Matematika A4 - Valószínűségszámítás 11. gyakorlat
Paraméterbecslés

$\mathcal{X}_n = (X_1, X_2, \ldots, X_n)$ 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!

Megoldásvázlat

(a) Átlag: $\overline{\mathcal{X}_8} = \frac{\sum\limits_{i=1}^{8} X_i}{8} = \frac{27}{8}$.

Tapasztalati szórás: $\hat{s} = \sqrt{\frac{1}{8}\sum\limits_{i=1}^{8}\left(X_i - \frac{27}{8}\right)^2} = \frac{\sqrt{191}}{8}$.

Torzítatlan becslés szórásra: $\hat{s}^* = \sqrt{\frac{1}{7}\sum\limits_{i=1}^{8}\left(X_i - \frac{27}{8}\right)^2} = \sqrt{\frac{n}{n-1}} \cdot \hat{s} = \sqrt{\frac{191}{56}}$.

(b) Először vegyük észre, hogy a valószínűségek nemnegativitása miatt $0 \leq \theta \leq \frac{1}{2}$. A log likelihood függvény:

$$L(\mathcal{X}_8, \theta) = \ln\left(\frac{5\theta}{4}\right) + 2\ln\left(\frac{2-2\theta}{4}\right) + 2\ln\left(\frac{1-2\theta}{4}\right) + 3\ln\left(\frac{1-\theta}{4}\right)$$

A $\frac{\partial}{\partial\theta}L(\mathcal{X}_8, \theta) = 0$ feltételből azt kapjuk, hogy

$$\frac{1}{\theta} - \frac{2}{1-\theta} - \frac{4}{1-2\theta} - \frac{3}{1-\theta} = 0.$$

Közös nevezőre hozás és a nevezővel szorzás után azt kapjuk, hogy $16\theta^2 - 12\theta + 1 = 0$. Ennek gyökei $\frac{3\pm\sqrt{5}}{8}$, amiből a nagyobb több, mint $\frac{1}{2}$, ami nem lehet hiszen ekkor $\mathbb{P}(X = 4|\theta) < 0$ lenne. Vagyis $\hat{\theta} = \frac{3-\sqrt{5}}{8}$.

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

Megoldásvázlat $\ln f(x|a) = 2\ln a + \ln x - ax$, vagyis $L(\mathcal{X}_n, a) = 2n\ln a + \sum_{i=1}^{n} \ln x_i - a\sum_{i=1}^{n} x_i$.
A $\frac{\partial}{\partial a}L(\mathcal{X}_n, a) = 0$ feltételből $\frac{2n}{a} - \sum_{i=1}^{n} x_i = 0$. Vagyis $\hat{a} = \frac{2n}{\sum_{i=1}^{n} x_i} = \frac{2}{\overline{\mathcal{X}_n}}$.

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 $N(\mu, 1)$!

Megoldásvázlat

(a) $\overline{\mathcal{X}_{16}} = \frac{49.5}{16} = \frac{99}{32}$

(b) $\ln f(x|\lambda) = \ln\lambda - \lambda x$, így $\frac{\partial}{\partial\lambda}f(x|\lambda) = \frac{1}{\lambda} - x$. Tehát $\frac{\partial}{\partial\lambda}L(\mathcal{X}_{16}, \lambda) = \sum_{i=1}^{16}\left(\frac{1}{\lambda} - x_i\right) = \frac{16}{\lambda} - 49,5$.

$\frac{\partial}{\partial\lambda}L(\mathcal{X}_{16}, \lambda) = 0$ feltételből $\hat{\lambda} = \frac{16}{49,5} = \frac{1}{\overline{\mathcal{X}_{16}}}$.

(c) $\ln f(x|\mu) = \ln\left(\frac{1}{\sqrt{2\pi}}\right) - \frac{(x-\mu)^2}{2}$, így $\frac{\partial}{\partial\mu}f(x|\mu) = x - \mu$. Tehát $\frac{\partial}{\partial\mu}L(\mathcal{X}_{16}, \mu) = \sum_{i=1}^{16}(x_i - \mu) = 49.5 - 16\mu$.
$\frac{\partial}{\partial\mu}L(\mathcal{X}_{16}, \mu) = 0$ feltételből $\hat{\mu} = \frac{49,5}{16} = \overline{\mathcal{X}_{16}}$.

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?

Megoldásvázlat Legyen X az első fej megjelenéséig a dobások száma. Ekkor $X \sim \mathrm{GEO}(p)$, azaz $\mathbb{P}(X=k|p) = (1-p)^{k-1} \cdot p$. Így $\ln(\mathbb{P}(X=k\,|\,p)) = (k-1)\ln(1-p) + \ln p$. Ezt deriválva

$$\frac{\partial}{\partial p} \ln\left(\mathbb{P}(X=k\,|\,p)\right) = -\frac{k-1}{1-p} + \frac{1}{p} = \frac{1-kp}{p(1-p)}.$$

(a) Tehát $\frac{\partial}{\partial p} L(\mathcal{X}_{16}, p) = \sum\limits_{i=1}^{16} \frac{1-k_i p}{p(1-p)} = \frac{16}{p(1-p)} - \frac{41}{1-p}$.

