תאוריה סטטיסטית 1, 88-2 5

ד"ר ע. וישנה

מועד א', תשס"ז

פתרון

1. החיבור בשאלה זו הוא מודולו 2.

   (א) יהי $Y \sim b(\frac{1}{2})$ משתנה ברנולי. בווקטור $(X_1, \dots, X_n) = (Y, Y, \dots, Y)$ כל רכיב מתפלג כדרוש, בעוד שכל שני רכיבים הם תלויים.

   (ב) (כאשר $n = 3$) אם Y, Z משתני ברנולי בלתי-תלויים, אז $(Y, Z, Y + Z)$ הוא וקטור העונה על הדרישות.

   (ג) אם השלשה מערבת שלושה אינדקסים שונים, אז מתקבל וקטור מהצורה $\alpha + (R_i, R_j, R_k)$, שרכיביו בלתי תלויים לכל התפלגות של α; אחרת מדובר בשלשה כמו $(X_1 + R_1, X_2 + R_2, Y_1 + R_1)$, שרכיביה בלתי תלויים משום ש- X_1, Y_1 בלתי תלויים. לעומת זאת, ברביעיה $(X_1 + R_1, X_2 + R_2, Y_1 + R_1, Y_2 + R_2)$ סכום שני הרכיבים הראשונים שווה לסכום שני הרכיבים האחרונים, אם ניקח $X_1 = X_2$ ו- $Y_1 = Y_2$.

2. (א) בשיטת המומנטים עלינו להשוות $\bar{X} = \alpha + \frac{1}{2}\alpha\delta$ ו- $S^2 = \frac{1}{12}\alpha^2\delta^2$. לכן $\alpha\delta = \sqrt{12}S$, ומכאן $\hat{\alpha} = \bar{X} - \sqrt{3}S$, $\hat{\delta} = \frac{\sqrt{12}S}{\bar{X} - \sqrt{3}S}$.

   (ב) הצפיפות היא $f(x) = \frac{1}{\alpha\delta} I_{[\alpha, \alpha+\delta\alpha]}(x)$, ולכן הנראות היא $L(\vec{X}; \alpha, \delta) = \frac{1}{(\alpha\delta)^n} I_{[\alpha, \alpha+\delta\alpha]}(Y_1) I_{[\alpha, \alpha+\delta\alpha]}(Y_n)$.

   (ג) יש למזער את $\delta\alpha$, בכפוף לאילוצים $\alpha \le Y_1 \le Y_n \le \alpha + \delta\alpha$, שמהם נובע $\delta\alpha \ge Y_n - Y_1$. לכן נבחר $\hat{\delta} = \frac{Y_n - Y_1}{\alpha}$; כעת, האילוצים גוררים $\hat{\alpha} = Y_1$.

   (ד) בהנתן Y_1, Y_n, כל המשתנים מתפלגים באחידות תחת האילוץ $Y_1 \le X \le Y_n$, והם בלתי תלויים. התיאור הזה לא תלוי ב- α, δ.

   (ה) בסיכוי $\frac{1}{n}$, $X_k = Y_1$; בסיכוי $\frac{1}{n}$, $X_k = Y_n$; בסיכוי המשלים, $X_k \sim U(Y_1, Y_n)$.

3. (א) בתור רכיב של התפלגות נורמלית דו-ממדית, $X_i \sim N(0, \sigma^2)$.

   (ב) לפי סעיף א', הנראות של $(X_1, \dots, X_n)$ היא $L = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{X_i^2}{2\sigma^2}} = (2\pi\sigma^2)^{n/2} e^{-\frac{1}{2\sigma^2}\sum X_i^2}$. לכן $\log(L) = C - n\log(\sigma) - \frac{1}{2\sigma^2}\sum X_i^2$, והנגזרת היא $\frac{d}{d\sigma}\log(L) = -\frac{n}{\sigma} + \frac{1}{\sigma^3}\sum X_i^2 = \frac{n}{\sigma^3}(\frac{1}{n}\sum X_i^2 - \sigma^2)$. לפי משפט קרמר-ראו, $T_1 = \frac{1}{n}\sum X_i^2$ הוא UMVUE ל- σ^2. כנ"ל $T_2 = \frac{1}{n}\sum Y_i^2$.

