CRLB and Rao-Blackwell

Cramer-Rao lower bound

Thm: Let $X_1, \dots, X_n$ be a sample from a distribution with probability function $f(x \mid \theta)$, and $\theta \in \mathbb{R}$. And let $\hat{\theta}$ be an unbiased estimator of θ. Under regularity conditions,

$$\mathrm{Var}(\hat{\theta}) \geq \frac{1}{\mathcal{I}(\theta)}$$

Proof will use C-S inequality:

$$\left(\mathbb{E}\left[(X-\mu_X)(Y-\mu_Y)\right]\right)^2 \leq \mathbb{E}\left[(X-\mu_X)^2\right] \mathbb{E}\left[(Y-\mu_Y)^2\right]$$

$$(\mathrm{Cov}(X,Y))^2 \leq \mathrm{Var}(X)\,\mathrm{Var}(Y)$$

$$(\mathrm{Cov}(\hat{\theta}, u(\theta)))^2 \leq \mathrm{Var}(\hat{\theta})\,\mathrm{Var}(u(\theta))$$
$$= \mathrm{Var}(\hat{\theta})\,\mathcal{I}(\theta)$$

Pf: $\left(\mathrm{Cov}(\hat{\theta}, u(\theta))\right)^2 \leq \mathrm{Var}(\hat{\theta})\, \mathcal{I}(\theta)$

$$\mathrm{Cov}(\hat{\theta}, u(\theta)) = \mathbb{E}[\hat{\theta}\, u(\theta)] - \mathbb{E}[\hat{\theta}]\, \underbrace{\mathbb{E}[u(\theta)]}_{=0}$$

under regularity conditions

$$= \mathbb{E}[\hat{\theta}\, u(\theta)]$$

$$= \mathbb{E}\left[\hat{\theta}\ \frac{\partial}{\partial\theta} \log f(X_1, \dots, X_n \mid \theta)\right] = \mathbb{E}\left[\hat{\theta}\ \frac{\frac{\partial}{\partial\theta} f(X_1, \dots, X_n \mid \theta)}{f(X_1, \dots, X_n \mid \theta)}\right]$$

$$= \int \hat{\theta}(x_1, \dots, x_n) \left[\frac{\partial}{\partial\theta} f(x_1, \dots, x_n)\right] \cdot \frac{1}{\cancel{f(x_1, \dots, x_n)}}\ \cancel{f(x_1, \dots, x_n)}\ dx_1 \cdots dx_n$$

$$= \int \hat{\theta}(x_1, \dots, x_n) \left[\frac{\partial}{\partial\theta} f(x_1, \dots, x_n)\right] dx_1 \cdots dx_n$$

(regularity conditions) $=$

$$\frac{\partial}{\partial\theta} \underbrace{\int \hat{\theta}(x_1, \dots, x_n)\, f(x_1, \dots, x_n)\, dx_1, \dots, dx_n}_{\mathbb{E}[\hat{\theta}]} = \frac{\partial}{\partial\theta} \mathbb{E}[\hat{\theta}]$$

(unbiased) $= \dfrac{\partial}{\partial\theta}\theta = 1$

$$\Rightarrow \quad 1 \leq \mathrm{Var}(\hat{\theta})\, \mathcal{I}(\theta) \quad \Rightarrow \quad \mathrm{Var}(\hat{\theta}) \geq \frac{1}{\mathcal{I}(\theta)} \quad //$$

More generally:

$$\mathrm{Var}(\hat{\theta}) \geq \frac{\left(\frac{\partial}{\partial\theta} \mathbb{E}[\hat{\theta}]\right)^2}{\mathcal{I}(\theta)}$$

(requires regularity conditions, true for biased <u>and</u> unbiased estimators)

<u>Recall</u>: under certain conditions, the MLE $\hat{\theta}$ of θ has the following properties:

① $\hat{\theta} \xrightarrow{P} \theta$ (consistent)

② $\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{d} N(0, \mathcal{I}_1^{-1}(\theta))$

- $\hat{\theta}$ is asymptotically normal
- asymptotically unbiased
- asymptotically achieves CRLB

Rao-Blackwell

Let θ be a parameter of interest, and $\tau(\theta)$ be some function of θ.

Let $\hat{\tau}$ be some unbiased estimator of $\tau(\theta)$, and T be a sufficient statistic for θ.

Let $\tau^* = \mathbb{E}[\hat{\tau} \mid T]$. Then:

① $\mathbb{E}[\tau^*] = \tau$

② $\text{Var}(\tau^*) \leq \text{Var}(\hat{\tau})$

<u>Reminder</u>: If the conditional distribution of $X_1, \ldots, X_n \mid T$ does <u>not</u> depend on θ, then T is a <u>sufficient</u> statistic for θ

Law of iterated expectation:

$$\mathbb{E}_Y[\mathbb{E}_X[X|Y]] = \mathbb{E}_X[X]$$

Ex where T is not sufficient;
$\hat{\tau}$ unbiased, $T=1$
$\mathbb{E}[\hat{\tau}|T] = \mathbb{E}[\hat{\tau}] = \tau$
don't know the parameter!

Law of iterated variance:

$$Var(X) = Var_Y(\mathbb{E}_X[X|Y]) + \mathbb{E}_Y[Var_X(X|Y)]$$

Pf:

① $$\mathbb{E}[\tau^*] = \mathbb{E}[\mathbb{E}[\hat{\tau}|T]] = \mathbb{E}[\hat{\tau}] = \tau$$
(unbiased) ✓

② $$\begin{aligned} Var(\tau^*) &= Var(\mathbb{E}[\hat{\tau}|T]) + \overbrace{\mathbb{E}[Var(\hat{\tau}|T)]}^{\geq 0} \\ &\leq Var(\mathbb{E}[\hat{\tau}|T]) + \mathbb{E}[Var(\hat{\tau}|T)] \\ &= Var(\hat{\tau}) \end{aligned}$$ ✓

Why do we need sufficiency? $\tau^* = \mathbb{E}[\hat{\tau}|T]$

We need τ^* to <u>not</u> be a function of θ and $\tau(\theta)$

If T is sufficient, $(X_1, \ldots, X_n) | T$ does not depend on θ $\Rightarrow$ $\hat{\tau}|T$ does not depend on θ (or on $\tau(\theta)$)

Let $X_1, \dots, X_n \overset{iid}{\sim} \text{Poisson}(\lambda)$

sufficient statistic:

$$T = \sum_i X_i$$

consider $\hat{\lambda} = X_1$ $\quad \mathbb{E}[\hat{\lambda}] = \lambda \quad Var(\hat{\lambda}) = \lambda$

$$\lambda^* = \mathbb{E}[\hat{\lambda} \mid T]$$

$$\mathbb{E}[X_1 \mid T] = \mathbb{E}[X_2 \mid T] = \dots$$

$$\underbrace{\sum_i \mathbb{E}[X_i \mid T]}_{\mathbb{E}[\sum_i X_i \mid T] \; = \; T} = n\,\mathbb{E}[X_1 \mid T]$$

$$\mathbb{E}\left[\sum_i X_i \mid T\right] = T$$

$$\Rightarrow \sum_i \mathbb{E}[X_i \mid T] = T$$

$$\Rightarrow n\,\mathbb{E}[X_1 \mid T] = T$$

$$\Rightarrow \mathbb{E}[X_1 \mid T] = \frac{T}{n}$$

$$\lambda^* = \frac{T}{n} = \frac{1}{n} \sum_i X_i$$

(achieves CRLB)