6.4-8. Let $X_1, X_2, \ldots, X_n$ be a random sample from the distribution whose pdf is $f(x;\theta) = (1/\theta)x^{(1-\theta)/\theta}$, $0<x<1$, $0<\theta<\infty$.

(a) Show that the maximum likelihood estimator of θ is $\widehat{\theta} = -(1/n)\sum_{i=1}^{n} \ln X_i$.

(b) Show that $E(\widehat{\theta}) = \theta$ and thus that $\widehat{\theta}$ is an unbiased estimator of θ.

$$x_1^{\frac{1-\theta}{\theta}} \cdot x_2^{\frac{1-\theta}{\theta}} \cdots x_n^{\frac{1-\theta}{\theta}} = (x_1 . x_2 \ldots x_n)^{\frac{1-\theta}{\theta}}$$

SOL:

a) Define the likelihood function

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{1}{\theta} x_i^{\frac{1-\theta}{\theta}} = \left(\frac{1}{\theta}\right)^n \cdot \left(\prod_{i=1}^{n} x_i\right)^{\frac{1-\theta}{\theta}} = \theta^{-n} \cdot \left(\prod_{i=1}^{n} x_i\right)^{\frac{1-\theta}{\theta}}$$

Apply natural logarithm to $L(\theta)$ to get

$$\ln(L(\theta)) = -n\ln\theta + \frac{1-\theta}{\theta}.\ln\left(\prod_{i=1}^{n} x_i\right) = -n\ln\theta + \left(\frac{1}{\theta}-1\right).\ln\left(\prod_{i=1}^{n} x_i\right)$$

The 1^{st} derivative of $\ln(L(\theta))$ is:

$$\frac{d}{d\theta}\ln(L(\theta)) = -\frac{n}{\theta} - \frac{1}{\theta^2}\ln\left(\prod_{i=1}^{n} x_i\right)$$

The 1st order condition is then:

$$\frac{d}{d\theta} \ln(L(\theta)) = 0 \iff -\frac{n}{\theta} - \frac{1}{\theta^2} \ln\left(\prod_{i=1}^{n} x_i\right) = 0$$

Add $\frac{n}{\theta}$ to both sides

$$\iff \frac{n}{\theta} = -\frac{1}{\theta^2} \ln\left(\prod_{i=1}^{n} x_i\right)$$

multiply both sides by $\frac{\theta^2}{n}$

$$\iff \theta = -\frac{1}{n} \ln\left(\prod_{i=1}^{n} x_i\right)$$

$\ln(x \cdot y) = \ln x + \ln y$

$$\iff \boxed{\theta = -\frac{1}{n} \sum_{i=1}^{n} \ln x_i}$$

We determined the critical point of $\ln(L(\theta))$ [which the same as that of $L(\theta)$].

To check that the critical point is a maximum, use the 2nd derivative:

$$\frac{d^2}{d\theta^2}\ln(L(\theta)) = \frac{n}{\theta^2} + \frac{2}{\theta^3}\underbrace{\sum_{i=1}^{n}\ln x_i}_{\ln\left(\prod_{i=1}^{n} x_i\right)}$$

Evaluate this 2nd derivative at the critical point to get

$$\left.\frac{d^2}{d\theta^2}\ln(L(\theta))\right|_{\theta=-\frac{1}{n}\sum_{i=1}^{n}\ln x_i} = \frac{n}{\left(-\frac{1}{n}\sum\ln x_i\right)^2} + \frac{2}{\left(-\frac{1}{n}\sum\ln x_i\right)^3}\sum_{i=1}^{n}\ln x_i$$

$$= \frac{n^3}{\left(\sum\ln x_i\right)^2} - \frac{2n^3}{\left(\sum\ln x_i\right)^2} = -\frac{n^3}{\left(\sum\ln x_i\right)^2} < 0$$

