## Theorem: Cramer - Rao Lower Bound

If $\hat{\theta}$ is an unbiased estimator of $\tau(\theta)$ then

$$Var(\hat{\theta}) \geq \left(\frac{\partial \tau}{\partial \theta}\right)^2 \Big/ I_N(\theta)$$

(*) If $f_\theta$ is regular enough

---

Pf. ① $\Rightarrow$ $$Cor(A,B) = \frac{Cov(A,B)}{\sqrt{Var(A)Var(B)}}$$

$$Cov(A,B) = E[AB] - E[A]E[B]$$
$$= E[(A - E(A))(B - E(B))]$$

② $\Rightarrow$ $$-1 \leq Cor(A,B) \leq 1$$

then $$0 \leq Cor(A,B)^2 \leq 1$$

③ Combine ① and ②

$$Cor(A,B)^2 = \frac{Cov(A,B)^2}{Var(A)Var(B)} \leq 1$$

and rearrange

$$\text{Var}(A) \geq \frac{\text{Cov}(A,B)^2}{\text{Var}(B)}$$

to prove theorem, let $A = \hat{\theta}$, $B = S_\theta$

① we know $\text{Var}(S_\theta) = I_N(\theta)$

② need to show $\text{Cov}(\hat{\theta}, S_\theta) = \frac{\partial \tau}{\partial \theta}$

$$\text{Cov}(\hat{\theta}, S_\theta) = \mathbb{E}[\hat{\theta} S_\theta] - \cancel{\mathbb{E}[\hat{\theta}]\mathbb{E}[S_\theta]}^{\,0}$$

(*) $\mathbb{E}[S_\theta] = 0$

$$= \int \hat{\theta} S_\theta f_\theta(\underset{\sim}{x})\, d\underset{\sim}{x}$$

recall:

$$S_\theta = \frac{\partial}{\partial \theta} \log f_\theta(\underset{\sim}{x})$$

and

$$\frac{\partial}{\partial \theta} \log f_\theta = \frac{\frac{\partial}{\partial \theta} f_\theta(\underset{\sim}{x})}{f_\theta(\underset{\sim}{x})}$$

$$= \int \hat{\theta} \frac{\frac{\partial}{\partial \theta} f_\theta(\underset{\sim}{x})}{\cancel{f_\theta(\underset{\sim}{x})}} \cancel{f_\theta(\underset{\sim}{x})}\, d\underset{\sim}{x}$$

$$= \int \hat{\theta} \frac{\partial}{\partial \theta} f_\theta(\underset{\sim}{x})\, d\underset{\sim}{x}$$

$$= \frac{\partial}{\partial \theta} \underbrace{\int \hat{\theta} f_\theta(\underset{\sim}{x})\, d\underset{\sim}{x}}_{\mathbb{E}[\hat{\theta}] = \tau(\theta)} \quad (*)$$

$$= \frac{\partial}{\partial \theta} \tau(\theta) = \frac{\partial \tau}{\partial \theta}$$

$$\mathrm{Var}(A) \geq \frac{\mathrm{Cov}(A,B)^2}{\mathrm{Var}(B)}$$ $A = \hat{\theta}$ and $B = S_\theta$

Shown: $\mathrm{Cov}(\hat{\theta}, S_\theta) = \frac{\partial \tau}{\partial \theta}$

$\mathrm{Var}(S_\theta) = I_N(\theta)$

$$\mathrm{Var}(\hat{\theta}) = \left(\frac{\partial \tau}{\partial \theta}\right)^2 \Big/ I_N(\theta).$$

---

Said last time:

(1) If I have $\hat{\theta}$ unbiased for $\tau(\theta)$ and I can show that

$$\mathrm{Var}(\hat{\theta}) = \mathrm{CRLB} = \left(\frac{\partial \tau}{\partial \theta}\right)^2 \Big/ I_N(\theta)$$

then $\hat{\theta}$ is the UMVUE for $\tau(\theta)$.

(2) However, if all I know is that

$$\mathrm{Var}(\hat{\theta}) > \mathrm{CRLB}$$

I don't know if $\hat{A}$ is the UMVUE.

(Bound isn't always tight)
sometimes nothing achieves the CRLB

---

**Theorem:** If $X_n \overset{iid}{\sim} f_\theta$ and $f_\theta$ satisfies the CRLB theorem conditions.
And $\hat{\theta}$ is unbiased for $\tau(\theta)$ then $\hat{\theta}$ achieves the CRLB iff

$$S_\theta \propto_\theta \hat{\theta} - \tau(\theta)$$

proportional to

i.e. $$S_\theta = a(\theta)(\hat{\theta} - \tau(\theta))$$

some function of $\theta$ (not $\underset{\sim}{X}$)

hence $\hat{\theta}$ is the UMVUE for $\tau(\theta)$.

