MATHEMATICAL STATISTICS - 1 | CHAPTER 7

UNBIASED MINIMUM VARIANCE ESTIMATOR (UMVE)

$Y = v(X_1, X_2, \dots, X_n)$ will be called an UMVE of θ if

(i) Y is unbiased: $E(Y) = \theta$

(ii) $Var(Y) \leq$ Every other unbiased estimator of θ

7.1. Show that the mean $\bar{X}$ of a random sample of size n from a distribution having p.d.f. $f(x;\theta) = (1/\theta)e^{-(x/\theta)}$, $0 < x < \infty$, $0 < \theta < \infty$, zero elsewhere, is an unbiased estimator of θ and has variance θ^2/n.

7.2. Let $X_1, X_2, \dots, X_n$ denote a random sample from a normal distribution with mean zero and variance θ, $0 < \theta < \infty$. Show that $\sum_1^n X_i^2/n$ is an unbiased estimator of θ and has variance $2\theta^2/n$.

7.3. Let $Y_1 < Y_2 < Y_3$ be the order statistics of a random sample of size 3 from the uniform distribution having p.d.f. $f(x;\theta) = 1/\theta$, $0 < x < \theta$, $0 < \theta < \infty$, zero elsewhere. Show that $4Y_1$, $2Y_2$, and $\frac{4}{3}Y_3$ are all unbiased estimators of θ. Find the variance of each of these unbiased estimators.

7.4. Let Y_1 and Y_2 be two independent unbiased estimators of θ. Say the variance of Y_1 is twice the variance of Y_2. Find the constants k_1 and k_2 so that $k_1Y_1 + k_2Y_2$ is an unbiased estimator with smallest possible variance for such a linear combination.

7.1) $X_i \overset{i.i.d}{\sim}$ Exponential(θ) ; $E(X_i) = \theta$; $Var(X_i) = \theta^2$

$$E(\bar{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n}\cdot E(\sum X_i) = \frac{1}{n}\cdot \sum E(X_i) = \frac{1}{n}\cdot\sum\theta = \frac{1}{n}\cdot n\theta = \theta$$

$$Var(\bar{X}) = Var\left(\frac{\sum X_i}{n}\right) = \frac{1}{n^2}\cdot Var(\sum X_i) = \frac{1}{n^2}\cdot\sum(Var(X_i)) = \frac{1}{n^2}\cdot\sum\theta^2$$

$$= \frac{1}{n^2}\cdot n\cdot\theta^2 = \frac{\theta^2}{n}$$

7.2) Remember; $Z_i \overset{i.i.d}{\sim} Normal(0;1) \Rightarrow \sum Z_i^2 \sim \chi_n^2$

Then; $X_i \overset{i.i.d}{\sim} Normal(0;\sigma^2) \Rightarrow W = \sum \frac{X_i^2}{\sigma^2} \sim \chi_n^2$

$$E(W) = n \; ; \; Var(W) = 2n$$

We have, $X_i \sim Normal(0; \overset{=\sigma^2}{\theta})$

$$E\left(\frac{\sum X_i^2}{n}\right) = E\left(\frac{\theta}{n}\cdot\frac{\sum X_i^2}{\theta}\right) = \frac{\theta}{n}\cdot E\left(\frac{\sum X_i^2}{\theta}\right) = \frac{\theta}{n}\cdot n = \theta$$

$$Var\left(\frac{\sum X_i^2}{n}\right) = Var\left(\frac{\theta}{n}\cdot\frac{\sum X_i^2}{\theta}\right) = \frac{\theta^2}{n^2}\cdot Var\left(\frac{\sum X_i^2}{\theta}\right) = \frac{\theta^2}{n^2}\cdot 2n = \frac{2\theta^2}{n}$$

7.3) Remember: $Y_1 < Y_2 < \dots < Y_k < \dots < Y_n$ order statistics with $f(x)$, pdf of Y_j, $j = 1,2,\dots,n$ are;

$$Y_1 : g_1(y) = n\cdot[1-F(y)]^{n-1}\cdot f(y)$$

$$Y_k : g_k(y) = \frac{n!}{(k-1)!\,.\,(n-k)!}\cdot[F(y)]^{k-1}\cdot[1-F(y)]^{n-k}\cdot f(y)$$

$$Y_n : g_n(y) = n.\,[F(y)]^{n-1}\cdot f(y)$$

We have; $X_i \sim Uniform(0;\theta)$ $\quad f(x) = \frac{1}{\theta}$ and $F(x) = \frac{x}{\theta}$

$Y_1 < Y_2 < Y_3$ and $\hat{\theta}_1 = 4Y_1$; $\hat{\theta}_2 = 2Y_2$; $\hat{\theta}_3 = \frac{4}{3}Y_3$ are estimators of θ.

$$Y_1 : \quad g_1(y) = 3\cdot\left[1-\frac{y}{\theta}\right]^2\cdot\frac{1}{\theta}$$

$$E(Y_1) = \int_0^\theta y\cdot 3\left[1-\frac{y}{\theta}\right]^2\cdot\frac{1}{\theta}\,dy = \frac{3}{\theta}\int_0^\theta\left(y - \frac{2y^2}{\theta} + \frac{y^3}{\theta^2}\right)dy$$

$$= \frac{3}{\theta}\left(\frac{y^2}{2} - \frac{2y^3}{3\theta} + \frac{y^4}{4\theta^2}\right]_0^\theta = \frac{3}{\theta}\cdot\left(\frac{\theta^2}{12} - \frac{2\theta^2}{3} + \frac{\theta^2}{4}\right) = \frac{3}{\theta}\cdot\frac{\theta^2}{12} = \frac{\theta}{4}$$

$$E(Y_1^2) = \int_0^\theta y^2 \cdot 3\left(1-\frac{y}{\theta}\right)^2 \cdot \frac{1}{\theta}\, dy = \frac{3}{\theta}\int_0^\theta \left(y^2 - \frac{2y^3}{\theta} + \frac{y^4}{\theta^2}\right) dy$$

$$= \frac{3}{\theta}\left(\frac{\theta^3}{3} - \frac{2\theta^4}{4\theta} + \frac{\theta^5}{5\theta^2}\right) = \frac{3}{\theta}\cdot\frac{2\theta^3}{60} = \frac{\theta^2}{10}$$

$$Var(Y_1) = E(Y_1^2) - E^2(Y_1) = \frac{\theta^2}{10} - \frac{\theta^2}{16} = \frac{3\theta^2}{80}$$

Then; $E(\hat{\theta}_1) = E(4Y_1) = 4E(Y_1) = 4\cdot\frac{\theta}{4} = \theta$ ✓unbiased

$$Var(\hat{\theta}_1) = Var(4Y_1) = 16\,Var(Y_1) = 16\cdot\frac{3\theta^2}{80} = \frac{3\theta^2}{5}$$

Calculations for $\hat{\theta}_2$ and $\hat{\theta}_3$ are similar.

