Lecture 8 : UMVUEs

Defn: Uniformly Minimum Variance Unbiased Estimator (UMVUE)

Note: If $B(\hat{\theta}) = 0$ then $MSE(\hat{\theta}) = Var\,\hat{\theta}$

We call θ^* the UMVUE of $\tau(\theta)$

some fn of θ

e.g. $\theta^2, \log\theta, e^\theta$

if
(1) unbiased for $\tau(\theta)$

$$\mathbb{E}[\theta^*] = \tau(\theta)$$

(2) minimum variance — uniformly

$$Var\,\theta^* \leq Var\,\hat{\theta} \quad \forall\, \theta \in \Theta$$

for all $\hat{\theta}$ that are unbiased for $\tau(\theta)$

Defn: Score Basically $\frac{\partial \ell}{\partial \theta}$ but viewed as random

recall: $\underset{\sim}{x}$ deterministic, $\underset{\sim}{X}$ is random

If $X_n \overset{iid}{\sim} f_\theta$ when $\theta \in \Theta$ then the score is

$$S = S_\theta = S_\theta(\underset{\sim}{X}) = \frac{\partial \log f_\theta(\underset{\sim}{X})}{\partial \theta} = \text{"}\frac{\partial \ell}{\partial \theta}\text{"}$$

$f_\theta(\underset{\sim}{x})$

Ex. $X_n \overset{iid}{\sim} \text{Exp}(\lambda)$ where $\lambda > 0$

then $L(\lambda) = \prod_{n=1}^{N} \lambda e^{-\lambda x_n} \mathbb{1}(x_n > 0)$

$= \lambda^N e^{-\lambda \sum_n x_n} \mathbb{1}(x_{(1)} > 0)$

$\ell(\lambda) = \log L(\lambda) = N \log \lambda - \lambda \sum_n x_n + \log \mathbb{1}(x_{(1)} > 0)$

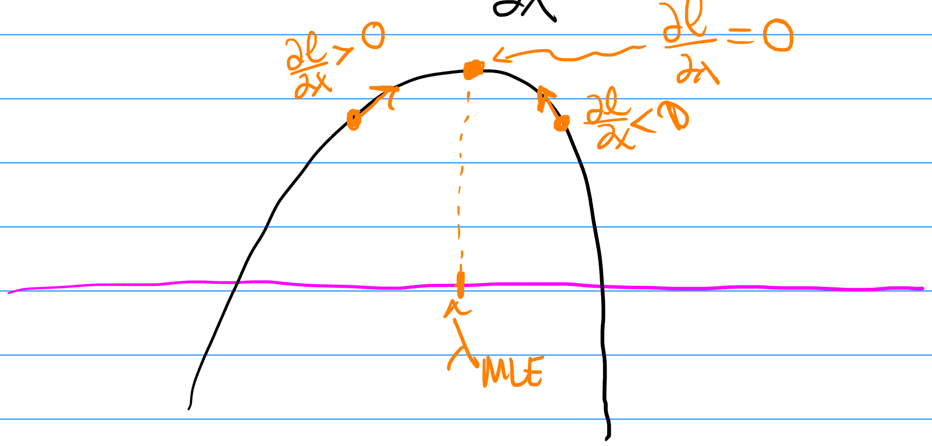
$\frac{\partial \ell}{\partial \lambda} = \frac{N}{\lambda} - \sum_n x_n$

Score promotes x_n to X_n

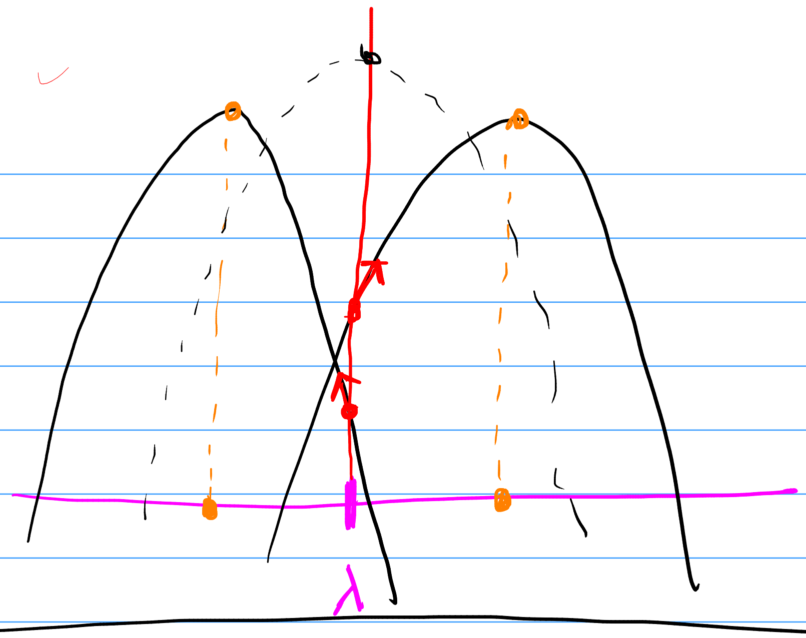
Sc $\boxed{S_\lambda = \frac{N}{\lambda} - \sum_n X_n}$ ← now random