---

Ex. let $X_n \overset{iid}{\sim} \text{Bernoulli}(p)$

and let $\hat{p} = \bar{X}$ and $\tau(p) = p$

then $\hat{p}$ is unbiased for $\tau(p)$.

$$\left(E[\bar{X}] = p\right)$$

point dist.

$$S_\theta(\underset{\sim}{x}) = \frac{\partial}{\partial\theta} \log f_\theta(\underset{\sim}{x})$$

$$f_\theta(\underset{\sim}{x}) = \prod_{n=1}^{N} p^{x_n}(1-p)^{1-x_n}$$

$$\log f_\theta(\underset{\sim}{x}) = \sum_{n=1}^{N} \log\left(p^{x_n}(1-p)^{1-x_n}\right)$$

$$= \sum_{n=1}^{N}\left[x_n \log(p) + (1-x_n)\log(1-p)\right]$$

$$= \log(p)\sum_{n=1}^{N} x_n + \log(1-p)\underbrace{\sum_{n=1}^{N}(1-x_n)}_{N - \sum_{n=1}^{N} x_n}$$

$$\frac{\partial}{\partial p}\log f_\theta(\underset{\sim}{x}) = \frac{1}{p}\underbrace{\sum_{n=1}^{N} x_n}_{N\overline{X}} - \frac{1}{1-p}\left(N - \underbrace{\sum_{n=1}^{N} x_n}_{N\overline{X}}\right)$$

$$= \frac{N\overline{X}}{p} - \frac{N(1-\overline{X})}{1-p}$$

$$= \frac{(1-p)N\overline{X} - pN(1-\overline{X})}{p(1-p)}$$

$$= \frac{N\overline{X} - \cancel{pN\overline{X}} - pN + \cancel{pN\overline{X}}}{p(1-p)}$$

$$S_\theta = \frac{N}{p(1-p)}(\overline{X} - p)$$

$$a(p)\left(\hat{p} - \tau(p)\right)$$

So by our theorem $\hat{p}$ is the UMVUE for $p$.

---

Pf. From proof of CRLB

$$Var(\hat{\theta}) \geq \frac{Cov(\hat{\theta}, S_\theta)}{Var(S_\theta)}$$

$>$ or $=$

wait to prove when equal

Want:

$$Var(\hat{\theta}) = \frac{Cov(\hat{\theta}, S_\theta)}{Var(S_\theta)}$$

rewrite:

$$\underbrace{\frac{Cov(\hat{\theta}, S_\theta)}{Var(S_\theta) Var(\hat{\theta})}}_{Cor(\hat{\theta}, S_\theta)^2} = 1$$

$$Cor(\hat{\theta}, S_\theta)^2 = 1$$

So $Cor(\hat{\theta}, S_\theta) = \pm 1$

meaning $\hat{\theta}$ and $S_\theta$ are linear functions of each other. So

$$S_\theta = b + a\hat{\theta}$$

Recall: $\mathbb{E}[S_\theta] = 0$

So $\mathbb{E}[S_\theta] = \mathbb{E}[b + a\hat{\theta}] = 0$

$$\Rightarrow b + a\,\mathbb{E}[\hat{\theta}] = 0$$

$$\Rightarrow b + a\,\tau(\theta) = 0$$

$$\Rightarrow b = -a\,\tau(\theta)$$

$$S_\theta = -a\,\tau(\theta) + a\hat{\theta}$$

$$= a(\hat{\theta} - \tau(\theta)).$$

---

Ex. $X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ where $\sigma^2$ is known.

$\hat{\mu} = \bar{X}$ then $\mathbb{E}[\hat{\mu}] = \mu$.

---

$$f(\underline{x}) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_n - \mu)^2\right)$$

$$= (2\pi)^{-N/2}(\sigma^2)^{-N/2} \exp\left(-\frac{1}{2\sigma^2}\sum_{n=1}^{N}(x_n - \mu)^2\right)$$

$$\log f(\underline{x}) = -\frac{N}{2}\log(2\pi) - \frac{N}{2}\log\sigma^2 - \frac{1}{2\sigma^2}\sum_{n=1}^{N}(x_n - \mu)^2$$

$$\frac{\partial}{\partial\mu}\log f(\underset{\sim}{x}) = \frac{1}{\sigma^2}\sum_{n=1}^{N}(x_n-\mu)$$

hence

$$S_\mu = \frac{1}{\sigma^2}\sum_{n=1}^{N}(x_n-\mu) = \frac{1}{\sigma^2}N\bar{x} - N\mu$$

$$= \underbrace{\frac{N}{\sigma^2}}_{a(\mu)}(\bar{x}-\mu)$$

$$a(\mu)(\hat{\mu}-\mu)$$

So $\bar{X}$ is the UMVUE for $\mu$.