$\Rightarrow$ $\hat{\theta} = -\frac{1}{n}\sum_{i=1}^{n}\ln x_i$ is the MLE of θ

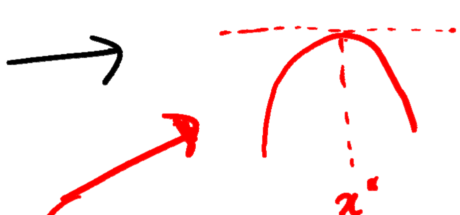
2nd derivative test

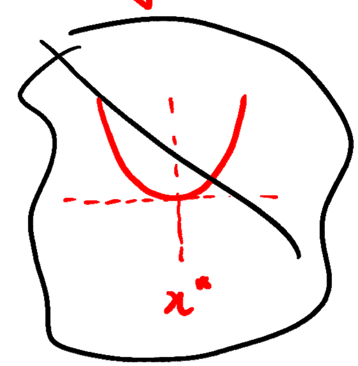
differentiable on I

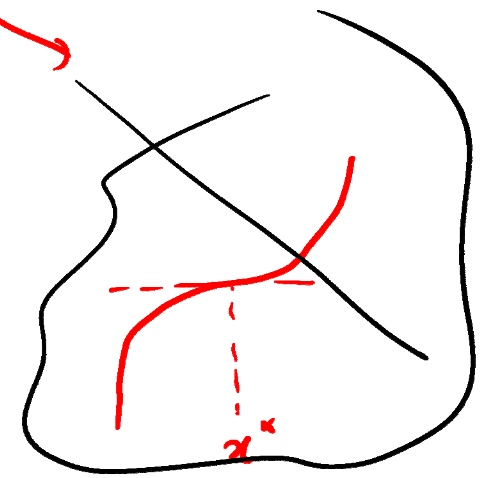
$f : I \to \mathbb{R}$, $I \subseteq \mathbb{R}$ interval

critical pt x^* $\boxed{f'(x^*) = 0}$ ← At the critical pt the tangent line is horizontal

3 possibilities







2nd derivative test : if f has a single critical pt x^* and $f''(x^*) < 0$

$\Rightarrow$ f has a maximum at x^*

6.4-5. Let $X_1, X_2, \ldots, X_n$ be a random sample from distributions with the given probability density functions. In each case, find the maximum likelihood estimator $\widehat{\theta}$.

(a) $f(x;\theta) = (1/\theta^2)\,x\,e^{-x/\theta}, \quad 0 < x < \infty, \quad 0 < \theta < \infty.$

(b) $f(x;\theta) = (1/2\theta^3)\,x^2\,e^{-x/\theta}, \quad 0 < x < \infty, \quad 0 < \theta < \infty.$

Key properties of ln:

$\ln(x \cdot y) = \ln x + \ln y \quad , \; x, y > 0$

$\ln(x^p) = p \ln x \quad , \; x > 0, \; p \in \mathbb{R}$

b) Define the likelihood function:

$$L(\theta) = \prod_{i=1}^{n} f(x_i, \theta) = \prod_{i=1}^{n} \underbrace{\frac{1}{2\theta^3}} \cdot x_i^2 \cdot e^{-x_i/\theta} = \underbrace{\left(\frac{1}{2\theta^3}\right)^n}_{(2\theta^3)^{-n}} \cdot \left(\prod_{i=1}^{n} x_i\right)^2 e^{-\frac{1}{\theta}\sum_{i=1}^{n} x_i}$$

Apply natural logarithm to get

$$\ln(L(\theta)) = -n \underbrace{\ln(2\theta^3)} + 2\ln\left(\prod_{i=1}^{n} x_i\right) - \frac{1}{\theta}\sum_{i=1}^{n} x_i$$

$$= -n \ln 2 - 3n \ln \theta + 2 \ln\left(\prod_{i=1}^{n} x_i\right) - \underbrace{\frac{1}{\theta}\sum_{i=1}^{n} x_i}$$

Compute the 1^{st} derivative of $\ln(L(\theta))$:

$$\frac{d}{d\theta} \ln(L(\theta)) = -\frac{3n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i$$

The 1^{st} order condition is then:

$$\frac{d}{d\theta} \ln(L(\theta)) = 0 \iff -\frac{3n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i = 0$$