I get $\hat{\lambda}_{MLE}$ by setting $\frac{\partial \ell}{\partial \lambda} = 0$

$\frac{\partial \ell}{\partial \lambda}$



Score ≈ random ver. of $\frac{\partial \ell}{\partial \lambda}$



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

Back to ex,

$$S_\lambda = \frac{N}{\lambda} - \sum_n X_n .$$

$$\mathbb{E}[S_\lambda] = \frac{N}{\lambda} - \sum_n \mathbb{E}[X_n] = \frac{N}{\lambda} - \sum_n \frac{1}{\lambda} = \frac{N}{\lambda} - \frac{N}{\lambda} = 0$$

My MLE is $\hat{\lambda} = 1/\bar{X}$

and $\mathbb{E}\hat{\lambda} = \mathbb{E}[1/\bar{X}] \neq \lambda$ ← MLE is not unbiased for λ

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

Pf:

lazy notation $\cdots \int \cdots \int \cdots dx_1 dx_2 \cdots$

$$\mathbb{E}[S_\theta(\underset{\sim}{x})] = \int S_\theta(\underset{\sim}{x}) f(\underset{\sim}{x}) d\underset{\sim}{x}$$

$$\mathbb{E}g(x) = \int g(x) f(x) dx$$

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

<u>Aside:</u> $\frac{\partial \ell}{\partial \theta} = \frac{\partial}{\partial \theta} \log f_\theta(\underset{\sim}{x})$

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

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

$$= \int \frac{\partial f}{\partial \theta} d\underset{\sim}{x}$$

Need enough "regularity" i.e. f_θ has to be nice enough

$$(*) = \frac{\partial}{\partial \theta} \int f(\underset{\sim}{x}) d\underset{\sim}{x}$$

This works for Exp. Fams.

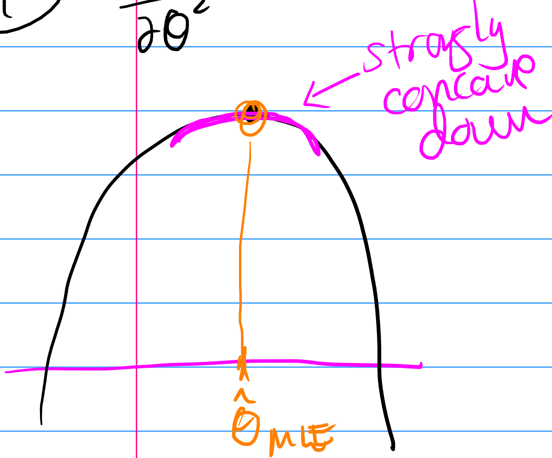
$$= \frac{\partial}{\partial \theta} (1)$$

$$= 0$$

What about $\frac{\partial^2 \ell}{\partial \theta^2}$?

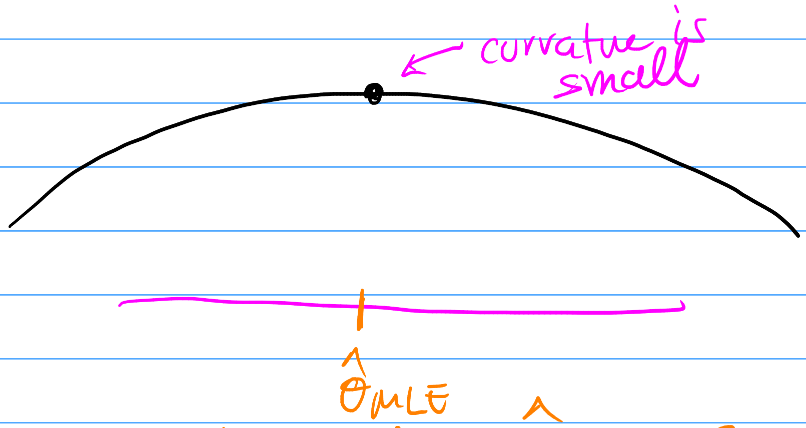
<u>Two possiblities</u>

(1) $\frac{\partial^2 \ell}{\partial \theta^2} \ll 0$



Strongly prefer $\hat{\theta}$ over other choices

(2) $\frac{\partial^2 \ell}{\partial \theta^2} < 0$ but close to zero



Weakly prefer $\hat{\theta}$ over other choices

Theorem: (*) also needs regularity of f_θ

$$\mathrm{Var}(S_\theta) = \mathbb{E}[S_\theta^2] = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial \theta^2}\right]$$

b/c $ES_\theta = 0$

viewing as random
promote $\underset{\sim}{x}$ to $\underset{\sim}{X}$

So if $S_\theta = \frac{\partial \ell}{\partial \theta}$ then

$$\mathbb{E}\left[\left(\frac{\partial \ell}{\partial \theta}\right)^2\right] = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial \theta^2}\right]$$

