L07 Efficiency and relative efficiency

1. Efficiency and relative efficiency

   (1) Efficiency

   For statistic $T \sim (E_\theta(T),\, \mathrm{var}_\theta(T))$, $\mathrm{eff}_\theta(T) = \dfrac{\mathrm{CRLB}_{\theta(E_\theta(T))}}{\mathrm{var}_\theta(T)}$ is the efficiency.

   Comments: T is a statistic, but not a vector of statistics. $\mathrm{effF}_\theta(T)$ is a function of θ and $0 \le \mathrm{eff}_\theta(T) \le 1$.

   (2) Property

   $\mathrm{eff}_\theta(T) \equiv 1 \Longleftrightarrow \mathrm{var}_\theta(T) \equiv \mathrm{CRLB}_\theta(E_\theta(T)) \Longleftrightarrow T$ is efficient $\Longrightarrow T$ is MVUE for $E_\theta(T)$.

   (3) Relative efficiency

   For estimators T_1 and T_2 for $\eta_1(\theta)$ and $\eta_2(\theta)$, $e(T_1,\, T_2) = \dfrac{\mathrm{MSE}_{\theta(T_2)}}{\mathrm{MSE}_{\theta(T_1)}}$ is the relative efficiency of T_1 to T_2.

   Comments: T_1 and T_2 could be of different dimensions. $e_\theta(T_1,\, T_2)$ is a function of θ and $e_\theta(T_1,\, T_2) \ge 0$.

$$
\begin{aligned}
e_\theta(T_1,\, T_2) < 1 &\iff \mathrm{MSE}_\theta(T_1) > \mathrm{MSE}_\theta(T_2),\\
e_\theta(T_1,\, T_2) = 1 &\iff \mathrm{MSE}_\theta(T_1) = \mathrm{MSE}_\theta(T_2),\\
e_\theta(T_1,\, T_2) > 1 &\iff \mathrm{MSE}_\theta(T_1) < \mathrm{MSE}_\theta(T_2).
\end{aligned}
$$

   These comparisons make sense only if T_1 and T_2 are for the same $\eta(\theta)$.

   (4) Properties

   If both T_1 and T_2 are unbiased estimators for their targeted parameters $\eta_1(\theta)$ and $\eta_2(\theta)$, then $e_\theta(T_1,\, T_2) = \dfrac{\text{Total-variance}(T_2)}{\text{Total-variance}(T_1)}$.

   If both T_1 and T_2 are unbiased estimators for their targeted parameters in R, then $e_\theta(T_1,\, T_2) = \dfrac{\mathrm{var}_\theta(T_2)}{\mathrm{var}_\theta(T_1)}$.

   If both T_1 and T_2 are unbiased estimators for the same parameter in R, then $e_\theta(T_1,\, T_2) = \dfrac{[\mathrm{CRLB}_{\theta(E_\theta(T_1))}/\mathrm{var}_\theta(T_1)]}{\mathrm{CRLB}_{\theta(E_\theta(T_1))}/\mathrm{var}_\theta(T_2)]} = \dfrac{\mathrm{eff}_\theta(T_1)}{\mathrm{eff}_\theta(T_2)}$.

   Ex1: Let $X_1, ..., X_n$ be random sample of size $n > 2$ from Poisson(λ) with mean λ and variance λ. By HW03, $\mathrm{CRLB}(\lambda) = \frac{\lambda}{n}$.

   For $\frac{X_1+X_2}{2} \sim \left(\lambda,\, \frac{\lambda}{2}\right)$, $\mathrm{eff}\left(\frac{X_1+X_2}{2}\right) = \dfrac{\mathrm{CRLB}(\lambda)}{\mathrm{var}\left(\frac{X_1+X_2}{2}\right)} = \frac{\lambda/n}{\lambda/2} = \frac{2}{n} < 1$.

   $\overline{X} \sim \left(\lambda,\, \frac{\lambda}{n}\right)$, $\mathrm{eff}(\overline{X}) = \dfrac{\mathrm{CRLB}(\lambda)}{\mathrm{var}(\overline{X})} = \frac{\lambda/n}{\lambda/n} = 1$. Hence $\overline{X}$ is efficient and is MVUE for λ.

   $e\left(\frac{X_1+X_2}{2},\, \overline{X}\right) = \dfrac{\mathrm{eff}\left(\frac{X_1+X_2}{2}\right)}{\mathrm{eff}(\overline{X})} = \dfrac{\mathrm{var}(\overline{X})}{\mathrm{var}\left(\frac{X_1+X_2}{2}\right)} = \frac{2}{n}$.

2. More on information

   (1) Same information

   If $Y = g(X)$ is a 1-1 mapping, then $I_X(\theta) = I_Y(\theta)$.

Proof: Consider continuous case only. Suppose $X \in A \Longleftrightarrow Y \in g(A) = B$.