Add $\frac{3n}{\theta}$ to both sides

$$\iff \frac{3n}{\theta} = \frac{1}{\theta^2} \sum_{i=1}^{n} x_i$$

Multiply both sides by $\frac{\theta^2}{3n}$

$$\iff \theta = \frac{1}{3n} \sum_{i=1}^{n} x_i$$

is the critical pt of $\ln(L(\theta))$ [and $L(\theta)$ too]

To check that the critical pt is indeed a maximum, check 2nd derivative:

$$\frac{d^2}{d\theta^2} \ln(L(\theta)) = \frac{3m}{\theta^2} - \frac{2}{\theta^3} \sum_{i=1}^{m} x_i$$

Evaluate at the critical pt to get:

$$\frac{d^2}{d\theta^2} \ln(L(\theta)) \Big|_{\theta = \frac{1}{3m}\sum_{i=1}^{m} x_i} = \frac{3m}{\left(\frac{1}{3m}\sum_{i=1}^{m} x_i\right)^2} - \frac{2}{\left(\frac{1}{3m}\sum_{i=1}^{m} x_i\right)^3} \cdot \sum_{i=1}^{m} x_i$$

$$= \frac{(3m)^3}{\left(\sum_{i=1}^{m} x_i\right)^2} - 2 \cdot \frac{(3m)^3}{\left(\sum_{i=1}^{m} x_i\right)^2} = -\frac{(3m)^3}{\left(\sum_{i=1}^{m} x_i\right)^2} < 0$$

Hence, $\hat{\theta} = \frac{1}{3m}\sum_{i=1}^{m} x_i$ is MLE of θ

6.5-3. The midterm and final exam scores of ten students in a statistics course are tabulated as shown.

(a) Calculate the least squares regression line for these data.

(b) Plot the points and the least squares regression line on the same graph.

(c) Find the value of $\widehat{\sigma^2}$.

Midterm (x_i)	Final (y_i)	Midterm	Final
70	87 (y_1)	67	73
74	79 (y_2)	70	83
80	88 (y_3)	64	79
84	98	74	91
80	96	82	94 (y_{10})

$n = 10$ data points

$(x_1, y_1), \ldots, (x_{10}, y_{10})$

a) $$\boxed{y = \hat{\alpha} + \hat{\beta}(x - \bar{x})}$$

where :

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$= \frac{1}{10}(70 + 74 + 80 + 84 + 80 + \cdots + 82)$$

$$\hat{\alpha} = \frac{1}{n} \sum y_i$$

$$= \frac{1}{10}(87 + 79 + 88 + \cdots + 94)$$

x_i	y_i	$x_i y_i$	x_i^2	y_i^2
$\sum x_i$	$\sum y_i$	$\sum x_i y_i$	$\sum x_i^2$	$\sum y_i^2$

given

$$\hat{\beta} = \frac{\sum x_i y_i - \frac{1}{n}\left(\sum x_i\right)\left(\sum y_i\right)}{\sum x_i^2 - \frac{1}{n}\left(\sum x_i\right)^2}$$

c) Use 2nd formula from formula sheet

$$\hat{\sigma}^2 = \frac{1}{n}\sum y_i^2 - \frac{1}{n^2}\left(\sum y_i\right)^2 - \frac{\hat{\beta}}{n}\sum x_i y_i + \frac{\hat{\beta}}{n^2}\sum x_i \sum y_i$$

If we were not given the sums in the table (as in this exercise!)