$\frac{\partial}{\partial p} L(\mathcal{X}_{16}, p) = 0$ feltételből $\hat{p} = \frac{16}{41}$.

(b) A fenti számolás alapján $\frac{\partial}{\partial p} L(\mathcal{X}_n, p) = \sum\limits_{i=1}^{n} \frac{1-k_i p}{p(1-p)} = \frac{n}{p(1-p)} - \frac{\sum\limits_{i=1}^{n} k_i}{1-p}$.

$\frac{\partial}{\partial p} L(\mathcal{X}_n, p) = 0$ feltételből $\hat{p} = \frac{n}{\sum\limits_{i=1}^{n} k_i} = \frac{1}{\overline{\mathcal{X}_n}}$.

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 $f(x,y)=0$

Megoldásvázlat

(a) Ellenőrizhető, hogy $\iint\limits_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$, így ez egy sűrűségfüggvény. Korábbi feladatokban látott képleteket használva számolhatjuk a kérdezett mennyiségeket. (A számolások felét megspórolhatjuk, ha észrevesszük, hogy az eloszlás szimmetrikus a változókban.)

- $f_X(x) = \frac{6x+2}{5}$, ha $0<x<1$, egyéb x esetén pedig $f_X(x)=0$.
  Hasonlóan (szimmetria miatt) $f_Y(y) = \frac{6y+2}{5}$, ha $0<y<1$, egyéb y esetén pedig $f_Y(y)=0$.
  $\mathbb{E}(X) = \mathbb{E}(Y) = \frac{3}{5}$
- Jól látszódik, hogy $f(x,y) = f_X(x)\cdot f_Y(y)$ nem teljesül minden x,y esetén, így X és Y nem független.
- $f_{X|Y}(x|y) = \frac{2(x+xy+y)}{3y+1}$, míg $f_{Y|X}(y|x) = \frac{2(x+xy+y)}{3x+1}$ (mindkettő $0<x<1$, $0<y<1$ esetén).
  $\mathbb{E}(X|Y=y) = \frac{5y+2}{9y+3}$, $\mathbb{E}(Y|X=x) = \frac{5x+2}{9x+3}$
  $\mathbb{E}\left(X^2|Y=y\right) = \frac{7y+3}{18y+6}$, $\mathbb{E}\left(Y^2|X=x\right) = \frac{7x+3}{18x+6}$
  $\mathbb{D}^2(X|Y=y) = \frac{13y^2+8y+1}{18(3y+1)^2}$, $\mathbb{D}^2(Y|X=x) = \frac{13x^2+8x+1}{18(3x+1)^2}$
- $\mathbb{E}(XY) = \frac{16}{45}$, $\mathrm{COV}(X,Y) = \frac{-1}{225}$.

(b) Ellenőrizhető, hogy $\iint\limits_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1$, így ez egy sűrűségfüggvény. Korábbi feladatokban látott képleteket használva számolhatjuk a kérdezett mennyiségeket. (A számolások felét megspórolhatjuk, ha észrevesszük, hogy az eloszlás szimmetrikus a változókban.)

- $f_X(x) = \lambda e^{-\lambda x}$, ha $0<x$ (egyéb x esetén pedig $f_X(x)=0$), vagyis $X \sim \mathrm{EXP}(\lambda)$. Hasonlóan $Y \sim \mathrm{EXP}(\lambda)$. Emiatt $\mathbb{E}(X) = \mathbb{E}(Y) = \frac{1}{\lambda}$.
- Jól látszódik, hogy $f(x,y) = f_X(x)\cdot f_Y(y)$ teljesül minden x,y esetén, így X és Y független.
- Függetlenség miatt a feltételes sűrűségfüggvények, várható értékek és szórásnégyzetek megegyeznek a peremeloszlás ($\mathrm{EXP}(\lambda)$) megfelelő mennyiségeivel:

$$f_{X|Y}(x|y) = \lambda e^{-\lambda x}, \quad f_{Y|X}(y|x) = \lambda e^{-\lambda y} \qquad 0<x, 0<y,$$

$$\mathbb{E}(X|Y=y) = \mathbb{E}(Y|X=x) = \frac{1}{\lambda},$$

$$\mathbb{D}^2(X|Y=y) = \mathbb{D}^2(Y|X=x) = \frac{1}{\lambda^2}.$$

- $\mathbb{E}(XY) \overset{\text{függetlenség}}{=} \mathbb{E}(X)\mathbb{E}(Y) = \frac{1}{\lambda^2}$, $\mathrm{COV}(X,Y) = 0$.