Matematika A4 - Valószínűségszámítás 11. gyakorlat
Paraméterbecslés

$\mathcal{X}_n = (X_1, X_2, \dots, X_n)$ n elemű minta alapján akarjuk az X valószínűségi változó θ paraméterét becsülni.

- Átlag: $\overline{\mathcal{X}_n} = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$ torzítatlan becslés $\mathbb{E}(X)$-re
- Tapasztalati szórás: $\hat{s} = \sqrt{\frac{1}{n}\sum\limits_{i=1}^{n}(X_i - \overline{\mathcal{X}_n})^2}$ torzított becslés $\mathbb{D}(X)$-re
- Torzítatlan becslés $\mathbb{D}(X)$-re: $\hat{s}^* = \sqrt{\frac{1}{n-1}\sum\limits_{i=1}^{n}(X_i - \overline{\mathcal{X}_n})^2}$

Maximum Likelihood Becslés

Azt a θ értéket akarjuk meghatározni, amire a látott $\mathcal{X}_n$ realizáció a legvalószínűbb. Ehhez maximalizálni kell θ-ban a $l(\mathcal{X}_n, \theta)$ **likelihood függvény**t:

- Diszkrét eset: $l(\mathcal{X}_n, \theta) = \prod\limits_{i=1}^{n} \mathbb{P}(X_i|\theta)$, ahol $\mathbb{P}(X_i|\theta)$ annak a valószínűsége, hogy ha az igazi paraméter θ, akkor X_i-t látunk
- Folytonos eset: $l(\mathcal{X}_n, \theta) = \prod\limits_{i=1}^{n} f(X_i|\theta)$, ahol $f(X_i|\theta)$ az X változó sűrűségfüggvénye X_i-ben, ha az igazi paraméter θ

Ehhez elég a $L(\mathcal{X}_n, \theta) = \log(l(\mathcal{X}_n, \theta))$ **log likelihood** függvényt maximalizálni. Vagyis keressük azt a $\hat{\theta}$ paramétert, amire $\frac{\partial}{\partial\theta} L\left(\mathcal{X}_n, \hat{\theta}\right) = 0$. Azaz

- Diszkrét: $\sum\limits_{i=1}^{n} \frac{\partial}{\partial\theta} \log\left(\mathbb{P}\left(X_i|\hat{\theta}\right)\right) = 0$
- Folytonos: $\sum\limits_{i=1}^{n} \frac{\partial}{\partial\theta} \log\left(f\left(X_i|\hat{\theta}\right)\right) = 0$

1. feladat Tegyük fel, hogy X súlyfüggvénye az ismeretlen θ paraméter függvényében az alábbi:

$$\mathbb{P}(X=0|\theta) = \frac{5\theta}{4}, \quad \mathbb{P}(X=2|\theta) = \frac{2-2\theta}{4}, \quad \mathbb{P}(X=4|\theta) = \frac{1-2\theta}{4}, \quad \mathbb{P}(X=5|\theta) = \frac{1-\theta}{4}.$$

Az alábbi mintát látjuk: $\mathcal{X}_8 = (0, 2, 4, 4, 5, 2, 5, 5)$.

(a) Mennyi a minta átlaga, tapasztalati szórása? Adjunk a szórásra torzítatlan becslést!

(b) Adjunk ML becslést θ paraméterre!

2. feladat Legyen $f(x) = a^2 x e^{-ax}$, ha $x \geq 0$. Adjunk ML becslést az a paraméterre tetszőleges $\mathcal{X}_n = (X_1, \dots, X_n)$ minta esetén!

3. feladat Egy berendezés élettartamára a következő mintánk van (években):

$$\mathcal{X}_{16} = (2, 3, 4, 2.5, 3, 3, 3.5, 4, 2, 4, 3, 3, 3.5, 2, 3, 4).$$

(a) Mennyi a minta átlaga?

(b) Adjunk ML becslést λ-ra, ha feltesszük, hogy a berendezés élettartamának eloszlása $\mathrm{EXP}(\lambda)$!

(c) Adjunk ML becslést μ-re, ha feltesszük, hogy a berendezés élettartamának eloszlása $N(\mu, 1)$!

4. feladat Egy cinkelt érmét dobálunk, ahol a fej valószínűsége p.

(a) A következő adatsorunk van az első fej megjelenésére:

$$\mathcal{X}_{16} = (2, 3, 1, 2, 3, 3, 2, 1, 4, 4, 4, 2, 3, 2, 1, 4).$$

Adjunk ML becslést p-re!

(b) Mi az ML becslés p-re tetszőleges $\mathcal{X}_n = (a_1, a_2, \ldots, a_n)$ minta esetén?

5. feladat Sűrűségfüggvények-e az alábbiak? Ha igen, adjuk meg a perem- és feltételes sűrűségfüggvényeket, várható értékeket, varianciákat illetve a kovarianciát!

(a) $f(x, y) = \frac{4}{5}(x + xy + y)$, ha $0 < x < 1, 0 < y < 1$, egyébként $f(x, y) = 0$

(b) $f(x, y) = \lambda^2 e^{-\lambda(x+y)}$, ha $x, y > 0$, egyébként $f(x, y) = 0$