<수리통계학Ⅱ> 6장_7절 과제

- 과제물은 11월 21일(화)까지 제출

▶ 6.7절 연습문제 #1, #3, #4 &

[A1] 위 연습문제 #1, #3, #4에서 구한 추정량이 최소분산비편향추정량(MVUE; Minimum Variance Unbiased Estimator)에 해당하는지 밝혀라.

** 확률분포가 지수족에 속하는지 확인하고, 완비충분통계량(complete sufficient statistic)을 구한 후, 해당 추정량이 MVUE인지 여부를 확인할 것

[A2] $X_1, X_2, \cdots, X_n$이 균일분포 $U(0, \theta)$에서 추출된 확률표본일 때, θ의 MME(적률추정량)이 $\tilde{\theta} = 2\overline{X}$이고, θ의 MLE(최대가능도추정량)가 $\hat{\theta} = X_{(n)} = \max\{X_1, X_2, \dots, X_n\}$임을 알고 있다.

1) 두 추정량 $\tilde{\theta} = 2\overline{X}$와 $\hat{\theta} = X_{(n)}$의 MSE(mean square error)를 각각 구하라.

2) 위에서 구한 MSE를 비교하여 어떤 추정량이 더 효율적(efficient)인가를 밝혀라.

(※ 다룬 MME $\tilde{\theta} = 2\overline{X}$이 비편향추정량이고, MLE $\hat{\theta} = X_{(n)}$이 편향(biased) 추정량이라는 것을 증명한 지난번 과제의 풀이 과정을 참고할 것)

[A3] $X_1, X_2, \cdots, X_n$이 정규분포 $N(\mu, \sigma^2)$에서 추출된 확률표본일 때, 아래 두 가지 σ^2에 대한 추정량으로 $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$와 $V = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$에 대해서

1) $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$와 $V = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$의 MSE를 각각 구하라

2) 위에서 구한 MSE를 비교해서 어떤 추정량이 더 효율적(efficient)인가 밝혀라.

슬라이드 p38

[A4] $(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ 이변량정규분포(bivariate Normal dist.)에서 추출된 확률표본일 때, 이분포가 이분포가 지수족(exponential family) 속하는 것을 보이고, 지수족 성질을 이용해 ($S_1 = \sum_{i=1}^{n} X_i$, $S_2 = \sum_{i=1}^{n} X_i^2$, $S_3 = \sum_{i=1}^{n} Y_i$, $S_4 = \sum_{i=1}^{n} Y_i^2$, $S_5 = \sum_{i=1}^{n} X_i Y_i$)이 모수 $\underline{\theta} = (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$에 대한 결합 완비충분통계량(joint complete sufficient statistic)임을 보여라

※ 6.4절 과제 중 비편향추정량(unbiased estimator) 관련 문제 풀어볼 것.

수리통계학Ⅱ(002) 과제5(6장7절) 소프트웨어학부 컴퓨터과학전공 2016133 이유진

#6.7-1

(a) $f(x_1\cdots x_n;\sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\theta}\sum x_i^2\right) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\theta}y\right)$

$= g(y;\sigma^2)\cdot 1$

↳ σ^2에 의존하지 않는 함수 $h(\underline{x})$

따라서 σ^2에 대한 충분통계량 $Y=\sum X_i^2$

(b) $\theta=\sigma^2$ 에 대해 $L(\theta;x_1\cdots x_n)=\prod_1^n f(x_i;\theta) = \left(\frac{1}{\sqrt{2\pi\theta}}\right)^n\exp\left(-\frac{1}{2\theta}\sum x_i^2\right)$

$\frac{d}{d\theta}\ln L(\theta) = \frac{d}{d\theta}\left(-\frac{n}{2}\ln(2\pi\theta)-\frac{1}{2\theta}\sum_1^n x_i^2\right)=0$ 이므로 정리하면

$-n\theta+\sum_1^n x_i^2=0$, 즉 $\theta=\frac1n\sum_1^n x_i^2$

따라서 σ^2의 최우추정량(MLE)은 $\hat{\sigma^2}=\frac1n\sum_1^n X_i^2$ ⇒ Y의 함수이다

(c) $E(\hat{\sigma^2}) = E\left(\frac1n\sum_1^n x_i^2\right) = E\left(\frac{\sigma^2}{n}\cdot\sum_1^n\frac{x_i^2}{\sigma^2}\right) = \frac{\sigma^2}{n}\cdot E\left(\sum_1^n \frac{x_i^2}{\sigma^2}\right) = \frac{\sigma^2}{n}\cdot n = \sigma^2$

↑ 이때 $\sum_1^n \frac{x_i^2}{\sigma^2} \sim \chi^2(n)$

$E(\hat{\sigma^2})=\sigma^2$ 이므로 σ^2의 최우추정량은 불편추정량이다.