7.4) $E(Y_1) = \theta$ $\quad$ $E(Y_2) = \theta$

$Var(Y_1) = 2\sigma^2$ $\quad$ $Var(Y_2) = \sigma^2$

Let; $\hat{\theta} = k_1 Y_1 + k_2 Y_2$

$$E(\hat{\theta}) = E(k_1 Y_1 + k_2 Y_2) = k_1 E(Y_1) + k_2 E(Y_2) = (k_1+k_2)\cdot\theta = \theta$$

$$\boxed{k_1 + k_2 = 1}$$

$$Var(\hat{\theta}) = Var(k_1 Y_1 + k_2 Y_2) = k_1^2 Var(Y_1) + k_2^2 Var(Y_2)$$

$$= 2k_1^2\sigma^2 + k_2^2\sigma^2 = 2k_1^2\sigma^2 + (1-k_1)^2\sigma^2$$

$$= (2k_1^2 + 1 - 2k_1 + k_1^2)\sigma^2 = (3k_1^2 - 2k_1 + 1)\sigma^2$$

Let $w(k) = 3k_1^2 - 2k_1 + 1$

$$\frac{d\,w(k)}{dk} = 6k_1 - 2 = 0$$

$$k_1 = \frac{1}{3} \Rightarrow k_2 = \frac{2}{3}$$

SUFFICIENT STATISTICS (SS)

Let $X_i \overset{i.i.d}{\sim} \{f(x;\theta),\ \theta \in \Omega\}$ and

$$y_i = u_1(X_1, X_2, \dots, X_n) \sim g_1(y;\theta)$$

y_1 is a sufficient statistics if and only if

$$h(\underline{x}|y) = \frac{f(x_1;\theta).f(x_2;\theta).\dots.f(x_n;\theta)}{g_1[u_1(x_1,x_2,\dots,x_n);\theta]} = \frac{L(\underline{x};\theta)}{g_1(y;\theta)} = \underbrace{H(x_1,x_2,\dots,x_n)}_{\text{independent of } \theta}$$

Example: let $X_i \overset{i.i.d}{\sim} Bernoulli(\theta)$

$$f(x_i;\theta) = \theta^x.(1-\theta)^{1-x} \quad x=0,1 \quad 0<\theta<1$$

let $y = \sum X_i$

Then: $Y \sim Binomial(n;\theta)$

$$g_1(y) = \binom{n}{y}.\theta^y.(1-\theta)^{n-y}$$

$$h(x_1,x_2,\dots,x_n|y) = \frac{\theta^{x_1}.(1-\theta)^{1-x_1}.\theta^{x_2}.(1-\theta)^{1-x_2}\dots\theta^{x_n}(1-\theta)^{1-x_n}}{\binom{n}{y}.\theta^y.(1-\theta)^{n-y}}$$

$$= \frac{\cancel{\theta^{\sum x_i}.(1-\theta)^{n-\sum x_i}}}{\binom{n}{y}.\cancel{\theta^y.(1-\theta)^{n-y}}} = \frac{1}{\binom{n}{y}} \text{ is independent of } \theta$$

Then, $Y = \sum X_i$ is a sufficient statistics.

Theorem: If Y is a SS for θ, then a 1-1 function of Y, let's say $k_1(Y)$ is also a S.S. for θ.

Example: Since $y=\sum X_i$ is a SS for θ of Bernoulli trials, $\bar{y}=\frac{\sum X_i}{n}$ is also a SS for θ.

Example: let $X_i \overset{i.i.d}{\sim} \text{Gamma}(\alpha=2;\ \beta=\theta)$

Remember, mgf of Gamma is $m(t)=(1-\beta t)^{-\alpha}$.

Show that $y=\sum X_i$ is a SS for θ.

Answer: mgf of X_i is $m(t)=(1-\theta t)^{-2}$ and

$$M_y(t)=E\left[e^{yt}\right]=E\left[e^{t(x_1+x_2+\cdots+x_n)}\right]=E(e^{tX_1}).E(e^{tX_2})\ldots E(e^{tX_n})$$

$$=\left[(1-\theta t)^{-2}\right]^n=(1-\theta t)^{-2n}$$

Then; $y\sim \text{Gamma}(\alpha=2n;\ \beta=\theta)$

Remember; $X\sim \text{Gamma}(\alpha;\beta)\Rightarrow \boxed{f(x)=\frac{1}{\Gamma(\alpha)\beta^{\alpha}}.x^{\alpha-1}.e^{-x/\beta}}$

Then; $f(x_i;\theta)=\frac{x_i.e^{-x_i/\beta}}{\Gamma(2).\theta^2}$ and

$$h(x_1,x_2,\ldots,x_n\,|\,y)=\frac{\frac{x_1.e^{-x_1/\theta}}{\Gamma(2).\theta^2}.\frac{x_2.e^{-x_2/\theta}}{\Gamma(2).\theta^2}.\cdots.\frac{x_n.e^{-x_n/\theta}}{\Gamma(2).\theta^2}}{\frac{y^{2n-1}.e^{-y/\theta}}{\Gamma(2n).\theta^{2n}}}$$

$$=\frac{\Gamma(2n)}{[\Gamma(2)]^n}.\frac{x_1.x_2.\ldots.x_n}{y^{2n-1}}$$ is independent of θ.

Then, $y=\sum X_i$ is a S.S. for θ.

Example let $X_1, X_2, \dots, X_n$ is a random sample from $f(x) = e^{-(x-\theta)}, x > \theta$. Show that $Y_{(1)}$: first order statistics is a SS for θ.

Answer Define $I_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{o.w.} \end{cases}$

Then, $f(x_i) = e^{-(x_i-\theta)} \cdot I_{(\theta,\infty)}(x_i) = e^{-(x_i-\theta)} \cdot I_{(\theta,\infty)}(\min(x_i))$

$$g_1(y) = n \cdot [1 - F(y)]^{n-1} \cdot f(y)$$

$$F(x) = \int_\theta^x e^{-(w-\theta)} dx = 1 - e^{-(x-\theta)} \qquad x \geq \theta$$

$$g_1(y) = n \cdot \left[e^{-(y-\theta)}\right]^{(n-1)} \cdot e^{-(y-\theta)} \cdot I_{(\theta;\infty)}(\min x_i)$$

So; $$h(\underline{x} \mid y) = \frac{L(\underline{x}, \theta)}{g_1(y)} = \frac{\prod e^{-(x_i-\theta)} \cdot I_{(\theta,\infty)}(\min x_i)}{n \cdot e^{-n \cdot (y-\theta)} \cdot I_{(\theta,\infty)}(\min x_i)}$$

$$= \frac{e^{-x_1-x_2-\dots-x_n}}{n \cdot e^{-ny}}$$ is independent of θ.

then, $Y_1 = \min(x_i)$ is a SS for θ.

Neyman Factorization Theorem

let $X_i \overset{iid}{\sim} \{f(x;\theta), \theta \in \Omega\}$. $Y = u(X_1, X_2, \dots, X_n)$ is a SS if and only if we can find two functions k_1 and k_2 such that

$$L(x_1, x_2, \dots, x_n; \theta) = k_1(y;\theta) \cdot k_2(x_1, x_2, \dots, x_n)$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Example: let $X_i \overset{i.i.d}{\sim} \text{Normal}(\theta, \sigma^2)$ show that $\bar{X}$ is a SS for θ.

