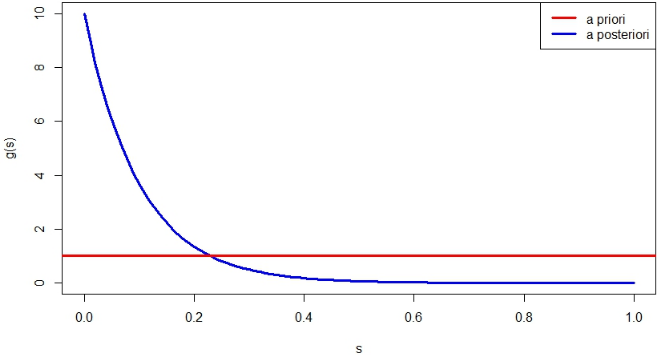
ahol p a paraméter a priori sűrűségfüggvénye. A példában $p(t) = \mathbb{I}(0 \leq t \leq 1)$, hiszen azt tettük fel előzetesen, hogy t egyenletes eloszlású a $[0,1]$ intervallumon.

Bayes-becslés

Behelyettesítve megkapjuk az a posteriori eloszlás sűrűségfüggvényét:

$$p^*(s|A) = \frac{\mathbb{P}(A|\lambda = s)p(s)}{\int_{-\infty}^{\infty} \mathbb{P}(A|\lambda = t)p(t)dt} = \frac{e^{-10s}\mathbb{I}(0 \le s \le 1)}{\int_0^1 e^{-10t}dt} = \frac{e^{-10s}\mathbb{I}(0 \le s \le 1)}{\frac{1}{10}(1 - e^{-10})}$$

Vagyis a megfigyelések alapján s-t olyan eloszlásúnak gondolhatjuk, aminek ez a fenti függvény a sűrűségfüggvénye.



Házi feladat március 21., hétfő, 10:30-ig

Legyenek $X_1, X_2, \ldots, X_n$ független valószínűségi változók, a sűrűségfüggvénye mindegyiknek

$$f(x) = \frac{\lambda^3 x^2}{2} e^{-\lambda x} \mathbb{I}(x > 0),$$

ahol $\mathbb{I}()$ értéke 1, ha igaz a benne szereplő feltétel, 0 különben.

Adjunk meg minél kevesebb, legfeljebb 3 elemből álló elégséges statisztikát ($X_1, X_2, \ldots$
maga mindig elégséges statisztika, de ez n elemből áll, nem jó most).