Momentummódszer (5. előadás)

Legyen $X_1, \ldots, X_n$ független azonos eloszlású minta.

1. Az eloszlás k. momentuma, ha ϑ az ismeretlen paraméter: $\mu_{k,\vartheta} = \mathbb{E}_\vartheta(X_1^k)$.
2. Legyen $\hat{\mu}_k = \frac{1}{n}\sum_{j=1}^n X_j^k$ az eloszlás k. tapasztalati momentuma.
3. Írjuk fel az alábbi egyenleteket a legkisebb olyan k-ig, amire az egyenletrendszer egyértelműen meghatározza ϑ-t (**bár nincs mindig ilyen k**):

$$\mathbb{E}_\vartheta(X_1) = \frac{1}{n}\sum_{j=1}^n X_j;$$

$$\mathbb{E}_\vartheta(X_1^2) = \frac{1}{n}\sum_{j=1}^n X_j^2;$$

$$\ldots$$

$$\mathbb{E}_\vartheta(X_1^k) = \frac{1}{n}\sum_{j=1}^n X_j^k.$$

4. A ϑ momentummódszerrel kapott becslése az a $\hat{\vartheta}$, ami megoldása a fenti egyenletrendszernek. **Nem mindig létezik, nem mindig egyértelmű, nem feltétlenül hatásos.**

Momentummódszer: Poisson– és exponenciális eloszlás

$X_1, \ldots, X_n$ független **Poisson-eloszlásúak** ismeretlen $\lambda > 0$ paraméterrel. A $k = 1$-hez tartozó egyenlet:

$$\mathbb{E}_\lambda(X_1) = \overline{X}.$$

Mivel a λ paraméterű Poisson-eloszlás várható értéke λ:

$$\hat{\lambda} = \overline{X}.$$

Momentummódszer: Poisson– és exponenciális eloszlás

$X_1, \ldots, X_n$ független **Poisson-eloszlásúak** ismeretlen $\lambda > 0$ paraméterrel. A $k = 1$-hez tartozó egyenlet:

$$\mathbb{E}_\lambda(X_1) = \overline{X}.$$

Mivel a λ paraméterű Poisson-eloszlás várható értéke λ:

$$\hat{\lambda} = \overline{X}.$$

$X_1, \ldots, X_n$ független **exponenciális** eloszlásúak ismeretlen $\lambda > 0$ paraméterrel. A $k = 1$-hez tartozó egyenlet:

$$\mathbb{E}_\lambda(X_1) = \frac{1}{\lambda} = \frac{1}{n} \sum_{j=1}^{n} X_j = \overline{X}.$$

Ez egyértelműen oldható meg λ-ra:

$$\hat{\lambda} = \frac{1}{\overline{X}}.$$

Momentummódszer: normális eloszlás

$X_1, \ldots, X_n$ független $N(m, \sigma^2)$ eloszlású minta (azaz normális eloszlású m várható értékkel és σ szórással).

A $k = 1$-hez és $k = 2$-höz tartozó egyenletek:

$$\mathbb{E}_{m,\sigma}(X_1) = m = \overline{X};$$
$$\mathbb{E}_{m,\sigma}(X_1^2) = \sigma^2 + m^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2.$$

A másodikba beírva az elsőt: $\sigma^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2 - \overline{X}^2 = s_n^2$ (a tapasztalati szórásnégyzet). Tehát az első két egyenlet együtt egyértelműen oldható meg, a momentummódszerrel kapott becslés:

$$\hat{m} = \overline{X}; \qquad \hat{\sigma} = s_n.$$

Az egyenletes eloszlás várható értéke és szórása

Az egyenletes eloszlás sűrűségfüggvénye: $f(x) = \frac{1}{b-a}$, ha $a \leq x \leq b$, és 0 különben.

A várható értéke:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x) dx = \int_a^b \frac{x}{b-a} dx = \left[\frac{x^2}{2(b-a)} \right]_{x=a}^{b}$$
$$= \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

Az egyenletes eloszlás várható értéke és szórása

Az egyenletes eloszlás sűrűségfüggvénye: $f(x) = \frac{1}{b-a}$, ha $a \leq x \leq b$, és 0 különben.

A várható értéke:

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_a^b \frac{x}{b-a}dx = \left[\frac{x^2}{2(b-a)}\right]_{x=a}^b$$
$$= \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

A négyzetének a várható értéke:

$$\mathbb{E}(X^2) = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx = \int_a^b \frac{x^2}{b-a}dx = \frac{1}{b-a}\left[\frac{x^3}{3}\right]_{x=a}^b$$
$$= \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}.$$

Az egyenletes eloszlás várható értéke és szórása

Az egyenletes eloszlás sűrűségfüggvénye: $f(x) = \frac{1}{b-a}$, ha $a \leq x \leq b$, és 0 különben.