Answer: $\sum (x_i-\theta)^2 = \sum [(x_i-\bar{x})+(\bar{x}-\theta)]^2$

$$= \sum (x_i-\bar{x})^2 + 2\underbrace{\sum (x_i-\bar{x})(\bar{x}-\theta)}_{=(\bar{x}-\theta).\sum(x_i-\bar{x})=0} + \sum (\bar{x}-\theta)^2$$

$$= \sum (x_i-\bar{x})^2 + n.(\bar{x}-\theta)^2$$

Remember; $f(x_i) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left\{-\frac{1}{2}\left(\frac{x_i-\theta}{\sigma}\right)^2\right\}$

Then;

$$L(x_1,x_2,\ldots,x_n;\theta) = \left(\frac{1}{\sigma\sqrt{2\pi}}\right)^n \cdot \exp\left\{-\sum_{i=1}^{n}\frac{(x_i-\theta)^2}{2\sigma^2}\right\}$$

$$= \underbrace{\exp\left\{\frac{-n.(\bar{x}-\theta)^2}{2\sigma^2}\right\}}_{k_1(\bar{x};\theta)} \cdot \underbrace{\frac{1}{(\sigma\sqrt{2\pi})^n} \cdot \exp\left\{-\sum\frac{(x_i-\bar{x})^2}{2\sigma^2}\right\}}_{k_2(x_1,x_2,\ldots,x_n)}$$

Then, $\bar{X}$ is a SS for θ by factorization thm.

Example: let $X_i \overset{i.i.d}{\sim}$ ~~[illegible]~~ $\text{Beta}(\alpha=\theta;\beta=1)$. Find a SS for θ.

$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot x^{\alpha-1} \cdot (1-x)^{\beta-1}$

Answer: $f(x;\theta) = \theta . x^{\theta-1} \quad 0 < x < 1$

$L(x_1,x_2,\ldots,x_n;\theta) = \theta^n . (x_1 . x_2 . \ldots . x_n)^{\theta-1}$ let $y = x_1 . x_2 . \ldots . x_n$

$$= \underbrace{\theta^n . y^{\theta-1} = \theta^n . y^\theta}_{k_1(y;\theta)} . \underbrace{\frac{1}{x_1 . x_2 . \ldots . x_n}}_{k_2(x_1,x_2,\ldots,x_n)}$$

y is a SS for θ

7.10. Let $X_1, X_2, \ldots, X_n$ be a random sample from the normal distribution $N(0, \theta)$, $0 < \theta < \infty$. Show that $\sum_1^n X_i^2$ is a sufficient statistic for θ.

7.11. Prove that the sum of the observations of a random sample of size n from a Poisson distribution having parameter θ, $0 < \theta < \infty$, is a sufficient statistic for θ.

7.12. Show that the nth order statistic of a random sample of size n from the uniform distribution having p.d.f. $f(x; \theta) = 1/\theta$, $0 < x < \theta$, $0 < \theta < \infty$, zero elsewhere, is a sufficient statistic for θ.

7.10) $X_i \overset{i.i.d}{\sim} Normal(0;\theta)$

$$f(x_i;\theta) = \frac{1}{\sqrt{2\pi\theta}} \cdot \exp\left\{-\frac{x_i^2}{2\theta}\right\} \quad ; \; \underline{y = \sum x_i^2}$$

$$L(x_1, x_2, \ldots, x_n;\theta) = \underbrace{(2\pi\theta)^{-n/2} \cdot \exp\left\{-\frac{y}{2\theta}\right\}}_{k_1(y;\theta)} \cdot \underbrace{1}_{k_2(x_1, x_2, \ldots, x_n)}$$

Then; $y = \sum x_i^2$ is a S.S. for θ.

7.11) $X_i \overset{i.i.d}{\sim} Poisson(\theta)$

$$f(x_i;\theta) = \frac{e^{-\theta} \cdot \theta^{x_i}}{x_i!} \quad ; \; y = \sum x_i$$

$$L(x_1, x_2, \ldots, x_n;\theta) = \frac{e^{-n\theta} \cdot \theta^{y}}{x_1! . x_2! . \ldots . x_n!} = \underbrace{e^{-n\theta} \cdot \theta^{y}}_{k_1(y;\theta)} \cdot \underbrace{\frac{1}{x_1! . x_2! . \ldots . x_n!}}_{k_2(x_1, x_2, \ldots, x_n)}$$

Then, $y = \sum x_i$ is a S.S. for θ.

7.12) $X_i \overset{iid}{\sim} \text{Uniform}(0;\theta)$

$f(x_i;\theta) = \frac{1}{\theta} \cdot I_{(0;\theta)}(x_i)$ $\qquad y = \max(X_i)$

$$L(x_1, x_2, \ldots, x_n;\theta) = \left(\frac{1}{\theta}\right)^n \cdot \prod_{i=1}^{n} I_{(0;\theta)}(x_i) = \underbrace{\left(\frac{1}{\theta}\right)^n \cdot I_{(0;\theta)}(y)}_{k_1(y;\theta)} \cdot \underbrace{1}_{k_2(x_1, x_2, \ldots, x_n)}$$

Then, $y = \text{Max}(X_i)$ is a SS for θ

Theorem: Let $X_i \overset{iid}{\sim} \{ f(x;\theta), \theta \in \Omega \}$. If a ~~SS~~ for θ exist and if MLE of θ also exists uniquely, $\hat{\theta}_{MLE}$ is a function of SS y.

Example4 For 7.11, we have

$L(\theta) = \frac{e^{-n\theta} \cdot \theta^y}{\prod x_i!}$ where $y = \sum X_i$ is a SS.

$\ln L(\theta) = -n\theta + y \ln\theta - \sum \ln x_i$

$\frac{d}{d\theta} \ln L(\theta) = -n + \frac{y}{\theta} = 0 \Rightarrow \hat{\theta}_{MLE} = \frac{y}{n} = \bar{X}$

For 7.12, we have

$L(\theta) = \left(\frac{1}{\theta}\right)^n \quad 0 < x_i < \theta$

$L(\theta)$ is Maximized when θ is minimum. But since ~~always~~ $x_i < \theta$, the minimum value θ can take is $\max(X_i)$

$\hat{\theta}_{MLE} = y = \text{Max}(X_i)$

RAO and BLACKWELL THEOREM

Let $X_i \overset{iid}{\sim} \{ f(x;\theta), \theta \in \Omega \}$ and let $Y_1 = u_1(X_1, X_2, \dots, X_n)$ is a SS for θ. Let $Y_2 = u_2(X_1, X_2, \dots, X_n)$ is an unbiased estimator of θ: $E(Y_2) = \theta$ which is NOT a function of Y_1 alone. Then; $\varphi(Y_1) = E(Y_2 | Y_1)$ is a statistics such that (i) is a function of SS.

(ii) is unbiased: $E[\varphi(Y_1)] = \theta$

(iii) $Var[\varphi(Y_1)] \le Var(Y_2)$

Example 4 Let $X_i \sim \{ f(x;\theta), \theta \in \Omega \}$ and let $E(X) = 2\theta$ and $Var(X) = 2\theta^2$. Let $Y = \sum X_i$ is a SS and $Y_1 = \frac{\sum X_i}{2n}$ and $Y_2 = \frac{X_1 + X_n}{4}$ are two statistics.

a) Show that Y_1 and Y_2 are unbiased

b) Show that $Var(Y_1) \le Var(Y_2)$

c) Show that $Var(E(Y_2|Y_1)) \le Var(Y_2)$

(Hint: Use Conditional Variance Formula:

$$Var(X) = E[Var(X|Y)] + Var[E(X|Y)]$$)

Answer 4 a) $E(Y) = E(\sum X_i) = \sum E(X_i) = 2n \cdot \theta$

$$E(Y_1) = E\left(\frac{Y}{2n}\right) = \frac{1}{2n} E(Y) = \frac{1}{2n} \cdot 2n\theta = \theta$$

$$E(Y_2) = E\left(\frac{X_1 + X_n}{4}\right) = \frac{1}{4}[E(X_1) + E(X_n)] = \frac{1}{2}(2\theta + 2\theta) = \theta$$

b) $Var(Y) = \sum Var(X_i) = 2n\theta^2$

$Var(Y_1) = Var\left(\frac{Y}{2n}\right) = \frac{1}{4n^2} Var(Y) = \frac{1}{4n^2} \cdot 2n\theta^2 = \frac{\theta^2}{2n}$

