Maximumlikelihood-módszer

Definíció (Likelihood-függvény)

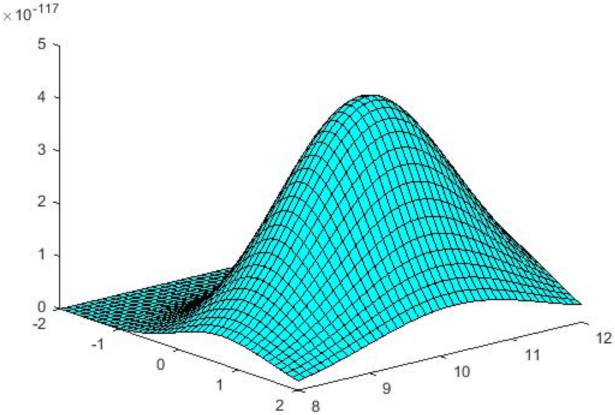
Ha az $(Y_1, \dots, Y_n)$ *független minta diszkrét (a lehetséges értékeinek száma véges vagy megszámlálható sok), akkor a likelihood-függvénye:*

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

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:*

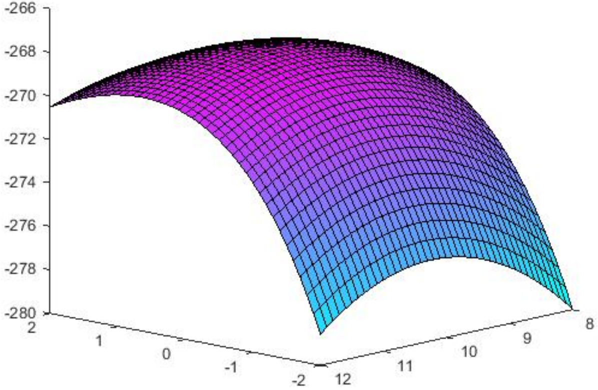
$$L_{n,\vartheta}(t_1, \dots, t_n) = \prod_{j=1}^{n} f_{j,\vartheta}(t_j) \qquad (t_1, \dots, t_n \in \mathbb{R}).$$

Likelihoodfüggvény



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

Log-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

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

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

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

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

ML-becslés: normális eloszlás

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

A σ szerinti parciális derivált:

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

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

Tehát az ML-becslés:

$$\hat{m} = \overline{X}; \qquad \hat{\sigma} = s_n = \sqrt{\frac{1}{n}\sum_{j=1}^{n} X_j^2 - \overline{X}^2}.$$

Tehát normális eloszlásnál az m paraméter becslése a mintaátlag, a szórásé a tapasztalati szórás. Ebben a speciális esetben a momentum módszer és az ML-módszer ugyanazt adja eredményül.

ML-becslés: egyenletes eloszlás

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

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

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

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

ML-becslés: egyenletes eloszlás

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

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

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

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

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

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

Maximum-likelihood becslések

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

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 ϑ-ra.
- Az alábbi egyenlet a maximumlikelihood-egyenlet:

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

  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).

A maximum-likelihood becslés tulajdonságai

Ha likelihoodfüggvény teljesít bizonyos regularitási feltételeket, akkor a ϑ paraméternek az $X_1, X_2, \dots, X_n$ 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.

Momentumok

Definíció

Az X valószínűségi változó k. momentuma:

$$\mathbb{E}(X^k) \qquad (k \geq 1),$$

ha ez a várható érték létezik.

Ha az X valószínűségi változó diszkrét, és lehetséges értékei: $x_1, x_2, \ldots$, akkor

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

Ha az X valószínűségi változó abszolút folytonos és sűrűségfüggvénye f, akkor

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

Momentummódszer

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:

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