A várható értéke:

$$\mathbb{E}(X) = \frac{a+b}{2}.$$

A négyzetének a várható értéke:

$$\mathbb{E}(X^2) = \frac{a^2+ab+b^2}{3}.$$

A szórásnégyzete:

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

Momentummódszer: egyenletes eloszlás

Legyen $X_1, \ldots, X_n$ független minta az $[a, b]$ intervallumon egyenletes eloszlásból. Ennek várható értéke $(a+b)/2$, szórása $(b-a)/\sqrt{12}$. Ezek alapján a $k = 1$-hez és $k = 2$-höz tartozó egyenlet:

$$\mathbb{E}_{a,b}(X_1) = \frac{a+b}{2} = \overline{X};$$

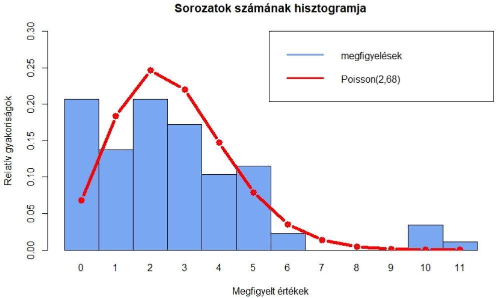
$$\mathbb{E}_{a,b}(X_1^2) = \frac{(b-a)^2}{12} + \left(\frac{a+b}{2}\right)^2 = \frac{1}{n}\sum_{j=1}^{n} X_j^2.$$

A másodikba beírva az elsőt: $\frac{(b-a)^2}{12} = \frac{1}{n}\sum_{j=1}^{n} X_j^2 - \overline{X}^2 = s_n^2$, amiből

$$\hat{a} = \overline{X} - \sqrt{3}s_n; \qquad \hat{b} = \overline{X} + \sqrt{3}s_n.$$

Hátránya: előfordulhat, hogy ezek nem is lehetséges értékek: ha $\hat{a}$ nagyobb a legkisebb megfigyelésnél, vagy $\hat{b}$ kisebb a legnagyobb megfigyelésnél.

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:

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

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

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 $\sum_{j=1}^{n} X_j = 131$):

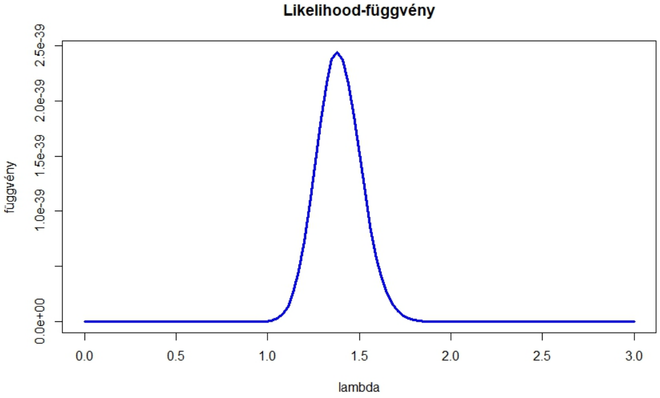
$$3, \quad 0, \quad 2, \quad 2, \quad 1, \quad 3, \ldots, 2.$$

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

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

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

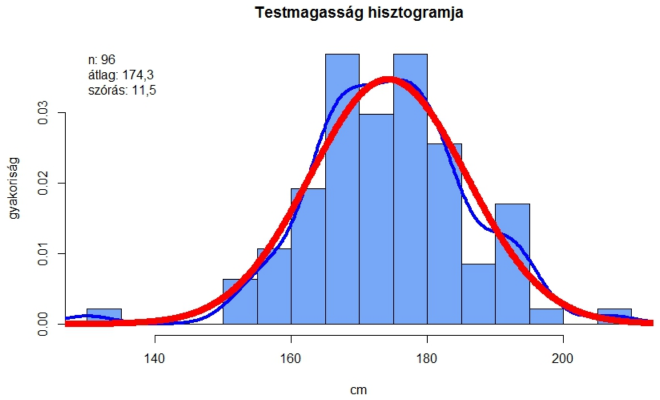
Poisson-eloszlás paraméterének becslése



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

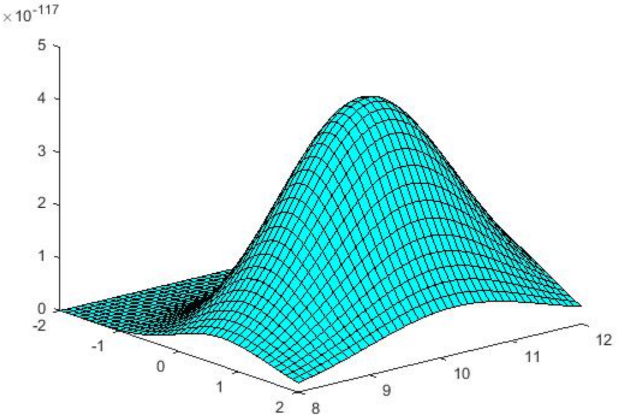
mintaátlag: $\overline{X} = \frac{131}{95} = 1,379$

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

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

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

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

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

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: $D_\vartheta(T^*) \leq D_\vartheta(T)$ 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):

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

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

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

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

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

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).

Házi feladat április 4., kedd, 12:00-ig

Tegyük fel, hogy $X_1, X_2, \dots, X_n$ független valószínűségi változók, melyek sűrűségfüggvénye:

$$f(x) = \begin{cases} \frac{8}{3}x^3 e^{-2x}, & \text{ha} x > 0; \\ 0, & \text{különben.} \end{cases}$$

Írjuk fel a likelihood-függvényt és keressünk elégséges statisztikát.