Southern University of Science and Technology
Department of Mathematics

MA204: Mathematical Statistics

Tutorial 6: Examples/Solutions

A. UMVUE and Efficient Estimator

A.1 An estimator of θ is called a **uniformly minimum variance unbiased estimator** (UMVUE) if it is unbiased and has the smallest variance among all unbiased estimators of θ.

A.2 An unbiased estimator for θ is an **efficient estimator** if it has variance equal to the Cramér–Rao lower bound.

A.3 Obviously, an efficient estimator for θ is a UMVUE for θ. However, a UMVUE for θ does not necessarily imply that it is an efficient estimator for θ.

B. Sufficiency

B.1 **Definition**

A statistic $T(\boldsymbol{X})$ is a **sufficient statistic** of θ if the conditional distribution of $X_1, \ldots, X_n$, given $T = t$, does not depend on θ for any value of t. For discrete cases, this means

$$\Pr\{X_1 = x_1 \ldots, X_n = x_n; \theta \mid T(\boldsymbol{X}) = t\} = h(\boldsymbol{x})$$

does not depend on θ.

B.2 **Factorization Theorem**

A statistic $T(\boldsymbol{X})$ is a sufficient statistic of θ if and only if the joint pdf (or pmf) can be written in the form

$$f(x_1, \ldots, x_n; \theta) = f(\boldsymbol{x}; \theta) = g(T(\boldsymbol{x}); \theta) \times h(\boldsymbol{x}),$$

where $h(\boldsymbol{x})$ does not depend on θ, $g(T; \theta)$ is a function of both T and θ, and it depends on $x_1, \ldots, x_n$ only through T.

C. Completeness

C.1 **Definition**

A statistic $T(\boldsymbol{X})$ is said to be **complete** if

$$E[h(T)] = 0,\ \forall\theta \in \Theta \quad \Rightarrow \quad \Pr\{h(T) = 0\} = 1,\ \forall\theta \in \Theta,$$

where a function $h(T)$ is a statistic.

C.2 **Lehmann–Scheffé Theorem**

Let $T(\boldsymbol{X})$ be a complete sufficient statistic of θ. If $g(T)$ is an unbiased estimator of $\tau(\theta)$, then $g(T)$ is the unique UMVUE for $\tau(\theta)$.

Example T6.1 (A geometric distribution). Let $X_1, \ldots, X_n$ be a random sample from the geometric distribution with density

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

(a) Show that $T(\boldsymbol{X}) = \sum_{i=1}^n X_i$ is a sufficient statistic of θ.

(b) Show that $T(\boldsymbol{X}) \sim \text{NBinomial}(n, \theta)$, whose density is

$$g(t;\theta) = \binom{n+t-1}{t}\theta^n(1-\theta)^t, \quad t = 0,\ 1,\ \ldots.$$

Hint: For a positive integer n,

$$(x+a)^{-n} = \sum_{y=0}^{\infty}(-1)^y\binom{n+y-1}{y}x^y a^{-n-y}, \quad \text{for } |x| < a.$$

(c) Show that $T(\boldsymbol{X})$ is complete for θ.

(d) Find the UMVUE for $\tau(\theta) = (1-\theta)/\theta$ by the Lehmann–Scheffé Theorem. Show that it is also an efficient estimator for $\tau(\theta)$.

(e) Show that $I_{\{0\}}(X_j)$ is an unbiased estimator for θ.

(f) Find the UMVUE for θ.

Solution: (a) The joint pmf of $X_1, \dots, X_n$ is

$$f(\boldsymbol{x};\theta) = \prod_{i=1}^{n} \theta(1-\theta)^{x_i} = \theta^n (1-\theta)^{T(\boldsymbol{x})} \times 1 = g(T(\boldsymbol{x});\theta) \times h(\boldsymbol{x}),$$

where $g(t;\theta) = \theta^n(1-\theta)^t$ and $h(\boldsymbol{x}) = 1$. Therefore, $T(\boldsymbol{X})$ is sufficient for θ.

