Néhány megoldás 1., 2024

1. feladatsor, 8. feladat Legyen $X_1, X_2, \dots, X_n$ exponenciális eloszlásból vett minta, és $Y_1 = nX_1^*, Y_2 = (n-1)(X_2^* - X_1^*), \dots, Y_{n-1} = 2(X_{n-1}^* - X_{n-2}^*)$, $Y_n = X_n^* - X_{n-1}^*$.
(a) Mutassuk meg, hogy $Y_1, Y_2, \dots, Y_n$ eloszlása megegyezik $X_1, X_2, \dots, X_n$ eloszlásával.
(b) Határozzuk meg az $X_1^*, X_2^*, \dots, X_n^*$ rendezett minta várható értékét és kovarianciamátrixát.

Megoldási javaslat: Csak 1 paraméterű exponenciális eloszlásra nézzük meg. A $H = \{(x_1, \dots, x_n) : x_1 < \dots < x_n\}$ halmazt bevezetve tudjuk, hogy az $(X_1^*, \dots, X_n^*)$ együttes sfgvénye

$$n! e^{-x_1} \dots e^{-x_n} 1_H.$$

Viszont a $(Y_1, \dots, Y_n)$-re való áttérés deriváltmátrixa alsó háromszögmátrix, melyen a főátló elemeinek szorzata $n!$, tehát ez a determináns, amivel *osztani* kell amikor az $(Y_1, \dots, Y_n)$ együttes sfgvényét felírjuk. A H halmaz képe $(0, \infty)^n$, ezért tehát a keresett együttes sfgvény $y_i \in (0, \infty)$, $i = 1, \dots, n$-re:

$$e^{-y_1 - y_2 - \dots - y_n},$$

hiszen $x_1 + \dots + x_n = y_1 + \dots + y_n$ az új koordinátákra való áttérésnél.

Ezek után tudjuk, hogy $EY_i = 1$ minden i-re, valamint az Y_i-k függetlenek. Innen $EX_1^* = 1/n$, $EX_2^* = 1/n + 1/(n-1)$, stb. Megkaptuk a várható értékeket. A kovarianciákkal most nem vesződünk, de megnézzük a szórásnégyzeteket. $EY_i^2 = 2$, ennek fényében $E(X_1^*)^2 = 2/n^2$, tehát $D^2(X_1^*) = 1/n^2$. Függetlenségből

$$D^2(X_2^*) = D^2(Y_1/n + Y_2/(n-1)) = D^2(Y_1/n) + D^2(Y_2/(n-1)) = 1/n^2 + 1/(n-1)^2.$$

Etc.

4. feladatsor 4. feladat A $(0, \vartheta)$ intervallumon egyenletes eloszlásból vett n elemű minta esetén blackwellizáljuk a $\vartheta > 0$ paraméter $T(X) = (n+1)X_1^*$ becslését. Hatásos-e?

Megoldási javaslat: $S = X_n^*$ elégséges és teljes. Igen, a blackwellizált hatásos lesz, mert teljes elégséges statisztikára blackwellizálunk.

Legyenek $Z_1, \dots, Z_{n+1}$ független exponenciálisok. Akkor $U_k = (Z_1 + \dots + Z_k)/(Z_1 + \dots + Z_{n+1})$, $k = 1, \dots, n$ együttes eloszlása megegyezik $(X_1^*, \dots, X_n^*)$-éval, lásd az első feladatsor 3. feladatát. Innen látható, hogy az $U_k = c$ feltétel mellett (ami azt jelenti, hogy $Z_{n+1} = \frac{1-c}{c}[Z_1 + \dots + Z_n]$) az U_1^* feltételes eloszlása a $cZ_1/(Z_1 + \dots + Z_n)$ eloszlásával egyezik meg. Használva ezt (és megin az 1. feladatsor 3. feladatát), azt látjuk, hogy X_1^* feltételes eloszlása $X_n^* = c$ mellett ugyanaz, mint egy $(n-1)$ elemű $[0, c]$-n egyenletes minta minimumának eloszlása. Azaz a feltételes várható érték c/n. Innen adódik, hogy behelyettesítve $c = X_n^*$-ot, a blackwellizált $E[T|S] = X_n^* \frac{n+1}{n}$, ez a hatásos becslés.

4. feladatsor 5. feladat Tekintsük az $E(\theta, \theta+1)$ eloszláscsaládot és számoljuk ki a θ torzítatlan becslésének, $\lfloor X_1 \rfloor$-nak a blackwellizáltját (n elemű mintából)!

Megoldási javaslat: Ehhez az X_1 feltételes eloszlására van szükség az $X_1^* = c_1$, $X_n^* = c_2$ feltételek mellett, hiszen (X_1^*, X_n^*) elégséges statisztika. Vegyük észre, hogy ha a rendezett minta elemei $y_1, \dots, y_n$, akkor X_1 rájuk vonatkozó feltételes eloszlása egyenletes az $\{y_1, \ldots, y_n\}$ halmazon (mivel folytonos az eloszlás, feltehető, hogy ezek mind különböző értékek). Innen

$$P(X_1 = c_1 | X_1^* = c_1, X_n^* = c_2) = P(X_1 = c_2 | X_1^* = c_1, X_n^* = c_2) = 1/n.$$

Az 1. feladatsor 6. feladatát kétszer alkalmazva látjuk, hogy $(X_2^*, \ldots, X_{n-1}^*)$ feltételes eloszlása $X_1^* = c_1$, $X_n^* = c_2$ mellett ugyanaz, mint a $[c_1, c_2]$-n egyenletes eloszlásból vett $(n-2)$ elemű rendezett mintáé.

Kis kitérő: leellenőrizhető az 1. feladatsorból, hogy ha f_k a $[0, 1]$-en egyenletes rendezett minta k elemének sűrűségfüggvénye, akkor $\sum_{k=1}^{n} f_k = konstans$. Ezt a $[c_1, c_2]$-n alkalmazva kijön, hogy X_1 feltételes eloszlása $X_1^* = c_1, X_2^* = c_2, X_1 \neq X_1^*$, $X_1 \neq X_n^*$ feltételek mellett egyenletes a $[c_1, c_2]$, hiszen a (feltételes) sűrűségfüggvénye az $X_2^*, \ldots, X_{n-1}^*$ sűrűségfüggvényeinek azonos súlyokkal vett lineáris kombinációja, ezek összege pedig a fentiek szerint konstans.

Száz szónak is egy a vége, a kérdéses blackwellizált az $\lfloor X_1 \rfloor$ feltételes várható értéke, tehát:

$$\frac{\lfloor X_1^* \rfloor + \lfloor X_n^* \rfloor}{n} + \frac{\int_{X_1^*}^{X_n^*} \lfloor x \rfloor \, dx}{X_n^* - X_1^*} \frac{n-2}{n}.$$

Érdekesség, hogy ha $\lfloor X_1^* \rfloor = \lfloor X_n^* \rfloor$, akkor a blackwellizált meggyezik az eredeti becsléssel.