$Var(Y_2) = Var\left(\frac{X_1+X_2}{4}\right) = \frac{1}{16}(2\theta^2 + 2\theta^2) = \frac{\theta^2}{4}$

$$\frac{\theta^2}{2n} \leq \frac{\theta^2}{4} \text{ for } n \geq 2$$

c) $\varphi(Y_1) = E(Y_2 / Y_1)$; $Var(Y_2) = E[Var(Y_2|Y_1)] + Var(E(Y_2|Y_1))$

$Var(\varphi(Y_1)) = Var(Y_2) - E(Var(Y_2|Y_1)) \leq Var(Y_2)$

7.20. If X_1, X_2 is a random sample of size 2 from a distribution having p.d.f. $f(x;\theta) = (1/\theta)e^{-x/\theta}$, $0 < x < \infty$, $0 < \theta < \infty$, zero elsewhere, find the joint p.d.f. of the sufficient statistic $Y_1 = X_1 + X_2$ for θ and $Y_2 = X_2$. Show that Y_2 is an unbiased estimator of θ with variance θ^2. Find $E(Y_2|y_1) = \varphi(y_1)$ and the variance of $\varphi(Y_1)$.

7.21. Let the random variables X and Y have the joint p.d.f. $f(x, y) = (2/\theta^2)e^{-(x+y)/\theta}$, $0 < x < y < \infty$, zero elsewhere.

(a) Show that the mean and the variance of Y are, respectively, $3\theta/2$ and $5\theta^2/4$.

(b) Show that $E(Y|x) = x + \theta$. In accordance with the theory, the expected value of $X + \theta$ is that of Y, namely, $3\theta/2$, and the variance of $X + \theta$ is less than that of Y. Show that the variance of $X + \theta$ is in fact $\theta^2/4$.

7.20) Let $g(y_1, y_2)$ is joint pdf of Y_1, Y_2 and $g_1(y)$ and $g_2(y)$ is pdf of Y_1 and Y_2 respectively.

Since $Y_1 = X_1 + X_2$ and $Y_2 = X_2$,

$g_2(y) = f(y) = \frac{1}{\theta} \cdot e^{-x/\theta}$ and $X \sim$ Exponential(θ)

$E(X) = \theta$

$Y_2 \sim$ Exponential (θ)

$Var(X) = \theta^2$

$E(Y_2) = \theta$ and $Var(Y_2) = \theta^2$

$$f_{12}(x_1,x_2) = f(x_1)\cdot f(x_2) = \frac{1}{\theta^2}\cdot e^{-(x_1+x_2)/\theta}$$

$$\left.\begin{aligned} y_1 &= x_1 + x_2 \\ y_2 &= x_2 \end{aligned}\right\} \quad \begin{aligned} x_1 &= y_1 - y_2 \\ x_2 &= y_2 \end{aligned} \qquad |J| = \begin{array}{c|cc|} & y_1 & y_2 \\ x_1 & 1 & -1 \\ x_2 & 0 & 1 \end{array} = 1$$

$x_1 + x_2 = y_1 - y_2 + y_2 = y_1$

$$g(y_1,y_2) = f_{12}(y_1 - y_2, y_2)\cdot |J| = \frac{1}{\theta^2}\cdot e^{-y_1/\theta} \qquad 0<y_2<y_1$$

$$g_1(y_1) = \int_0^{y_1} g(y_1,y_2)\,dy_2 = \int_0^{y_1} \frac{1}{\theta^2}\cdot e^{-y_1/\theta}\,dy_2 = \frac{1}{\theta^2} e^{-y_1/\theta}\int_0^{y_1} dy_2$$

$$= \frac{1}{\theta^2} e^{-y_1/\theta}\; y_2\Big]_0^{y_1} = \frac{y_1}{\theta^2}\cdot e^{-y_1/\theta}, \quad 0<y_1$$

$$h(y_2|y_1) = \frac{g(y_1,y_2)}{g_1(y_1)} = \frac{\frac{1}{\theta^2}\cdot e^{-y_1/\theta}}{\frac{y_1}{\theta^2}\cdot e^{-y_1/\theta}} = \frac{1}{y_1} \qquad 0<y_1$$

$$\varphi(y_1) = E(y_2|y_1) = \int_0^{y_1} y_2\cdot\frac{1}{y_1}\cdot dy_2 = \frac{1}{y_1}\int_0^{y_1} y_2\,dy_2 = \frac{1}{y_1}\cdot\frac{y_2^2}{2}\Big]_0^{y_1} = \frac{y_1}{2}$$

$$E(\varphi(y_1)) = E\left(\frac{y_1}{2}\right) = \frac{1}{2}\cdot E(y_1) = \frac{1}{2}E(X_1+X_2) = \frac{1}{2}\cdot 2\theta = \theta$$

$\varphi(y_1)$ is unbiased!

$$Var[\varphi(y_1)] = Var\left[\frac{y_1}{2}\right] = \frac{1}{4}Var(y_1) = \frac{1}{4}Var(X_1+X_2)$$

$$= \frac{1}{4}(\theta^2+\theta^2) = \frac{\theta^2}{2} < \theta^2 = Var(y_2)$$

7.21) $f(x,y) = \frac{2}{\theta^2} \cdot e^{-(x+y)/\theta} \qquad 0 < x < y < \infty$

a) $g_1(x) = \int_x^\infty f(x,y)\,dx = \int_x^\infty \frac{2}{\theta^2} \cdot e^{-(x+y)/\theta}\,dy = \frac{2}{\theta^2} \cdot (-\theta) \cdot e^{-(x+y)/\theta} \Big]_x^\infty$

$= \frac{2}{\theta} \cdot e^{-2x/\theta} \qquad 0 < x < \infty$

Then, $X \sim \text{Exponential}\left(\frac{\theta}{2}\right)$; $E(X) = \frac{\theta}{2}$; $Var(X) = \frac{\theta^2}{4}$

$g_2(y) = \int_0^y f(x,y)\,dx = \int_0^y \frac{2}{\theta^2} \cdot e^{-(x+y)/\theta}\,dx = \frac{2}{\theta^2} \cdot (-\theta) \cdot e^{-(x+y)/\theta} \Big]_0^y$

$= \frac{2}{\theta} \cdot e^{-y/\theta} - \frac{2}{\theta} \cdot e^{-2y/\theta}$

$E(y) = \int_0^\infty y \cdot g_2(y)\,dy = 2\int_0^\infty y \cdot \underbrace{\frac{1}{\theta} \cdot e^{-y/\theta}}_{\substack{Exp(\theta) \\ \mu=\theta \\ \sigma^2=\theta^2}}\,dy - \int_0^\infty y \cdot \underbrace{\frac{2}{\theta} \cdot e^{-2y/\theta}}_{\substack{Exp\left(\frac{\theta}{2}\right) \\ \mu=\frac{\theta}{2} \quad \sigma^2=\frac{\theta^2}{4}}}\,dy$