(b) The mgf of X_i $(i = 1, \dots, n)$ is

$$\begin{aligned} M_{X_i}(t) &= E(\mathrm{e}^{tX_i}) \\ &= \sum_{x=0}^{\infty} \mathrm{e}^{tx}\theta(1-\theta)^x \\ &= \theta \sum_{x=0}^{\infty} [\mathrm{e}^t(1-\theta)]^x \\ &= \frac{\theta}{1-\mathrm{e}^t(1-\theta)}, \quad t < -\log(1-\theta). \end{aligned}$$

So the mgf of $T(\boldsymbol{X})$ is

$$M_T(t) = \prod_{i=1}^{n} M_{X_i}(t) = \left[\frac{\theta}{1-e^t(1-\theta)}\right]^n, \quad t < -\log(1-\theta).$$

On the other hand, let $Y \sim \text{NBinomial}(n,\theta)$. The mgf of Y is

$$\begin{aligned} M_Y(t) &= E(\mathrm{e}^{tY}) \\ &= \sum_{y=0}^{\infty} \mathrm{e}^{ty}\binom{n+y-1}{y}\theta^n(1-\theta)^y \\ &= \theta^n \times \sum_{y=0}^{\infty} \binom{n+y-1}{y}[\mathrm{e}^t(1-\theta)]^y \\ &= \theta^n \times [1-\mathrm{e}^t(1-\theta)]^{-n} \\ &= \left(\frac{\theta}{1-\mathrm{e}^t(1-\theta)}\right)^n, \qquad t < -\log(1-\theta). \end{aligned}$$

Since $M_T(t) = M_Y(t)$, $T(\boldsymbol{X}) \sim \text{NBinomial}(n,\theta)$.

(c) Assume that

$$\begin{aligned} & E(h(T)) = 0 \\ \Rightarrow & \sum_{t=0}^{\infty} h(t)\binom{n+t-1}{t}\theta^n(1-\theta)^t = 0 \\ \Rightarrow & \sum_{t=0}^{\infty} \binom{n+t-1}{t} h(t)(1-\theta)^t = 0, \quad 0 < \theta < 1. \end{aligned} \tag{1}$$

The equation (1) is a polynomial of $(1-\theta)$ and $(1-\theta)$ must be nonzero. The fact that it equals to zero implies that all its coefficients are zero, i.e.,

$$\binom{n+t-1}{t} h(t) = 0, \quad t = 0, 1, \ldots .$$

Hence, $h(T) = 0$, i.e. $\Pr\{h(T) = 0\} = 1$. Therefore, $T(\boldsymbol{X})$ is complete for θ.

(d)

$$E[T(\boldsymbol{X})] = \frac{d\, M_T(t)}{dt}\bigg|_{t=0} = \frac{ne^t\theta^n(1-\theta)}{[1-e^t(1-\theta)]^{n+1}}\bigg|_{t=0} = n \cdot \frac{1-\theta}{\theta} = n \cdot \tau(\theta).$$

Denote $\bar{X} = T(\boldsymbol{X})/n$. It implies that $E(\bar{X}) = \tau(\theta)$, and thus $\bar{X}$ is an unbiased estimator for $\tau(\theta)$. Because $T(\boldsymbol{X})$ is sufficient and complete, according to the Lehmann–Scheffé Theorem, $\bar{X}$ is the unique UMVUE for $\tau(\theta) = (1-\theta)/\theta$.