---

<u>Ex.</u> $X_n \overset{iid}{\sim} \mathrm{Exp}(\lambda)$. Let $\underbrace{\hat{\lambda} = \bar{X}}$ and $\tau(\lambda) = 1/\lambda$.

Then

$$f_\lambda(\underset{\sim}{x}) = \prod_{n=1}^{N}\lambda e^{-\lambda x_n} = \lambda^N e^{-\lambda\sum_{n=1}^{N}x_n}$$

$$\log f_x(\underset{\sim}{x}) = N\log\lambda - \lambda\underbrace{\sum_{n=1}^{N}x_n}_{N\bar{x}}$$

$$= N\log\lambda - \lambda N\bar{x}$$

$$\frac{\partial}{\partial x}\log f_x(\underset{\sim}{x}) = \frac{N}{\lambda} - N\bar{x}$$

$$S_\lambda = \frac{N}{\lambda} - N\bar{x} = -N(\bar{x} - 1/\lambda)$$

$a(x) \quad \hat{\lambda} \quad T(x)$

So $\bar{X}$ is the UMVUE for $1/\lambda$.

---

<u>Theorem:</u> Attainment for Exp. Fams.

If $X_n \overset{iid}{\sim} f_\theta$ and $f_\theta$ is an exp. fam so that

$$f_\theta(\underset{\sim}{x}) = c(\theta) h(\underset{\sim}{x}) \exp\left(w(\theta) T(\underset{\sim}{x})\right)$$

Let $\tau(\theta) = \mathbb{E}[T]$.
Then $T$ is the UMVUE for $\tau(\theta)$.

---

<u>Pf:</u> Note that

$$\log f_\theta(\underset{\sim}{x}) = \log c(\theta) + \log h(\underset{\sim}{x}) + w(\theta) T(\underset{\sim}{x})$$

$$\frac{\partial}{\partial\theta} \log f_\theta(\underset{\sim}{x}) = \frac{c'(\theta)}{c(\theta)} + w'(\theta) T(\underset{\sim}{x})$$

So $$S_\theta = \frac{c'(\theta)}{c(\theta)} + w'(\theta) T(\underset{\sim}{X})$$

$\mathbb{E} S_\theta = 0$

$c(\theta)$

$$= w'(\theta)\Big(T(\underline{X}) - \underbrace{\frac{-c'(\theta)}{c(\theta)w'(\theta)}}_{\tau(\theta)}\Big)$$

$E S_\theta = 0$
so
$E\left[T - \frac{-c'(\theta)}{c(\theta)w'(\theta)}\right] = 0$
hence
$E[T] = \frac{-c'(\theta)}{c(\theta)\,w'(\theta)}$

$$= \underbrace{w'(\theta)}_{a(\theta)}\big(T - \tau(\theta)\big)$$

---

DONE w/ CRLB

---

## Review of Conditional/Iterated Expectation

(Leemis' Probability Book Section 6.3)

If $(X, Y)$ have a joint dist $f(x,y)$ and marginals $f(x)$ and $f(y)$
and the conditional dist. of $X$ given $Y=y$ is

$$f(x|y) = \frac{f(x,y)}{f(y)}$$

We can think of $\underbrace{X|Y=y}_{\text{Univariate RV } Z}$ as a univariate RV.

$\to Z$ has a dist. $f(x|y)$

↳ $z$ has a dist. $f(x|y)$

→ $z$ has an expectation

$$E[z] = E[X|Y=y]$$

$$= \int x f(x|y)\, dx$$

Notice that if $y$ changes so does $X|Y=y$

and hence $E[X|Y=y]$

a number that changes depending on $y$ (i.e. a function)

Could call

$$g(y) = E[X|Y=y]$$

$$g: \mathbb{R} \to \mathbb{R}$$

---

Plug $Y$ into $g$, to get $g(Y)$

$$g(Y) = E[X|Y=Y]$$

a rv

notation is awkward

$$= E[X|Y]$$

a rv.

---

$\mathbb{E}[X|Y]$ is $\mathbb{E}[X|Y=y]$

but treat $y$ as random

e.g. if

$$\mathbb{E}[X|Y=y] = y^2$$

then $\mathbb{E}[X|Y] = Y^2$.

---

Punchline:

$\mathbb{E}[X|Y=y]$ a number

$\mathbb{E}[X|Y]$ a rv.

---

Theorem: Iterated Expectation

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

---