#6.7-3

$X_1\cdots X_n$ random sample from gamma$(1,\frac1\theta)$ → $f(x)=\begin{cases}\theta\cdot e^{-\theta x}, & 0\le x<\infty,\ \frac1\theta>0\\ 0, & \text{o.w}\end{cases}$ 이므로

$f(\underline{x};\frac1\theta)=\prod_1^n f(x_i;\frac1\theta) = \theta^n\exp[-\theta\sum x_i] = \exp[-\theta\sum x_i + n\ln\theta] = g(y;\frac1\theta)\cdot \underbrace{h(\underline{x})}_{1}$

따라서 $Y=\sum_1^n X_i$는 충분통계량

$Y=X_1+\cdots+X_n$ 에서 $E(e^{tY}) = E(e^{tx_1}\cdot e^{tx_2}\cdot\cdots\cdot e^{tx_n}) = E(e^{tx_1})\cdot\cdots\cdot E(e^{tx_n})$ (∵ 독립)

$= M_{X_1}(t)\cdot\cdots\cdot M_{X_n}(t) = (1-\frac1\theta t)^{-1}\cdot\cdots\cdot(1-\frac1\theta t)^{-1} = \underbrace{(1-\frac1\theta t)^{-n}}_{\text{gamma}(n,\frac1\theta)\text{의 mgf}}$

따라서 $Y\sim$ gamma$(n,\frac1\theta)$이다. → $\mu=\frac n\theta$, $\sigma^2=\frac{n}{\theta^2}$

$E\left(\frac{n-1}{Y}\right) = (n-1)E\left(\frac1Y\right) = (n-1)E\left(\frac{1}{\sum x_i}\right) = (n-1)\int_0^\infty \frac1y f_Y(y)\,dy$

$=(n-1)\int_0^\infty \frac{1}{\Gamma(n)\theta^{-n}} y^{n-1}e^{-y\theta}\,dy = (n-1)\cdot\frac{\theta\,\Gamma(n-1)}{\Gamma(n)} = \frac{n-1}{n-1}\cdot\theta=\theta$ → 비편향추정량이다.

#6.7-4

$X_1\cdots X_n$ random sample from $f(x;p)=p(1-p)^{x-1}$, $x=1,2,\cdots$, $0<p<1$ → $X\sim Geo(p)$

(a) $f(x_1\cdots x_n;p)=\prod_1^n p(1-p)^{x_i-1} = p^n(1-p)^{\sum_1^n x_i - n} = p^n(1-p)^{y-n} = g(y;p)\cdot h(\underline{x})$, 이때 $h(\underline{x})=1$

따라서 $Y=\sum_1^n X_i$는 p에 대한 충분통계량

(b) $\frac{d}{d\theta}L(\theta) = \frac{d}{d\theta}\left(-n\ln\theta + \left(\sum_1^n x_i - n\right)\left(\ln\left(1-\frac1\theta\right)\right)\right)=0$ 이므로 정리하면

$-\frac{n}{\cancel{\theta}} + \left(\sum_1^n x_i - n\right)\cdot\frac{1}{\cancel{\theta}(\theta-1)}=0$, $(\theta-\cancel{1})\cdot n = \sum_1^n x_i - \cancel{n}$, 즉 $\theta=\frac1n\sum_1^n x_i=\bar X$ → MLE

$E(\hat\theta)=E\left(\frac1n\sum_1^n x_i\right) = \frac1n E\left(\sum_1^n x_i\right) = \frac{n\mu}{n}=\mu$ → 비편향

∴ $Y=\bar X$

A1

6.7-1) ① $f(\underline{x};\sigma^2) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \exp\left(-\frac{1}{2\theta}\Sigma x_i^2\right) = \exp\left[-\frac{1}{2\theta}\Sigma x_i^2 - n\ln(\sigma\sqrt{2\pi})\right]$ ← 지수족

따라서 σ^2에 대한 통계량 $Y=\sum_{i}^{n} X_i^2$는 완비충분통계량

② (c)에 의해 $E(\hat{\sigma}^2) = \sigma^2$이므로 비편향추정량

⇒ 이 추정량은 MVUE (최소분산비편향추정량) 이다.