$= 2\theta - \frac{\theta}{2} = \frac{3\theta}{2}$

$E(y^2) = \int_0^\infty y^2 \cdot g_2(y)\,dy = 2\int_0^\infty y^2 \cdot \frac{1}{\theta} \cdot e^{-y/\theta}\,dy - \int_0^\infty y^2 \cdot \frac{2}{\theta} \cdot e^{-2y/\theta}\,dy$

$= 2\underbrace{[\theta^2 + \theta^2]}_{\sigma^2+\mu^2} - \underbrace{\left[\frac{\theta^2}{4} + \frac{\theta^2}{4}\right]}_{\sigma^2+\mu^2} = 4\theta^2 - \frac{\theta^2}{2} = \frac{7\theta^2}{2}$

$Var(y) = E(y^2) - E^2(y) = \frac{7\theta^2}{2} - \left(\frac{3\theta}{2}\right)^2 = \frac{5\theta^2}{4}$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

b) $$h(y|x) = \frac{f(x,y)}{g_1(x)} = \frac{\frac{2}{\theta^2}\cdot e^{-(x+y)/\theta}}{\frac{2}{\theta}\cdot e^{-2x/\theta}} = \frac{e^{-y/\theta}}{\theta\cdot e^{-x/\theta}} \quad y>x$$

$$E(y|x) = \int_x^\infty y\cdot h(y|x)\,dy = \int_x^\infty \frac{y}{\theta\cdot e^{-x/\theta}}\cdot e^{-y/\theta}\,dy$$

LAPTÜ

$$= \frac{1}{\theta\cdot e^{-x/\theta}}\cdot\int_x^\infty y\cdot e^{-y/\theta}\,dy \qquad u=y \quad dv=e^{-y/\theta}dy \qquad du=dy \quad v=-\theta\cdot e^{-y/\theta}$$

$$\int u\,dv = uv-\int v\,du$$

$$= \frac{1}{\theta\cdot e^{-x/\theta}}\left[\left(-y\theta\cdot e^{-y/\theta}\right]_x^\infty + \int_x^\infty \theta\cdot e^{-y/\theta}dy\right]_x^\infty$$

$$= \frac{1}{\theta\cdot e^{-x/\theta}}\cdot\left[x\theta e^{-x/\theta}-0\right] + \frac{1}{\theta\cdot e^{-x/\theta}}\cdot\left[-\theta^2\cdot e^{-y/\theta}\right]_x^\infty$$

$$= x+\theta$$

$$E[E(y|X)] = E(X+\theta) = E(X)+\theta = \frac{\theta}{2}+\theta = \frac{3\theta}{2} = E(y)$$

$$Var[E(y|X)] = Var(X+\theta) = Var(X) = \frac{\theta^2}{4} < \frac{5\theta^2}{4} = Var(y)$$

COMPLETENESS & UNIQUENESS

let $Z \sim \{h(z;\theta);\theta \in \Omega\}$. If $E[u(Z)]=0$ requires $u(z)=0$ on the set Z has NON-zero probabilities, then the family $\{h(z;\theta):\theta\in\Omega\}$ is called a complete family of pdf's.

Example: let $X_i \overset{iid}{\sim}$ Poisson (θ); $f(x)=\dfrac{e^{-\theta}\cdot\theta^x}{x!}$

we have shown that $Y=\sum X_i$ is a SS with distribution

$$Y\sim \text{Poisson}(n\theta);\quad f(y;\theta)=\frac{(n\theta)^y\cdot e^{-(n\theta)}}{y!}\quad \theta>0$$

Suppose that $E[u(Y)]=0$

We have, $E[u(Y)]=\sum_{y=0}^{\infty} u(y)\dfrac{(n\theta)^y\cdot e^{-(n\theta)}}{y!}$

$$=e^{-n\theta}\cdot\left[u(0)+u(1)\cdot\frac{(n\theta)}{1!}+u(2)\cdot\frac{(n\theta)^2}{2!}+\cdots\right]=0$$

Since $e^{-n\theta}$ and $\dfrac{(n\theta)^y}{y!}$ terms are NON-zero, then each coefficient $u(y)$ must be zero. So, $E[u(Y)]=0$ requires $u(0)=u(1)=u(2)=\cdots\cdots=0$

Example: $X\sim$ Bernoulli(θ); $f(x;\theta)=\theta^x\cdot(1-\theta)^{1-x}$ $x=0,1$ $0<\theta<1$

$$E(u(x))=\sum_0^1 u(x)\cdot\theta^x(1-\theta)^{1-x}=\underbrace{u(0)}_{=0}\cdot(1-\theta)+\underbrace{u(1)}_{=0}\cdot\theta=0$$

$\Rightarrow u(0)=u(1)=0$ then, $f(x;\theta)$ is a complete family.

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

Example 4 Let $X \sim Uniform(-\theta, \theta)$, $\theta > 0$

$$f(x) = \frac{1}{2\theta}, \quad -\theta \leq x \leq \theta$$

Since $E(X) = 0$ and $u(x) = X \neq 0$, The family of pdf's is NOT complete

Example 4 likewise, $X \sim Normal(0; \sigma^2)$ is NOT complete because $E(X) = 0$

7.23. If $az^2 + bz + c = 0$ for more than two values of z, then $a = b = c = 0$. Use this result to show that the family $\{b(2, \theta) : 0 < \theta < 1\}$ is complete.

7.24. Show that each of the following families is not complete by finding at least one nonzero function $u(x)$ such that $E[u(X)] = 0$, for all $\theta > 0$.

(a) $f(x; \theta) = \frac{1}{2\theta}, \quad -\theta < x < \theta$, where $0 < \theta < \infty$,

$= 0$ elsewhere.

(b) $N(0, \theta)$, where $0 < \theta < \infty$.

7.23) $X \sim Binomial(n=2; \theta)$; $f(x;\theta) = \binom{2}{x} \theta^x (1-\theta)^{2-x}$ $x = 0,1,2$

$$E[u(X)] = \sum_{x=0}^{2} u(x) \cdot \binom{2}{x} \theta^x (1-\theta)^{2-x}$$

$$= u(0) \cdot (1-\theta)^2 + u(1) \cdot 2\theta(1-\theta) + u(2) \cdot \theta^2$$

$$= u(0) \cdot (1 - 2\theta + \theta^2) + 2\theta\, u(1) - 2\theta^2 u(1) + \theta^2 u(2)$$

$$= \theta^2 \underbrace{[u(0) + u(2) - 2u(1)]}_{=0} + \theta \underbrace{[2u(1) - 2u(0)]}_{=0} + \underbrace{u(0)}_{=0} = 0$$

$u(1) = \frac{1}{2}[u(0) + u(2)]$ $\quad$ $u(1) = u(0) = 0$ (ii) $\quad$ $u(0) = 0$ (i)

$0 = 0 + \frac{u(2)}{2} \Rightarrow u(2) = 0$ (iii)

Then, the family is complete.

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

LEHMANN SCHEFFÉ THEOREM

Let $X_i \overset{iid}{\underset{\theta}{\sim}} \{f(x;\theta): \theta \in \Omega\}$, let $Y = u(X_1, X_2, \ldots, X_n)$ is a SS for θ and let the family $\{g(y,\theta): \theta \in \Omega\}$ is complete. If there is a function of Y which is an unbiased estimator of θ, this function of Y is unique minimum varianced unbiased estimator (UMVUE) of θ.

7.27. Show that the first order statistic Y_1 of a random sample of size n from the distribution having p.d.f. $f(x;\theta) = e^{-(x-\theta)}$, $\theta < x < \infty$, $-\infty < \theta < \infty$, zero elsewhere, is a complete sufficient statistic for θ. Find the unique function of this statistic which is the unbiased minimum variance estimator of θ.