To prove that $\bar{X}$ is an efficient estimator for $\tau(\theta)$, we need to show that $\mathrm{Var}(\bar{X})$ equals to the Cramér–Rao lower bound. Since

$$E([T(\boldsymbol{X})]^2) = \frac{d^2\, M_T(t)}{dt^2}\bigg|_{t=0} = \frac{ne^t\theta^n(1-\theta)[1-ne^t(1-\theta)]}{[1-e^t(1-\theta)]^{n+2}}\bigg|_{t=0} = \frac{n(1-\theta)[1-n(1-\theta)]}{\theta^2}.$$

we have

$$\therefore\ \mathrm{Var}[T(\boldsymbol{X})] = E\{[T(\boldsymbol{X})]^2\} - \{E[T(\boldsymbol{X})]\}^2 = \frac{n(1-\theta)}{\theta^2}.$$

Hence,

$$\mathrm{Var}(\bar{X}) = \frac{\mathrm{Var}(T(\boldsymbol{X}))}{n^2} = \frac{1-\theta}{n\theta^2}.$$

The log-likelihood function is

$$\ell(\theta; \boldsymbol{X}) = \log f(\boldsymbol{X}; \theta) = n \log\theta + T(\boldsymbol{X}) \log(1-\theta).$$

Thus, the Fisher information is

$$I_n(\theta) = E\left[-\frac{d^2\ell(\theta; \boldsymbol{X})}{d\theta^2}\right] = E\left[\frac{n}{\theta^2} + \frac{T(\boldsymbol{X})}{(1-\theta)^2}\right] = \frac{n}{\theta^2(1-\theta)}.$$

Since

$$\tau'(\theta) = \frac{d\tau(\theta)}{d\theta} = -\frac{1}{\theta^2},$$

the Cramér–Rao lower bound is

$$\upsilon(\theta) = \frac{[\tau'(\theta)]^2}{I_n(\theta)} = \frac{1-\theta}{n\theta^2}.$$

Since $\mathrm{Var}(\bar{X}) = \upsilon(\theta)$, $\bar{X}$ is an efficient estimator for $\tau(\theta) = (1-\theta)/\theta$.

(e) Let

$$\boxed{I_{\{0\}}(X_j) = \begin{cases} 1, & \text{if } X_j = 0, \\ 0, & \text{otherwise.} \end{cases}}$$

From the pmf of X_j, we obtain that

$$E[I_{\{0\}}(X_j)] = 1 \times \Pr\{X_j = 0\} + 0 \times (1 - \Pr\{X_j = 0\}) = f(0;\theta) = \theta.$$

Therefore, $I_{\{0\}}(X_j)$ is an unbiased estimator for θ.

(f) Denote

$$T = \sum_{i=1}^{n} X_i \quad \text{and} \quad T_{-j} = \sum_{i=1, i\neq j}^{n} X_i = T - X_j.$$

From (b), we know that $T \sim \mathrm{NBinomial}(n, \theta)$ and $T_{-j} \sim \mathrm{NBinomial}(n-1, \theta)$. So

$$\begin{aligned}
\Pr\{X_j = 0 \mid T = t\} &= \frac{\Pr\{X_j = 0, T = t\}}{\Pr\{T = t\}} \\
&= \frac{\Pr\{X_j = 0, T - X_j = T_{-j} = t\}}{\Pr\{T = t\}} \\
&= \frac{\Pr\{X_j = 0\}\Pr\{T_{-j} = t\}}{\Pr\{T = t\}} \quad \text{(since } X_j \text{ and } T_{-j} \text{ are independent)} \\
&= \frac{\theta \cdot \binom{n+t-2}{t}\theta^{n-1}(1-\theta)^t}{\binom{n+t-1}{t}\theta^n(1-\theta)^t} \\
&= \frac{n-1}{t+n-1}.
\end{aligned}$$

Let

$$g(T) = E[I_{\{0\}}(X_j) \mid T] = \Pr\{X_j = 0 \mid T\} = \frac{n-1}{T+n-1}.$$

Since $E[g(T)] = E\{E[I_{\{0\}}(X_j) \mid T]\} = E[I_{\{0\}}(X_j)] = \theta$, $g(T)$ is an unbiased estimator for θ. Because T is sufficient and complete, according to the Lehmann–Scheffé Theorem,

$$g(T) = \frac{n-1}{\sum_{i=1}^{n} X_i + n - 1}$$

is the unique UMVUE for θ. ‖