3. feladatsor, matematikai statisztika, 2024. február 29.

1. A $(0, \theta)$ intervallumon egyenletes eloszlásból vett n elemű minta alapján adjunk becslést a $\theta > 0$ paraméterre (több megoldás is lehetséges). Torzítatlan-e? Ha nem, tegyük azzá! Melyik becslésnek legkisebb a négyzetes hibája?

2. Tekintsük a következő diszkrét eloszlást:

$$P_\theta(X = k) = \binom{3}{k} \frac{(\theta - 1)^{3-k}}{\theta^3},\ k = 0, 1, 2, 3,$$

   ahol $\theta > 1$ ismeretlen paraméter. Mutassuk meg, hogy semmilyen n-re sem lehet n elemű mintából torzítatlanul becsülni θ-t!

3. A $(\theta, \theta + 1)$ intervallumon egyenletes eloszlás esetén becsüljük a θ valós paramétert egyetlen megfigyelés alapján a $T_c(X) = [X + c] - c$ statisztikával.
   (a) Mely c-re lesz ez torzítatlan?
   (b) Számítsuk ki a szórásnégyzetét!
   (c) Létezik-e hatásos becslés?

4. Torzítatlanok-e az alábbi becslések? Ha nem, gondolkodjunk azon, hogyan tehetnénk torzítatlanná.
   (a) Mintaátlag a várható értékre.
   (b) Tapasztalati szórásnégyzet a szórásnégyzetre.
   (c) A mintaátlag reciproka az exponenciális eloszlás paraméterére.
   (d) A mintaátlag reciproka a geometriai eloszlás paraméterére.

5. Egy ékszerműhelyben fülbevalókat készítenek, majd kettesével összecsomagolják őket. Mindegyik (darab) fülbevaló, egymástól függetlenül, p valséggel rendelkezik valamilyen apró hibával. Jelölje X_i, hogy az i-edik párban hány hibás van.
   a) Mutassuk meg, hogy $T = \sum_{i=1}^{n} X_i$ elégséges statisztika p-re!
   b) $n = 4$ esetben írjuk fel a minta feltételes eloszlását a $T = 3$ feltételre nézve!

A 2. feladatsor 4. feladatának megoldása:

A rendezett minta mindig elégséges. Az együttes sűrűségfüggvény

$$(1/a^n) \prod_{i=1}^{n} 1_{\{x_i/a \in [1,2]\}} = (1/a^n) 1_{\{\max_i x_i \le 2a\}} 1_{\{\min_i x_i \ge a\}}$$

amiből (X_1^*, X_n^*) elégséges statisztika.

A további három lehetőség egyike sem elégséges statisztika. Legyen $n = 2$ az egyszerűség kedvéért.

Ha X_1 elégséges lenne, akkor a $P_a(X_2 \in A | X_1)$ feltételes eloszlás nem függene a-tól. Mivel X_2 és X_1 függetlenek, ez az eloszlás éppen $P_a(X_2 \in A)$, ami nagyon is függ a-tól, hiszen a sűrűségfüggvénye $(1/a) 1_{\{a \le x \le 2a\}}$.

Ha $(X_1 + X_2)/2$ elégséges lenne, akkor $P_a(X_1 < 1/2 | (X_1 + X_2)/2 > 1/4)$ nem függene a-tól. Márpedig $a = 1$ esetén ez a valószínűség 0, míg $a = 1/3$ esetén $P_a(X_1 < 1/2 \text{ és } (X_1 + X_2)/2 > 1/4) \ge P_{1/3}(X_1 < 1/2, X_2 > 1/2) > 0$, hiszen X_1, X_2 függetlenek és egyenletesek $[1/3, 2/3]$-on.

Ha $X_n^* - X_1^*$, a terjedelem, lenne elégséges, akkor $P_a(X_1 < 1/2 | X_2^* - X_1^* > 1/10)$ nem függene a-tól. Márpedig $a = 1$ esetén ez a valószínűség 0, míg $a = 1/3$ esetén $P_a(X_1 < 1/2 \text{ és } X_2^* - X_1^* > 1/10) \ge P_{1/3}(X_1 < 1/2, X_2 > 6/10) > 0$.