   (ג) לפי ההגדרה של התפלגות רב-נורמלית, פונקצית הצפיפות היא

$$f(X, Y; \sigma, \rho) = \frac{\exp\left(-\frac{1}{2\sigma^2(1-\rho^2)}(X, Y)\begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix}\right)}{2\pi\sigma^2\sqrt{1-\rho^2}},$$

ולכן הנראות היא

$$L(\vec{X},\vec{Y};\sigma,\rho) = \frac{1}{(2\pi)^n\sigma^{2n}(1-\rho^2)^{n/2}}\exp\left(-\frac{\sum X_i^2 - 2\rho\sum X_iY_i + \sum Y_i^2}{2\sigma^2(1-\rho^2)}\right).$$

נסמן $R = \sum X_i^2 - 2\rho\sum X_iY_i + \sum Y_i^2$, ונגזור את $\log(L)$:

$$\begin{aligned}\frac{d}{d\sigma}\log(L) &= \frac{d}{d\sigma}\left(C - 2n\log(\sigma) - \frac{R}{2\sigma^2(1-\rho^2)}\right)\\ &= -\frac{2n}{\sigma} + \frac{R}{\sigma^3(1-\rho^2)}\\ &= \frac{2n}{\sigma^3}\left(\frac{R}{2(1-\rho^2)n} - \sigma^2\right).\end{aligned}$$

לפי משפט קרמר-ראו,

$$T = \frac{R}{2(1-\rho^2)n} = \frac{1}{2(1-\rho^2)}(T_1+T_2) - \frac{\rho}{1-\rho^2}\cdot\frac{1}{n}\sum X_iY_i$$

הוא UMVUE של σ^2.

(ד) השונות של $T_1 = \frac{1}{n}\sum X_i^2$ שווה ל- $\frac{1}{n}V(X_1^2) = \frac{2\sigma^4}{n}$, וכנ"ל ל- T_2. לעומת זאת, $T = \frac{1}{n}\frac{1}{2(1-\rho^2)}\sum_i\left(X_i^2 + Y_i^2 - 2\rho X_iY_i\right)$. שונות הביטוי שבסוגריים היא

$$\begin{aligned}& V(X^2 - 2\rho XY + Y^2)\\ =\ & E((X^2 - 2\rho XY + Y^2)^2) - E(X^2 - 2\rho XY + Y^2)^2\\ =\ & E(X^4 + Y^4 + 2(1+2\rho^2)X^2Y^2 - 4\rho X^3Y - 4\rho XY^3) - 4(1-\rho^2)^2\sigma^4\\ =\ & \left(3 + 3 + 2(1+2\rho^2)^2 - 12\rho^2 - 12\rho^2 - 4(1-\rho^2)^2\right)\sigma^4\\ =\ & 4(1-\rho^2)^2\sigma^4,\end{aligned}$$

ולכן $V(T) = \frac{4n(1-\rho^2)^2\sigma^4}{4n^2(1-\rho^2)^2} = \frac{\sigma^4}{n}$.

(ה) נחשב: $\mathrm{Cov}(T_1,T_2) = \frac{1}{n^2}\sum_{i,j}\mathrm{Cov}(X_i^2,Y_j^2) = \frac{1}{n^2}\sum_i\mathrm{Cov}(X_i^2,Y_i^2) = \frac{1}{n}\mathrm{Cov}(X_1^2,Y_1^2) = \frac{1}{n}(E(X_1^2Y_1^2) - E(X_1^2)E(Y_1^2)) = \frac{2\rho^2\sigma^4}{n}$. לכן השונות של $\frac{1}{2}(T_1+T_2)$ היא $\frac{1}{4}(\frac{2\sigma^4}{n} + \frac{2\sigma^4}{n} + 2\mathrm{Cov}(T_1,T_2)) = \frac{(1+\rho^2)\sigma^4}{n}$. אם כך, T עדיף כאשר $\rho \neq 0$, כי $V(T) < V(\frac{1}{2}(T_1+T_2))$; אם $\rho = 0$, ממילא $T = \frac{1}{2}(T_1+T_2)$.