

11. Gyakorlat

Alapstatisztikák, paraméterbecslés, MLE

Alapstatisztikák

Definíció (Minta és statisztika) $\underline{X}_n = (X_1, \dots, X_n)$ egy n elemű mintának nevezzük, ha $X_1, \dots, X_n$ független azonos eloszlású (f.a.e.) valószínűségi változók. A minta egy realizációja $\underline{x}_n = (x_1, \dots, x_n) \in \mathbb{R}^n$. A $T_n : \mathbb{R}^n \rightarrow \mathbb{R}$ statisztika a minta egy függvénye: $T_n(\underline{X}_n)$.

Definíció (Mintaátlag, empirikus variancia, korrigált empirikus variancia) Az $\underline{X}_n = (X_1, \dots, X_n)$ minta esetén alapvető statisztikák

- A mintaátlag

$$\bar{X}_n = \frac{1}{n} \sum_{k=1}^n X_k.$$

- Az empirikus variancia és szórás

$$S_n^2 = \frac{1}{n} \sum_{k=1}^n (X_k - \bar{X}_n)^2 \quad \text{és} \quad S_n = \sqrt{S_n^2}.$$

- A korrigált empirikus variancia és szórás

$$S_n^{*2} = \frac{1}{n-1} \sum_{k=1}^n (X_k - \bar{X}_n)^2 \quad \text{és} \quad S_n^* = \sqrt{S_n^{*2}}.$$

Egy konkrét realizáción felvett értékek jelölése:

$$\bar{x}_n = \frac{1}{n} \sum_{k=1}^n x_k, \quad s_n^2 = \frac{1}{n} \sum_{k=1}^n (x_k - \bar{x}_n)^2, \quad s_n = \sqrt{s_n^2} \quad s_n^{*2} = \frac{1}{n-1} \sum_{k=1}^n (x_k - \bar{x}_n)^2, \quad s_n^* = \sqrt{s_n^{*2}}.$$

Paraméterbecslés

Definíció (Torzítatlan statisztika) Legyen $\underline{X}_n = (X_1, \dots, X_n)$ minta egy θ paraméterű eloszláscsaládból. Azt mondjuk, hogy a T_n statisztika torzítatlan $f(\theta)$ -ra, ha

$$\mathbb{E}[T_n(\underline{X}_n)] = f(\theta).$$

Állítás Legyen $\underline{X}_n = (X_1, \dots, X_n)$ minta X egy θ paraméteres eloszlásából. Ekkor $\bar{X}_n$ torzítatlan becslés $\mathbb{E}_\theta(X)$ -re, S_n^{*2} torzítatlan becslés $\mathbb{D}_\theta^2(X)$ -re és S_n^* torzítatlan becslés $\mathbb{D}_\theta(X)$ -re.

Definíció (Maximum likelihood becslés - MLE) Legyen $\underline{X}_n = (X_1, \dots, X_n)$ minta egy θ paraméterű eloszláscsaládból. Ekkor θ maximum likelihood becslése az a $\hat{\theta}$, amire a likelihood függvénye $\mathcal{L}_\theta : \mathbb{R}^n \rightarrow \mathbb{R}_0^+$ maximális (amennyiben létezik) az adott $\underline{x}_n = (x_1, \dots, x_n)$ realizáción.

- Amennyiben az eloszlás diszkrét

$$\mathcal{L}_\theta(x_1, \dots, x_n) = \prod_{k=1}^n p_\theta(x_k), \quad \text{ahol } p_\theta \text{ az eloszlás súlyfüggvénye.}$$

- Amennyiben az eloszlás folytonos

$$\mathcal{L}_\theta(x_1, \dots, x_n) = \prod_{k=1}^n f_\theta(x_k), \quad \text{ahol } f_\theta \text{ az eloszlás sűrűségfüggvénye.}$$

Megjegyzés Mivel $\ln$ monoton függvény, a maximumhelyhez elég a log-likelihood függvény maximumát megkeresni. A log-likelihood függvény

$$l_\theta = \ln(\mathcal{L}_\theta).$$

Feladatok

1. Legyen X eloszlása az alábbi egy ismeretlen θ paraméter esetén

$$\mathbb{P}_\theta(X=0) = \frac{2\theta}{3}, \quad \mathbb{P}_\theta(X=2) = \frac{1-\theta}{3}, \quad \mathbb{P}_\theta(X=3) = \frac{2-\theta}{3}.$$

Tekintsük az $\underline{x}_5 = (3, 0, 2, 0, 2)$ mintarealizációt az eloszlásból.

- (a) Adjunk a realizáció alapján torzítatlan becslést X várható értékére és szórására.
 - (b) Adjunk a realizáció alapján maximum likelihood becslést θ -ra.
2. Egy $0 < \theta < 1$ valószínűséggel sikeres kísérletet az első sikerig ismétlünk egymástól függetlenül.
- (a) Adjunk maximum likelihood becslést θ -ra tetszőleges realizáció esetén.
 - (b) Az $\underline{x}_{10} = (4, 3, 4, 1, 10, 6, 5, 2, 2, 3)$ 10 darab megfigyelés alapján adjunk becslést θ -ra.
3. Legyen X eloszlása egy ismeretlen $0 < \theta < 1$ paraméter esetén az alábbi.

$$\mathbb{P}_\theta(X=k) = (k-1)\theta^2(1-\theta)^{k-2}, \quad k = 2, 3, 4, \dots$$

- (a) Mi az eloszlása X -nek?
 - (b) Adjunk maximum likelihood becslést θ -ra tetszőleges realizáció esetén.
4. Legyen $\underline{X}_n = (X_1, \dots, X_n)$ egy n elemű f.a.e. minta, ahol $\theta > 0$ paraméter esetén a mintaelemek sűrűségfüggvénye.

$$f_\theta(x) = \begin{cases} \theta^2 x e^{-\theta x} & \text{ha } x > 0 \\ 0 & \text{egyébként} \end{cases}.$$

Adjunk maximum likelihood becslést θ -ra tetszőleges realizáció esetén.

5. Legyen $\underline{X}_n = (X_1, \dots, X_n)$ egy n elemű f.a.e. minta egy $\theta > 0$ paraméterű eloszlásból. Adjunk maximum likelihood becslést θ -ra tetszőleges realizáció esetén.
- (a) Amennyiben az eloszlás $\text{Exp}(\theta)$.
 - (b) Amennyiben az eloszlás $\text{Uni}(0, \theta)$.