7.27) $f(x) = e^{-(x-\theta)} \quad x > \theta$

$$F(x) = \int_\theta^x f(w)\,dw = 1 - e^{-(x-\theta)}$$

$$Y_1 : \; g_1(y) = n.[1-F(y)]^{n-1} \cdot f(y) = n.\left[e^{-(y-\theta)}\right]^{n-1} \cdot e^{-(y-\theta)}$$

$$g_1(y) = n.\, e^{-n(y-\theta)} \quad y > \theta$$

$$E[u(Y)] = \int_\theta^\infty u(Y).\,n.\,e^{-n.(y-\theta)}\,dy = 0$$

Since $n.e^{-n(y-\theta)}$ is positive for $y > \theta$, the integral is zero only if $u(Y) = 0$. The family is complete.

$$E(Y_1) = \int_\theta^\infty y.\,n.\,e^{-n(y-\theta)}\,dy = n.\int_\theta^\infty y.\,e^{-n(y-\theta)}\,dy \qquad \begin{matrix} du = dy \\ u = y \\ dv = e^{-n(y-\theta)}dy \\ v = \frac{-1}{n}e^{-n(y-\theta)} \end{matrix}$$

$$= n\left(\frac{-y}{n}.\,e^{-n(y-\theta)} + \frac{1}{n}\int e^{-n(y-\theta)}\,dy\Big]_\theta^\infty\right) = \theta + n \Rightarrow E(Y_1 - n) = \theta$$

$$\hat\theta = Y_1 - n$$

Alp Giray Özen | 0533 549 91 08 | alp@lecturemania.com | www.lecturemania.com

The Exponential Class of PDF's

let $\{f(x;\theta) : \theta \in \Omega\}$ $\Omega = \{\theta : \gamma < \theta < \delta\}$

CONTINUOUS

$$f(x;\theta) = \exp[p(\theta)K(x) + S(x) + q(\theta)] \quad a < x < b$$

Such a pdf is said to be a member of exponential class of pdf's. Regularity Conditions;

(i) Neither a nor b depends upon θ

(ii) $p(\theta)$ is a nontrivial continuous function of θ

(iii) Each of $K'(x) \not\equiv 0$ and $S(x)$ is cont. func. of x

DISCRETE

$$f(x;\theta) = \exp[p(\theta)K(x) + S(x) + q(\theta)] \quad x = a_1, a_2, a_3, \ldots$$

Such a pdf is said to represent a regular case of the exponential class of pdf's of discrete type if

The set $\{x : x = a_1, a_2, \ldots\}$ does NOT depend upon θ

$p(\theta)$ is a nontrivial continuous function of θ

$K(x)$ is a nontrivial function of x on $x = a_1, a_2, \ldots$

THEOREM: let $f(x;\theta)$, $\gamma < \theta < \delta$ be a pdf which represents a regular case of exponential class. Then if $X_1, X_2, \ldots, X_n$ is a random sample from $f(x;\theta)$, the statistics $Y = \sum K(X_i)$ is a CSS for θ.

Example: let $X \sim Normal(0;\theta)$. Find a CSS for θ.

Answer: $f(x;\theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{x^2}{2\theta}} = \exp\left[\underbrace{-\frac{1}{2\theta}}_{p(\theta)} \cdot \underbrace{x^2}_{K(x)} - \underbrace{\ln\sqrt{2\pi\theta}}_{q(\theta)}\right]$ $x \in \mathbb{R}$, $\theta > 0$, $S(x)=0$

Then, $y = \sum x_i^2$ is a CSS for θ.

Example: let $X \sim Normal(\theta;\sigma^2)$. Find a CSS for θ

Answer: $f(x;\theta) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(x-\theta)^2}{2\sigma^2}\right]$, $x \in \mathbb{R}$, $\theta \in \mathbb{R}$

$$= \exp\left(\underbrace{\frac{\theta}{\sigma^2}}_{=p(\theta)} \cdot \underbrace{X}_{=K(x)} \underbrace{- \frac{x^2}{2\sigma^2} - \ln\sqrt{2\pi\sigma^2}}_{S(x)} \underbrace{- \frac{\theta^2}{2\sigma^2}}_{=q(\theta)}\right)$$

Then, $y = \sum X_i$ is a CSS for θ.

Example: let $X \sim Poisson(\theta)$. Find UMVUE of θ.

Answer: $f(x;\theta) = \frac{\theta^x e^{-\theta}}{x!} = \exp\left[\underbrace{(\ln\theta)}_{p(\theta)} . \underbrace{x}_{K(x)} \underbrace{- \ln x!}_{S(x)} \underbrace{- \theta}_{q(\theta)}\right]$ $x = 0,1,2 \ldots$

Then $y = \sum X_i$ is a CSS for θ.

$E(y) = E(\sum X_i) = \sum E(X_i) = \sum \theta = n\theta$

$E\left(\frac{y}{n}\right) = E(\bar{X}) = \theta \Rightarrow \bar{X}$ is UMVUE of θ.

Example (Cont) → Find UMVUE of θ for $X \sim Normal(0;\theta)$

Answer: $E(y) = E\left[\sum X_i^2\right] = Var\left[\sum X_i\right] + \underbrace{E^2(\sum X_i)}_{=0} = \sum Var(X_i) = n.\theta$

$E\left(\frac{y}{n}\right) = \theta \Rightarrow \frac{y}{n} = \frac{\sum X_i^2}{n}$ is UMVUE of θ.

7.29. Write the p.d.f.

$$f(x;\theta) = \frac{1}{6\theta^4} x^3 e^{-x/\theta}, \quad 0 < x < \infty, \ 0 < \theta < \infty,$$

zero elsewhere, in the exponential form. If $X_1, X_2, \ldots, X_n$ is a random sample from this distribution, find a complete sufficient statistic Y_1 for θ and the unique function $\varphi(Y_1)$ of this statistic that is the unbiased minimum variance estimator of θ. Is $\varphi(Y_1)$ itself a complete sufficient statistic?

7.30. Let $X_1, X_2, \ldots, X_n$ denote a random sample of size $n > 1$ from a distribution with p.d.f. $f(x;\theta) = \theta e^{-\theta x}$, $0 < x < \infty$, zero elsewhere, and $\theta > 0$. Then $Y = \sum_1^n X_i$ is a sufficient statistic for θ. Prove that $(n-1)/Y$ is the unbiased minimum variance estimator of θ.

7.39. Let $X_1, X_2, \ldots, X_n$, $n > 2$, be a random sample from the binomial distribution $b(1, \theta)$.

(a) Show that $Y_1 = X_1 + X_2 + \cdots + X_n$ is a complete sufficient statistic for θ.

(b) Find the function $\varphi(Y_1)$ which is the unbiased minimum variance estimator of θ.

(c) Let $Y_2 = (X_1 + X_2)/2$ and compute $E(Y_2)$.

(d) Determine $E(Y_2 | Y_1 = y_1)$.

7.29) $f(x;\theta) = \frac{1}{6\theta^4} \cdot x^3 \cdot e^{-x/\theta} = \exp\Big[\underbrace{-\frac{1}{\theta}}_{p(\theta)} \cdot \underbrace{x}_{K(x)} + \underbrace{3\ln x - \ln 6}_{S(x)} \underbrace{- 4\ln\theta}_{q(\theta)}\Big]$

Then, $y = \sum X$ is a CSS for θ.