x_i	y_i	$\hat{y}_i$	$y_i - \hat{y}_i$	$(y_i - \hat{y}_i)^2$

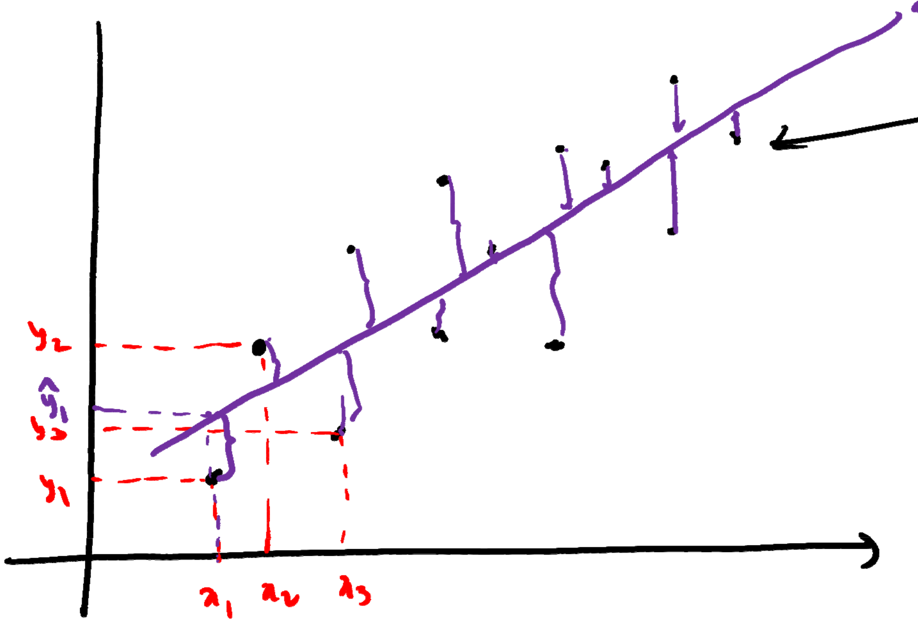
values are given

$$y = \hat{\alpha} + \hat{\beta}(x - \bar{x})$$ ← regression line

compute the estimates for y_i

$$\hat{y}_i = \hat{\alpha} + \hat{\beta}(x_i - \bar{x})$$

$$\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$$



(x_i, y_i)

$$y = \hat{\alpha} + \hat{\beta}(x - \bar{x})$$

$$\hat{y}_1 = \hat{\alpha} + \hat{\beta}(x_i - \bar{x})$$

$(y_1 - \hat{y}_1)^2$: square of distance from actual value y_1 to the estimate $\hat{y}_1$ given by regression line

$\widehat{\sigma^2}$ = average of distances squared

$$\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

6.7-5. Let $X_1, X_2, \ldots, X_n$ be a random sample from a gamma distribution with $\alpha = 1$ and $1/\theta > 0$. Show that $Y = \sum_{i=1}^{n} X_i$ is a sufficient statistic, Y has a gamma distribution with parameters n and $1/\theta$, and $(n-1)/Y$ is an unbiased estimator of θ.

Third part: Show that $\hat{\theta} = \dfrac{n-1}{Y}$ is an unbiased estimator of θ.

that is $\boxed{E[\hat{\theta}] = \theta}$

from next slide

But
$$E[\hat{\theta}] = E\left[\frac{n-1}{Y}\right] = (n-1) \cdot \underbrace{E\left[\frac{1}{Y}\right]}_{???\ \to\ \frac{\theta}{n-1}} = (n-1) \cdot \frac{\theta}{n-1} = \theta$$

Need to compute $E\left[\frac{1}{Y}\right]$ where Y is gamma with parameters n and $1/\theta$

$$E\left[\frac{1}{Y}\right] = \int_{-\infty}^{\infty} \frac{1}{y} \cdot f(y)\, dy = $$

$\uparrow$ $\underbrace{f(y) \text{ is the pdf of } Y}$

Recall:

if X has pdf $f(x)$ then

$$E[u(X)] = \int_{-\infty}^{\infty} u(x) \cdot f(x)\, dx$$

from formula sheet $\rightarrow$ $f(y) = \dfrac{y^{m-1}\, \theta^{m} \cdot e^{-\theta y}}{\Gamma(m)}, \quad y \geq 0$

gamma with parameter m and $1/\theta$

$\Gamma(m) = (m-1)\Gamma(m-1)$

$$= \int_0^{\infty} \frac{1}{y} \cdot \frac{y^{m-1}\, \theta^{m}\, e^{-\theta y}}{\Gamma(m)}\, dy = \int_0^{\infty} \frac{y^{m-2}\, \theta^{m}\, e^{-\theta y}}{\Gamma(m)}\, dy$$