6.7-3) ① $f(\underline{x};\frac{1}{\theta}) = \exp\left[-\theta\Sigma x_i + n\ln\theta\right]$ ← 지수족

따라서 $\frac{1}{\theta}$에 대한 추정량 $Y=\sum_{i}^{n} X_i$는 완비충분통계량

② $E\left(\frac{n-1}{Y}\right) = \theta$ 이므로 비편향추정량 (같은 partition을 가지는 일대일대응)

⇒ 이 추정량은 MVUE이다.

6.7-4) ① $f(\underline{x};p) = p^n(1-p)^{\sum_{i}^{n} x_i - n} = \exp\left[n\ln p + \left(\sum_{i}^{n} x_i - n\right)\ln(1-p)\right]$

$= \exp\left[\ln(1-p)\cdot\sum_{i}^{n} x_i + n\ln p - n\ln(1-p)\right]$ ← 지수족

따라서 p에 대한 통계량 $Y=\sum_{i}^{n} X_i^2$는 완비충분통계량

② (b)에 의해 $E(\bar{X}) = \mu$ 이므로 비편향추정량 (같은 partition을 가지는 일대일대응)

⇒ 이 추정량은 MVUE이다.

A2

$X_1 \cdots X_n$ random sample from $U(0,\theta) \rightarrow f(x;\theta) = \begin{cases} \frac{1}{\theta}, & 0<x<\theta \\ 0, & \text{o.w} \end{cases}$

MLE : $\hat{\theta} = Y = X_{(n)} = \max(X_1, \cdots, X_n) \rightarrow E(\hat{\theta}) = \frac{n}{n+1}\theta$

MME : $\tilde{\theta} = 2\bar{X} \rightarrow E(\tilde{\theta}) = \theta$, 비편향추정량

MME가 MLE보다 편향성은 좋지만 이것이 효율성도 좋다는 것을 의미하는 것은 아님 → MSE 계산필요

(1) $MSE(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right] = Var(\hat{\theta}) + \left(E(\hat{\theta})-\theta\right)^2 = Var(\hat{\theta}) + \left(-\frac{1}{n+1}\theta\right)^2$

이때 $Var(\hat{\theta}) = E(\hat{\theta}^2) - E(\hat{\theta})^2$

$= \int_0^\theta \left(t^2 f_{\hat{\theta}}(t)\right)dt - E(\hat{\theta})^2 = \frac{n}{\theta^n}\cdot\frac{\theta^{n+2}}{n+2} - E(\hat{\theta})^2 = \frac{n}{n+2}\theta^2 - \left(\frac{n}{n+1}\theta\right)^2$

$= \frac{n}{(n+2)(n+1)^2}\theta^2$

따라서 $MSE(\hat{\theta}) = \left(\frac{n}{(n+2)(n+1)^2} + \frac{1}{(n+1)^2}\right)\theta^2 = \frac{2(n+1)}{(n+2)(n+1)^2}\theta^2 = \frac{2}{(n+1)(n+2)}\theta^2$

$MSE(\tilde{\theta}) = E\left[(\tilde{\theta}-\theta)^2\right] = Var(\tilde{\theta}) + \left(E(\tilde{\theta})-\theta\right)^2 = Var(\tilde{\theta})$

이때 $Var(\tilde{\theta}) = Var(2\bar{X}) = 4\,Var\left(\frac{1}{n}\sum_{i}^{n} X_i\right) = \frac{4}{n}Var(X_i) = \frac{4}{n}\cdot\frac{1}{12}\theta^2 = \frac{1}{3n}\theta^2$ $\left(\because E(X_i) = \frac{1}{2}\theta,\right.$

따라서 $MSE(\tilde{\theta}) = \frac{1}{3n}\theta^2$ $\left. Var(X_i) = \frac{1}{12}\theta^2\right)$

(2) $MSE(\hat{\theta}) < MSE(\tilde{\theta})$ 이므로 MLE가 MME보다 효율적인 추정량이다. (단, $n>2$)

2/3

A3

$X_1 \cdots X_n$ random sample from $N(\mu, \sigma^2) \rightarrow S^2 = \frac{1}{n-1}\sum_1^n (X_i - \bar{X})^2$, $V = \frac{1}{n}\sum_1^n (X_i - \bar{X})^2$

