Want to find "best" estimators.

Fact: In general if I'm too permissive in what I allow to be an estimator — there is no "best" estimator.

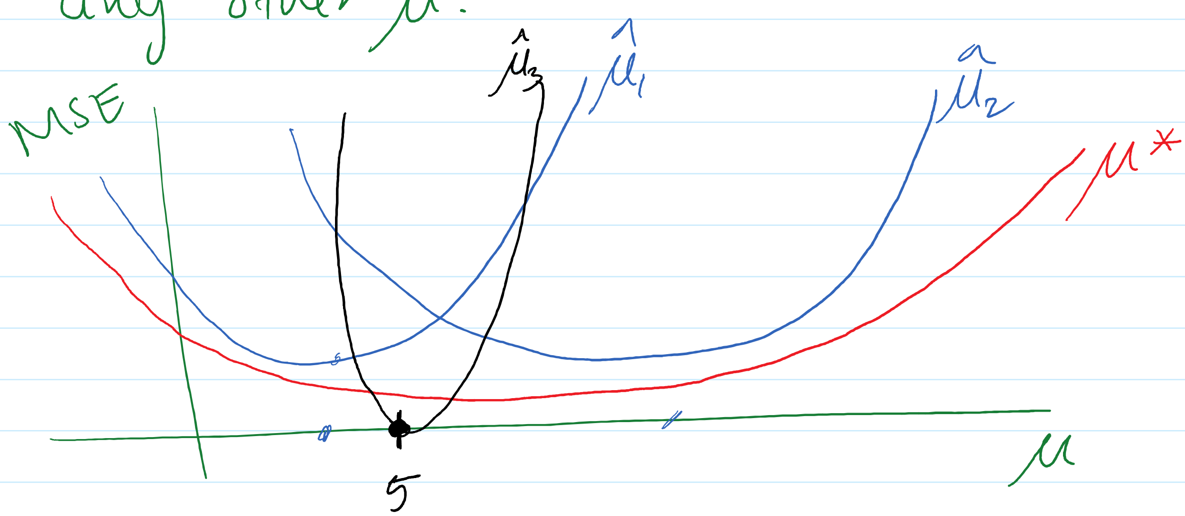
Ex. $X_n \overset{iid}{\sim} N(\mu, 1)$

Want to estimate μ.

i.e. want some μ^* so that

$$MSE_\mu(\mu^*) \le MSE_\mu(\hat{\mu}) \quad \forall \mu$$

for any other $\hat{\mu}$.



Let $\hat{\mu}_3 = 5$

$$\underset{\mu=5}{MSE}(\hat{\mu}_3) = E[(\hat{\mu}_3 - 5)^2] = 0$$

Need to restrict class of allowable estimators otherwise there is no "best" estimator.

One way: restrict "allowable" estimators to unbiased ests.

Defn: Uniformly Minimum-Variance Unbiased Estimator (UMVUE)

Note: $B_\theta(\hat{\theta}) = 0$ then $MSE_\theta(\hat{\theta}) = Var(\hat{\theta})$

We call $\hat{\theta}^*$ the UMVUE of $\tau(\theta)$ if

some fn of θ, e.g. $\theta^2, \log\theta, e^{-\theta}, \ldots$

① unbiased

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

② minimum variance

$$Var(\hat{\theta}^*) \le Var(\hat{\theta}) \quad \forall \theta \in \Theta$$

and all unbiased estimators $\hat{\theta}$ of $\tau(\theta)$

Defn: Score:

Basically $\frac{\partial \ell}{\partial \theta}$ but viewed as random.