Defn: Fisher Information

We define the fisher info. for θ contained in X $(N=1)$ is defined as

$$I(\theta) = -\mathbb{E}\left[\frac{\partial^2}{\partial \theta^2} \log f_\theta(X)\right]$$

$\frac{\partial \ell}{\partial \theta}$ w/ $N=1$

If I have N samples $X_n \overset{iid}{\sim} f_\theta$ then the Fisher info. contained about θ in $\underset{\sim}{X}$ is

$$I_N(\theta) = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial \theta^2}\right].$$

Theorem: $I_N(\theta) = N\, I(\theta)$

If I have N indep. samples - I have N times as much info.

Pf. $I_N(\theta) = -\mathbb{E}\left[\frac{\partial^2 \ell}{\partial \theta^2}\right]$

joint

$$= -\mathbb{E}\left[\frac{\partial^2}{\partial \theta^2} \log f(\underline{x})\right]$$

$$= -\mathbb{E}\left[\frac{\partial^2}{\partial \theta^2} \log\left(\prod_n f_\theta(x_n)\right)\right]$$

$$= -\mathbb{E}\left[\frac{\partial^2}{\partial \theta^2} \sum_n \log f(x_n)\right]$$

$$= \sum_n \underbrace{-\mathbb{E}\left[\frac{\partial^2}{\partial \theta^2} \log f(x_n)\right]}_{I(\theta)}$$

$$= N\, I(\theta)$$

Ex. $X_n \overset{iid}{\sim} \text{Pois}(\lambda)$, $\lambda > 0$

Find $I_N(\lambda)$.

Let's find $I(\lambda)$ and multiply by N.

(1) Find $\log f_\lambda(x)$

$$f_\lambda(x) = \frac{\lambda^x e^{-\lambda}}{x!} \Rightarrow \log f_\lambda = x \log \lambda - \lambda - \log(x!)$$

(2) Take two derivs.

$$\to \frac{\partial}{\partial x} \log f_\lambda = \frac{x}{\lambda} - 1$$

$$\to \frac{\partial^2}{\partial x^2} \log f_\lambda = -\frac{x}{\lambda^2}$$

promote x to X .... to get $-\frac{X}{\lambda^2}$

need to calc

$$I(x) = -\mathbb{E}\left[\frac{\partial^2 \log f_\lambda}{\partial \lambda^2}\right] = -\mathbb{E}\left[-\frac{X}{\lambda^2}\right] = \frac{1}{\lambda^2}\mathbb{E}X = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

So $\boxed{I_N(x) = N/\lambda}$.

$\mathbb{E}\bar{X} = \lambda$ (unbiased)

$\mathrm{Var}\,\bar{X} = \frac{\lambda}{N}$

<u>Ex.</u> Let $X_n \overset{iid}{\sim} N(\mu, \sigma^2)$ Known

Let's get $I_N(\mu)$.

(1) $f_\mu(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x-\mu)^2\right)$

(2) $\log f_\mu = -\frac{1}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}(x-\mu)^2$

(3) $\frac{\partial}{\partial \mu}\log f_\mu = -\frac{1}{2\sigma^2}(2)(x-\mu)(-1) = \frac{1}{\sigma^2}(x-\mu) = \frac{x}{\sigma^2} - \frac{\mu}{\sigma^2}$

(4) $\frac{\partial^2}{\partial \mu^2} \lg f_\mu = -\frac{1}{\sigma^2}$

(5) Promote x to X .... to get $-\frac{1}{\sigma^2}$

(6) $-\mathbb{E}[\cdots\cdots] = -\mathbb{E}\left[-\frac{1}{\sigma^2}\right] = \frac{1}{\sigma^2}$

$I(\mu) = 1/\sigma^2$

So $\boxed{I_N(\mu) = N I(\mu) = N/\sigma^2}$

$\mathbb{E}\bar{X} = \mu$ (unbiased)

$\mathrm{Var}\,\bar{X} = \sigma^2/N$

Ex. $X_n \overset{iid}{\sim} \mathrm{Pois}(\lambda)$, $\mathbb{E}X_n = \mathrm{Var}\,X_n = \lambda$

$sd(X_n) = \sqrt{\lambda} = \psi$

i.e. $\psi^2 = \lambda$

To get $I_N(\psi)$ I can re-parameterize Pois in terms of ψ and apply my procedure

$f_\lambda(x) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{(\psi^2)^x e^{-\psi^2}}{x!} = f_\psi(x)$