Note that $X \sim \text{Gamma}(\alpha = 4; \beta = \theta)$

$E(X) = \alpha\beta = 4\theta$ $\qquad f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} \cdot e^{-x/\beta}$

Then, $E(\sum X) = \sum E(X) = \sum 4\theta = 4n\theta$

$E\left(\frac{\sum X}{4n}\right) = \theta \Rightarrow \varphi(y) = \frac{y}{4n} = \frac{\bar{X}}{4}$ is UMVUE for θ

Yes, because $\varphi(y)$ is a 1-1 function of CSS y

7.30) $f(x;\theta)=\theta\cdot e^{-\theta x}=\exp\{\underbrace{-\theta}_{p(\theta)}\underbrace{x}_{K(x)}+\underbrace{\ln\theta}_{q(\theta)}\}$ $\quad S(x)=0$

Then, $Y=\sum X_i$ is a CSS for θ.

Note that, $X\sim Exp\left(\frac{1}{\theta}\right)=Gamma\left(1,\frac{1}{\theta}\right)$

Remember, if $X_i\overset{ind.}{\sim}Gamma(\alpha_i;\beta)\Rightarrow\sum X_i\sim Gamma(\sum\alpha_i;\beta)$

Then, $Y=\sum X_i\sim Gamma\left(n;\frac{1}{\theta}\right)$

$$f(y)=\frac{1}{\Gamma(n)\cdot\underbrace{\left(\frac{1}{\theta}\right)^n}_{=\theta\cdot\theta^{n-1}}}\cdot y^{n-1}\cdot e^{-y\theta}$$

$$E\left(\frac{n-1}{Y}\right)=\int_0^\infty\frac{n-1}{y}\cdot\frac{\theta^n}{\underbrace{\Gamma(n)}_{=(n-1)!}}\cdot y^{n-1}\cdot e^{-y\theta}dy$$

$(n-2)!$

$$=\theta\cdot\underbrace{\int_0^\infty\frac{\theta^{n-1}}{\Gamma(n-1)}\cdot y^{n-2}\cdot e^{-y\theta}dy}_{\text{pdf of } Gamma\left(n-1,\frac{1}{\theta}\right)}=\theta$$

7.39) a) $f(x;\theta)=\theta^x\cdot(1-\theta)^{1-x}=\theta^x\cdot\left(\frac{1}{1-\theta}\right)^{x-1}=\left(\frac{\theta}{1-\theta}\right)^x\cdot(1-\theta)$

$$=\exp\left[\underbrace{\ln\left(\frac{\theta}{1-\theta}\right)}_{p(\theta)}\cdot\underbrace{x}_{K(x)}+\underbrace{\ln(1-\theta)}_{q(\theta)}\right]\quad x=0,1;\ 0<\theta<1$$

Then, $Y_1=\sum X_i$ is a CSS for θ

b) $E(Y_1)=E(\sum X_i)=\sum E(X_i)=\sum\theta=n\theta$

Then; $\frac{Y_1}{n}=\bar{X}$ is UMVUE of θ.

c) $E(Y_2) = E\left(\frac{X_1+X_2}{2}\right) = \frac{1}{2}\left[E(X_1)+E(X_2)\right] = \frac{1}{2}\cdot 2\theta = \theta$

d) $E(Y_2 / Y_1 = y_1) = \varphi(Y_1) = \bar{X}$

FUNCTIONS of a PARAMETER

Example: let $X \sim Bernoulli(\theta)$.

a) what is UMVUE of θ^2?

b) what is UMVUE of $Var(X)$?

Answer: we have shown in 7.39 that $\bar{X}$ is UMVUE of θ.

We have, $E(X) = \theta$; $Var(X) = \theta.(1-\theta) = \theta - \theta^2$

$$Y = \sum X_i \Rightarrow E(Y) = n\theta \quad Var(Y) = n\theta(1-\theta)$$

a) $E(Y^2) = Var(Y) + E^2(Y) = n\theta(1-\theta) + n^2\theta^2 = \underbrace{n\theta}_{=E(Y)} - n\theta^2 + n^2\theta^2$

$$E(Y^2) = E(Y) + \theta^2(n^2-n)$$

$$E\left[\frac{Y^2-Y}{n^2-n}\right] = \theta^2 \Rightarrow \frac{Y^2-Y}{n^2-n} \text{ is UMVUE of } \theta^2$$

b) $E\left(\frac{Y}{n}\right) = \theta$; $E\left(\frac{Y^2-Y}{n^2-n}\right) = \theta^2$ Then $E\left[\frac{Y}{n} - \frac{Y^2-Y}{n^2-n}\right] = \theta - \theta^2$

So; $\frac{Y}{n} - \frac{Y(1-Y)}{n(1-1)} = \frac{Y}{n}\left[1 - \frac{1-Y}{n-1}\right]$ is UMVUE of $Var(X)$.

7.40. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution that is $N(\theta, 1)$, $-\infty < \theta < \infty$. Find the unbiased minimum variance estimator of θ^2.

Hint: First determine $E(\bar{X}^2)$.

7.41. Let $X_1, X_2, \ldots, X_n$ denote a random sample from a distribution that is $N(0, \theta)$. Then $Y_1 = \sum X_i^2$ is a complete sufficient statistic for θ. Find the unbiased minimum variance estimator of θ^2.

7.40) We have shown that $\bar{X}$ is UMVUE of θ.

$E(\bar{X}^2) = Var(\bar{X}) + E^2(\bar{X}) = \frac{1}{n} + \theta^2$

Because $X \sim Normal(\mu, \sigma^2) \Rightarrow \bar{X} \sim Normal(\mu, \frac{\sigma^2}{n})$

Then; $E(\bar{X}^2 - \frac{1}{n}) = \theta^2 \Rightarrow \bar{X}^2 - \frac{1}{n}$ is UMVUE of θ^2

7.41) $X \sim Normal(0;\theta)$

We have shown that $Y = \frac{\Sigma X^2}{n}$ is UMVUE of θ.

Also Remember $W = \frac{\Sigma X^2}{\theta} \sim \chi^2_n$ and $E(W) = n$, $Var(W) = 2n$

Y is a CSS of $\theta \Rightarrow Y^2$ is a CSS of θ^2

$$E(Y^2) = Var(Y) + E^2(Y) = Var\left[\frac{\Sigma X^2}{n}\right] + E^2\left[\frac{\Sigma X^2}{n}\right]$$

$$= Var\left[\frac{\theta}{n} \cdot \frac{\Sigma X^2}{\theta}\right] + E^2\left[\frac{\theta}{n} \cdot \frac{\Sigma X^2}{\theta}\right] = \frac{\theta^2}{n^2} Var(W) + \frac{\theta^2}{n^2} E^2(W)$$

$$= \frac{\theta^2}{n^2}\left[2n + n^2\right] = \frac{\theta^2}{n^2} \cdot n(n+2) = \frac{n+2}{n} \cdot \theta^2$$

Then, $E\left[\frac{n}{n+2} Y^2\right] = \theta^2$ and $\frac{n}{n+2}\left[\frac{\Sigma X^2}{n}\right]^2$ is UMVUE of θ^2

Example: Let $X \sim \{f(x;\theta) : \theta \in \Omega\}$

$E(X) = 2\theta$ and $Var(X) = 2\theta^2$

Also let $Y = \Sigma X_i$ is a CSS for θ.