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

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

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

random

Ex. $X_n \overset{iid}{\sim} \mathrm{Exp}(\lambda)$

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

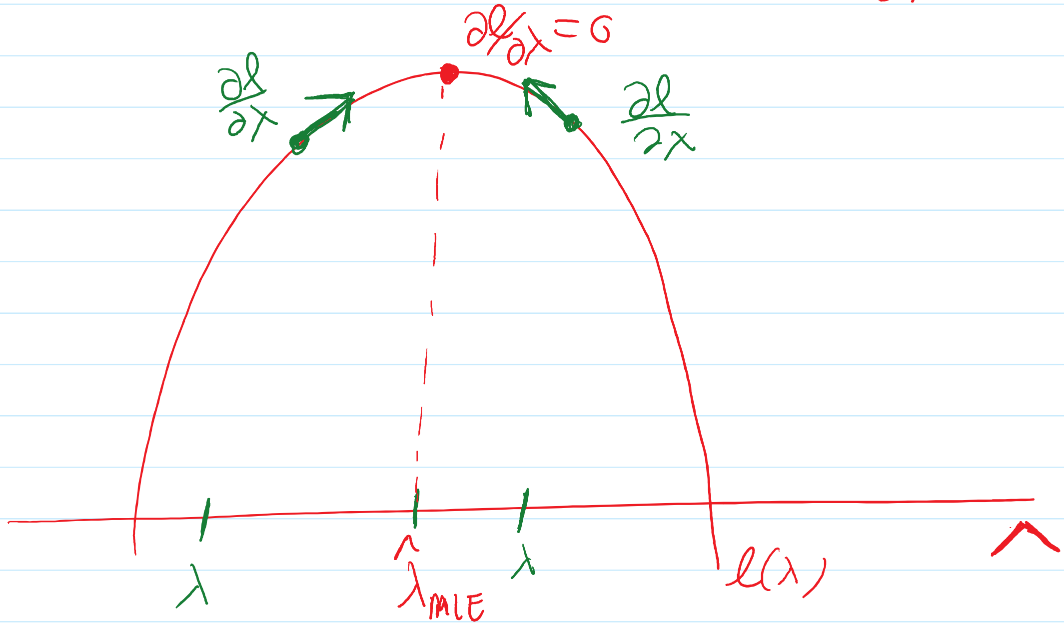
$$= \lambda^N e^{-\lambda \sum_n x_n} \mathbb{1}(x_{(1)} > 0)$$

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

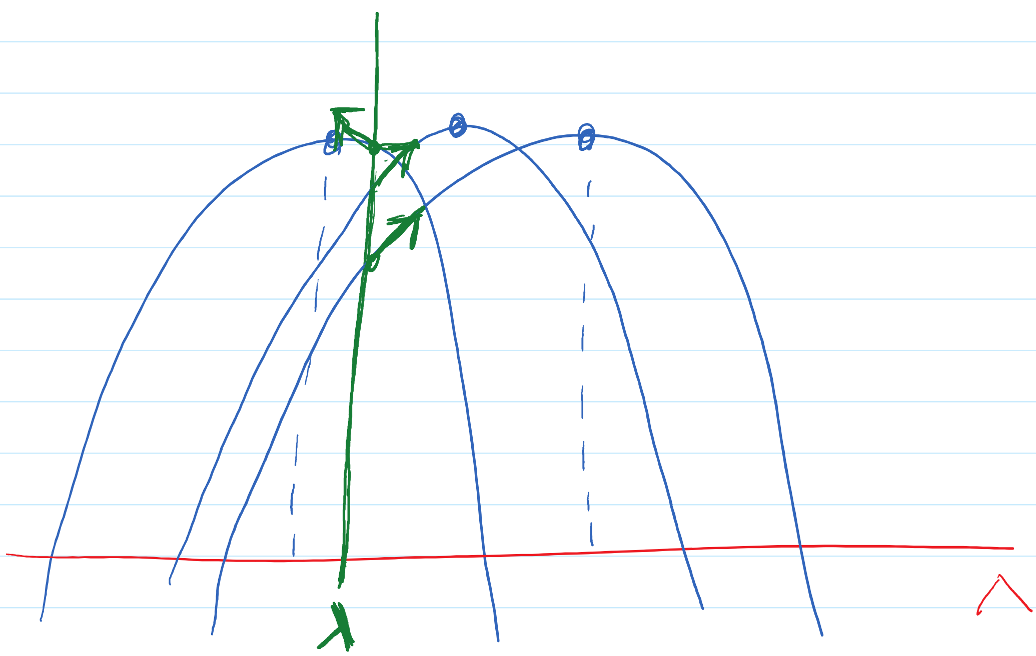
$$\frac{\partial \ell}{\partial \lambda} = \frac{N}{\lambda} - \sum_n x_n \longleftarrow \text{deriv of log-lik.}$$

Score: $$S_\lambda = \frac{N}{\lambda} - \sum_n X_n \longleftarrow \text{now random}$$

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



Score:



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

In prev. example:

$$\mathbb{E}[S_\lambda] = 0.$$

$$\mathbb{E}\Big[\frac{N}{\lambda} - \overbrace{\sum_n X_n}^{N\bar{X}}\Big] = 0$$

true: $\frac{N}{\lambda} - \sum_n \underbrace{\mathbb{E}[X_n]}_{1/\lambda} = 0$

My MLE: $\frac{N}{\lambda} - \sum X_n = 0$

and get $\hat{\lambda} = \frac{1}{\bar{X}}$

$$\mathbb{E}\Big[\frac{1}{\bar{X}}\Big] \neq \lambda$$

---

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

lazy notation: $\iiint \cdots dx_1 dx_2 \cdots$

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

$$\mathbb{E}[g(X)] = \int g(x) f(x)\, dx$$

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

Aside:

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

$$= \frac{\frac{\partial f}{\partial \theta}}{\ldots}$$

$\frac{d}{dx} \log u(x)$

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

$$= \frac{\frac{\partial f}{\partial \theta}}{f_\theta}$$

$$\frac{d}{dx} \log u(x) = \frac{u'(x)}{u(x)}$$

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

(*) Need enough regularity in f_θ — i.e. f_θ has to be nice enough

$$= \frac{\partial}{\partial \theta} \underbrace{\int f_\theta(\underset{\sim}{x})\, d\underset{\sim}{x}}_{1}$$

This works for Exp. families

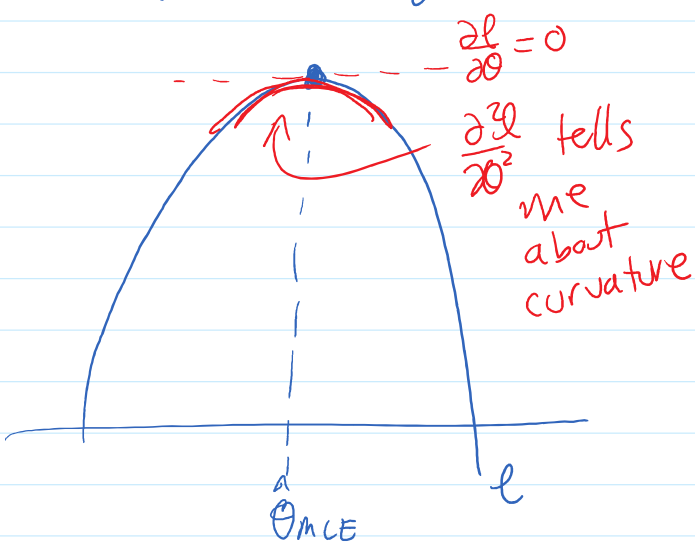
$$= \frac{\partial}{\partial \theta}(1) = 0$$

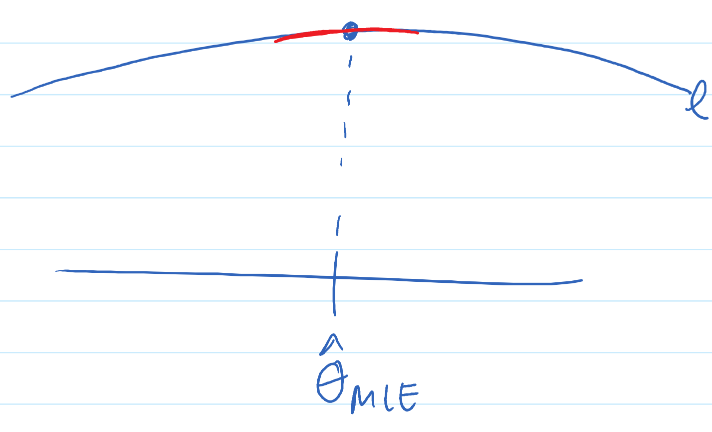
What about $\frac{\partial^2 \ell}{\partial \theta^2}$?

Two possibilites:

large (neg) $\frac{\partial^2 \ell}{\partial \theta^2}$

small $\frac{\partial^2 \ell}{\partial \theta^2}$





Theorem: (*) also needs regularity of f_θ

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

thinking of as random

If I think of S_θ as $\frac{\partial \ell}{\partial \theta}$ 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 $(N=1)$

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

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

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

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

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

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

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

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

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

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

Find $I_N(\lambda)$.

① Find $\log f_\lambda(x)$

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

② Find $\frac{\partial^2}{\partial \lambda^2} \log f_\lambda(x)$

$$\frac{\partial}{\partial \lambda} \rightsquigarrow \frac{x}{\lambda} - 1$$

$$\frac{\partial^2}{\partial \lambda^2} \rightsquigarrow -\frac{x}{\lambda^2}$$

(3) Form $-\mathbb{E}\left[\frac{\partial^2 \log f_X}{\partial \lambda^2}\right] = I(\lambda)$

$$-\mathbb{E}\left[-\frac{X}{\lambda^2}\right] = \frac{\mathbb{E}X}{\lambda^2} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

(4) $I_N(\lambda) = N\,I(\lambda)$

$$I_N(\lambda) = N/\lambda.$$