(1) $E\left(\sum_1^n (X_i - \bar{X})^2\right) = E\left(\sum_1^n X_i^2 - 2\bar{X}X_i + \bar{X}^2\right)$

$= \sum_1^n E(X_i^2) - nE(\bar{X}^2)$

$= n(\sigma^2 + \mu^2) - n(\mu^2 + \frac{\sigma^2}{n})$

$= (n-1)\sigma^2$

따라서 $\begin{cases} E(S^2) = E\left(\frac{1}{n-1}\sum_1^n (X_i - \bar{X})^2\right) = \sigma^2 \text{ (비편향추정량)} \\ E(V) = E\left(\frac{1}{n}\sum_1^n (X_i - \bar{X})^2\right) = \frac{n-1}{n}\sigma^2 \text{ (편향추정량)} \end{cases}$ 이다.

$MSE(S^2) = Var(S^2) + Bias(S^2) = \frac{2}{n-1}\sigma^4 + 0$

($\because$ $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$ 이므로 $Var\left(\frac{(n-1)S^2}{\sigma^2}\right) = 2(n-1)$,
정리하면 $Var(S^2) = \frac{2}{n-1}\sigma^4$)

$MSE(V) = Var(V) + Bias(V)$

$= Var\left(\frac{n-1}{n}S^2\right) + \left(\frac{n-1}{n}\sigma^2 - \sigma^2\right)^2 = \left(\frac{n-1}{n}\right)^2 \cdot Var(S^2) + \left(-\frac{1}{n^2}\sigma^2\right)^2 = \frac{2n-1}{n^2}\sigma^4$

(2) $MSE(S^2) > MSE(V)$ 이므로 V가 S^2보다 더 효율적인 추정량이다.

A4

$(X_1, Y_1), (X_2, Y_2), \cdots, (X_n, Y_n)$ random sample from $BVN(\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$

즉, $\underline{\theta} = (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$ 일때

$$f(\underline{x};\underline{\theta}) = \left(\frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\right)^n \cdot \exp\left[\frac{1}{2(1-\rho^2)} \cdot \left\{\frac{\Sigma(x_i-\mu_x)^2}{\sigma_x^2} - 2\rho\left(\frac{\Sigma(x_i-\mu_x)}{\sigma_x}\right)\left(\frac{\Sigma(y_i-\mu_y)}{\sigma_y}\right) + \left(\frac{\Sigma(y_i-\mu_y)^2}{\sigma_y^2}\right)\right\}\right]$$

$$= \exp\left[\frac{1}{2(1-\rho^2)} \cdot \left\{\frac{\Sigma(x_i-\mu_x)^2}{\sigma_x^2} - 2\rho\left(\frac{\Sigma(x_i-\mu_x)}{\sigma_x}\right)\left(\frac{\Sigma(y_i-\mu_y)}{\sigma_y}\right) + \left(\frac{\Sigma(y_i-\mu_y)^2}{\sigma_y^2}\right)\right\} - n\ln(2\pi\sigma_x\sigma_y\sqrt{1-\rho^2})\right]$$

→ 지수형태로 나타낼수 있으므로 지수족이다.

위 식을 정리하면

$$\exp\left[\frac{1}{\sigma_x^2}\Sigma x_i^2 + \frac{1}{\sigma_y^2}\Sigma y_i^2 - \frac{2\rho}{\sigma_x\sigma_y}\Sigma x_i\Sigma y_i - 2\frac{\mu_x\sigma_y - n\rho\mu_y\sigma_x}{\sigma_x^2\sigma_y}\Sigma x_i + 2\frac{n\rho\mu_x\sigma_y - 2\mu_y\sigma_x}{\sigma_x\sigma_y^2}\Sigma y_i + \frac{n\mu_x^2}{\sigma_x^2} + \frac{n\mu_y^2}{\sigma_y^2} - \frac{2\rho n^2\mu_x\mu_y}{\sigma_x\sigma_y} - n\ln(2\pi\sigma_x\sigma_y\sqrt{1-\rho^2})\right]$$

→ $\sum_1^n X_i^2$, $\sum_1^n Y_i^2$, $\sum_1^n X_i$, $\sum_1^n Y_i$, $\sum_1^n X_iY_i$는 $\underline{\theta} = (\mu_x, \mu_y, \sigma_x^2, \sigma_y^2, \rho)$의 완비충분통계량이다.