a) Find UMVUE of θ

b) Find UMVUE of θ^2

c) Find $E[(X_1 + X_2) | Y]$

Answer a) $E(y) = E(\Sigma X_i) = \Sigma E(X_i) = \sum 2\theta = 2n\theta$

Then, $E\left(\frac{y}{2n}\right) = \theta$ and $\frac{\Sigma X_i}{2n} = \frac{\bar{X}}{2}$ is UMVUE of θ

b) $E(y) = 2n\theta$ and $Var(y) = 2n\theta^2$

$E(y^2) = E[(\Sigma X_i)^2] = Var(y) + E^2(y) = 2n\theta^2 + 4n^2\theta^2 = \theta^2(4n^2+2n)$

Then, $E\left[\frac{y^2}{4n^2+2n}\right] = \theta^2$ and $\frac{(\Sigma X_i)^2}{2n(2n+1)}$ is UMVUE of θ^2

c) let $\varphi(y) = E[X_1 + X_2 \mid y]$

$E[\varphi(y)] = E\left[E[X_1+X_2 \mid y]\right] = E(X_1+X_2) = 2\theta + 2\theta = 4\theta$

Since y is CSS, thee can be only one function of y which is Unbiased Estimator of 4θ

$E(y) = 2n\theta \Rightarrow E\left(\frac{2y}{n}\right) = 4\theta$

Then; $\varphi(y) = E[X_1 + X_2 \mid y] = \frac{2y}{n} = 2\bar{X}$

Example let $X \sim f(x;\theta) = 3\theta^3 \cdot x^4 \quad x \geq \theta \quad \theta \geq 0$

$$E(X) = \frac{3\theta}{2} \qquad Var(X) = \frac{3\theta^2}{4}$$

Let $y = \min(X_i)$ has pdf $g(y;\theta) = 3n \cdot \theta^{3n} \cdot y^{-(3n+1)} \quad y \geq 0$

and $E(y) = \frac{3n\theta}{3n-1}$

a) Find a SS for θ

b) Show that y is a CSS

c) Find UMVUE of θ

d) Find UMVUE of $Var(X)$

Answer 4

a) $f(x;\theta) = 3\theta^3 \cdot x^4 \cdot I_{(\theta,\infty)}(x)$

$$L(x_1,x_2,\dots,x_n;\theta) = 3^n \cdot \theta^{3n} \cdot \prod_{i=1}^{n} x_i^4 \cdot \prod I_{(\theta,\infty)}(x_i)$$

$$= \underbrace{3^n \cdot \theta^{3n} \cdot I_{(\theta,\infty)}\min(x_i)}_{k_1(y,\theta)} \cdot \underbrace{\prod_{i=1}^{n} x_i^4}_{k_2(x_1,x_2,\dots,x_n)} \qquad y=\min(x_i)$$

Then, by factorization thm, $y=\min(X_i)$ is a SS for θ

b) $E[u(y)] = 0$ iff $u(y)=0 \quad \forall\, y \geq \theta,\ \theta \geq 0$

$$E[u(y)] = \int_{\theta}^{\infty} u(y) \cdot \underbrace{3n\,\theta^{3n}}_{\neq 0}\, y^{-(3n+1)}\, dy = 0$$

$$\int_{\theta}^{\infty} u(y) \cdot y^{-(3n+1)}\, dy = 0$$

If this integral is zero, its derivative w.r.t θ is also zero. Remember, Leibnitz's Theorem;

$$\frac{d}{dx}\int_{\alpha(x)}^{\beta(x)} f(x)\,dx = \left\{f[\beta(x)]\frac{d}{dx}\beta(x)\right\} - \left\{f[\alpha(x)]\frac{d}{dx}\alpha(x)\right\} + \int_{\alpha(x)}^{\beta(x)} \frac{d}{dx} f(x)\,dx$$

So, we have; $\frac{d}{d\theta}\int_{\theta}^{\infty} u(y) \cdot y^{-(3n+1)}\, dy = 0$

$$\underbrace{\lim_{c\to\infty}\left(u(c) \cdot c^{-(3n+1)} \cdot \underbrace{\frac{d}{d\theta}c}_{=0}\right)}_{=0} - u(\theta) \cdot \theta^{-(3n+1)} \cdot 1 + \int_{0} \underbrace{\frac{d}{d\theta}u(y)y^{-3n-1}}_{=0}\, dy = 0$$

$$-u(\theta) \cdot \theta^{-3n-1} = 0$$

$u(\theta) = 0$ Then $y=\min(X_i)$ is a CSS for θ

c) $E(y) = \frac{3n\theta}{3n-1} \Rightarrow E\left[\frac{3n-1}{3n} y\right] = \theta$

Then, $\frac{3n-1}{3n} y$ is UMVUE of θ.

d) $g(y) = 3n \cdot \theta^{3n} y^{-(3n+1)} \quad y \geq \theta$

y is CSS for $\theta \Rightarrow y^2$ is CSS for θ^2

$$E(y^2) = \int_\theta^\infty y^2 \cdot 3n\theta^{3n} y^{-(3n+1)} dy = 3n \cdot \theta^{3n} \int_\theta^\infty y^{-3n+1} dy$$

$$= 3n \cdot \theta^{3n} \cdot \left.\frac{y^{-3n+2}}{-3n+2}\right]_\theta^\infty = \frac{3n}{3n-2} \cdot \theta^2$$

Then, $E\left[\frac{3n-2}{3n} y\right] = \theta^2$ and $E\left[\frac{3}{4} \frac{3n-2}{3n} y\right] = \frac{3\theta^2}{4} = Var(X)$

So, $\frac{3n-2}{4n} \min(X_i)$ is UMVUE of $Var(X)$

Example Let $X_i \sim Poisson(\theta)$

Find UMVUE of $P(X=0)$

Answer: $X \sim Poisson(\theta) \Rightarrow y = \sum X_i \sim Poisson(n \cdot \theta)$

we have shown that $\bar{X} = \frac{y}{n}$ is UMVUE of θ.

$$f(x) = \frac{e^{-\theta} \cdot \theta^x}{x!}$$

So, we want the UMVUE of $P(X=0) = e^{-\theta}$

Since $y = \sum X_i$ is a CSS for θ, there must be a unique function $\varphi(y)$ which is UMVUE of $e^{-\theta}$.

Namely, we want $E[\varphi(y)] = e^{-\theta}$

$$E[\varphi(y)] = \sum_{y=0}^{\infty} \varphi(y) \cdot g(y) = \sum_{y=0}^{\infty} \varphi(y) \cdot \frac{e^{-n\theta} \cdot n\theta^{y}}{y!} = e^{-\theta}$$

$$\sum_{y=0}^{\infty} \frac{\varphi(y) \cdot e^{-n\theta} \cdot e^{\theta} \cdot (n\theta)^{y}}{y!} = 1$$

$$\sum_{y=0}^{\infty} \varphi(y) \cdot \underbrace{\frac{e^{-\theta(n-1)} \cdot (n\theta)^{y}}{y!}}_{\text{pmf of Poisson}(\theta(n-1))} = 1$$

$$\text{So; } (n\theta)^{y} = n^{y} \theta^{y} \cdot \frac{(n-1)^{y}}{(n-1)^{y}} = \underbrace{\left(\frac{n}{n-1}\right)^{y}}_{\varphi(y)} \cdot [\theta \cdot (n-1)]^{y}$$

$$\varphi(y) = \left(\frac{n}{n-1}\right)^{\Sigma x_i} \text{ is UMVUE of } e^{-\theta}$$