By the substitution of $x = g^{-1}(y)$ in the integration,

$$P(X \in A) = \iint_A f(x;\, \theta)\, dx = \iint_B f(g^{-1}(y),\, \theta)\, \text{abs}|J|\, dy = P(Y \in B).$$

$$\begin{aligned} \text{So} \qquad & f_Y(y;\, \theta) = f_X(x;\, \theta)\text{abs}|J| \Longrightarrow \ln f_Y(y;\, \theta) = \ln f_X(x;\, \theta) + \ln \text{abs}|J| \\ \Longrightarrow \quad & \nabla \ln f_Y(y;\, \theta) = \nabla \ln f_X(x;\, \theta) \Longrightarrow V(Y;\, \theta) = V(X;\, \theta) \\ \Longrightarrow \quad & I_X(\theta) = I_Y(\theta). \end{aligned}$$

Ex2: $X \sim N(\mu,\, I)$ and $Y = AX + b \sim N(A\mu + b,\, AA^T)$ where A is non-singular, verify that $I_X(\mu) = I_Y(\mu)$.

$f_X(x;\, \mu) = \frac{1}{(2\pi)^{p/2}} \exp\left[\frac{1}{-2}(x-\mu)^T(x-\mu)\right]$.

So $V(X;\, \mu) = \frac{\partial}{\partial \mu} \frac{1}{-2}(X-\mu)^T(X-\mu) \stackrel{*}{=} X - \mu \sim N(0,\, I)$. Thus $I_X(\mu) = I$.

$f_Y(y;\, \theta) = \frac{1}{(2\pi)^{p/2}|AA^T|^{1/2}} \exp\left[\frac{1}{-2}(Y - A\mu - b)^T(AA^T)^{-1}(Y - A\mu - b)\right]$.

$$\begin{aligned} \text{So } V(Y;\, \mu) \quad &= \quad \nabla \ln f_Y(Y;\, \mu) = \frac{\partial}{\partial \mu}\frac{1}{-2}(Y - A\mu - b)^T(AA^T)^{-1}(Y - A\mu - b) \\ &\stackrel{*}{=} \quad A^T(AA^T)^{-1}(Y - A\mu - b) \sim A^{-1}N(0,\, AA^T) = N(0,\, I). \end{aligned}$$

Thus $I_Y(\mu) = I = I_X(\mu)$.

Comment: *: Because $\frac{\partial\, Ax}{\partial\, x^T} = A$, $\frac{\partial\, x^TAx}{\partial\, x^T} = x^T(A^T + A)$ and $\frac{\partial\, z}{\partial\, x^T} = \frac{\partial\, z}{\partial\, y^T}\frac{\partial\, y}{\partial,\, x^T}$,

$\frac{\partial\,(x - B\mu)^TA(x-B\mu)}{\partial\, \mu^T} = \frac{\partial\,(x-B\mu)^TA(x-B\mu)}{\partial\,(x-B\mu)^T}\frac{\partial\,(x-B\mu)}{\partial\, \mu^T} = (x - B\mu)^T(A^T + A)(-B)$.

Thus $\frac{\partial\,(x-B\mu)^TA(x-B\mu)}{\partial\, \mu} = -B(A^T + A)(X - B\mu)$.

(2) If $\eta = \eta(\theta)$ is a 1-1 mapping, then $I(\eta) = \left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)^T I(\theta) \left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)$.

Proof: By chain rule, $\frac{\partial \ln f}{\partial\, \eta^T} = \left(\frac{\partial \ln f}{\partial\, \theta^T}\right)\left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)$. By transposing both sides

$\frac{\partial \ln f}{\partial\, \eta} = \left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)^T \left(\frac{\partial \ln f}{\partial\, \theta}\right)$, i.e., $V(X;\, \eta) = \left(\frac{\partial \theta}{\partial \eta^T}\right)^T V(X;\, \theta)$.

Thus $I(\eta) = \left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)^T I(\theta) \left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)$.

Ex3: By definition, $\text{CRLB}(\eta) = \left(\frac{\partial\, \eta}{\partial\, \theta^T}\right)[I(\theta)]^{-1}\left(\frac{\partial\, \eta}{\partial\, \theta^T}\right)^T$. By definition and (2)

$\text{CRLB}(\eta) = [I(\eta)]^{-1} = \left[\left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)^T I(\theta)\left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)\right]^{-1} = \left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)^{-1}[I(\theta)]^{-1}\left[\left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)^T\right]^{-1}$.

By comparison, $\left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)^{-1} = \frac{\partial\, \eta}{\partial\, \theta^T}$ which is true since $I = \frac{\partial\, \theta}{\partial\, \theta^T} = \left(\frac{\partial\, \theta}{\partial\, \eta^T}\right)\left(\frac{\partial\, \eta}{\partial\, \theta^T}\right)$.