$$= \int_0^{\infty} \frac{y^{m-2}\, \theta^{m-1} \cdot \theta \cdot e^{-\theta y}}{(m-1)\Gamma(m-1)}\, dy = \frac{\theta}{m-1} \int_0^{\infty} \frac{y^{m-2}\, \theta^{m-1}\, e^{-\theta y}}{\Gamma(m-1)}\, dy = \frac{\theta}{m-1} \cdot 1 = \frac{\theta}{m-1}$$

$\leftarrow$ pdf of gamma with $m-1$ and $1/\theta$

Cramér - Rao inequality :

if Y is an unbiased estimator of θ, then :

$E[Y] = \theta \leftarrow$ Y is unbiased

Var (Y) tells how much Y spreads about θ

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

if we happen to find an estimator with variance equal to $\frac{1}{I(\theta)}$, then

we have found a minimum - variance unbiased estimator

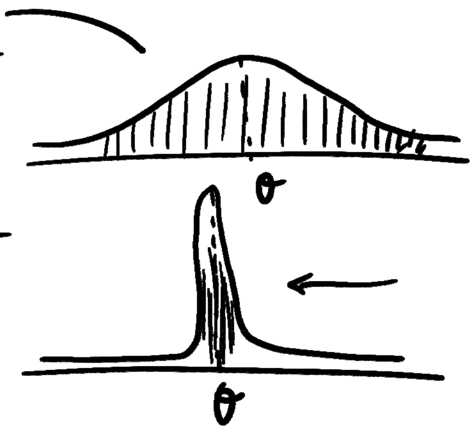
where $I(\theta)$ is the Fisher information

$$I(\theta) = \underbrace{n\, E\left[\left(\frac{\partial}{\partial\theta} \ln f(X,\theta)\right)^2\right]}_{\text{measures the amount of information that } f(x,\theta) \text{ contains about } \theta} = -n\, E\left[\frac{\partial^2}{\partial\theta^2} \ln f(X,\theta)\right]$$

Suppose $X \sim N(\underset{\text{unknown}}{\theta}, \underset{\text{known}}{\sigma^2})$

large σ^2

small σ^2



$$I(\theta) = n E\left[\left(\frac{\partial}{\partial \theta} \ln f(X,\theta)\right)^2\right]$$

$$f(x,\theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\theta)^2}{2\sigma^2}} \Rightarrow \ln f(x,\theta) = -\frac{1}{2}\ln(2\pi\sigma^2) - \frac{(x-\theta)^2}{2\sigma^2}$$

$$\Rightarrow \frac{\partial}{\partial \theta} \ln f(x,\theta) = \frac{(x-\theta)}{\sigma^2}$$

$$I(\theta) = n E\left[\left(\frac{X-\theta}{\sigma^2}\right)^2\right] = \frac{n}{\sigma^4} E\left[(X-\underset{E[X]}{\theta})^2\right] = \frac{n\sigma^2}{\sigma^4} = \frac{n}{\sigma^2}$$

Factorization theorem (to identify sufficient statistics).

Let $X_1, X_2, \dots, X_n$ be a random sample from a distr. with pmf/pdf $f(x, \theta)$:

The joint pdf of $X_1, \dots, X_n$ is then:

$$f_{joint}(x_1, x_2, \dots, x_n, \theta) = \prod_{i=1}^{n} f(x_i; \theta)$$

$Y = u(X_1, X_2, \dots, X_n)$ is a sufficient statistic for θ iff

$$f_{joint}(x_1, \dots, x_n, \theta) = \underbrace{\phi(y, \theta)}_{\phi \text{ depends on } x_1, \dots, x_n \text{ only through } y.} \cdot \underbrace{h(x_1, \dots, x_n)}_{\text{does not depend on } \theta}$$

EXAMPLE: $X_1, \dots, X_n$ random sample from

$$f(x,\theta) = \frac{1}{\theta} e^{-x/\theta}, \quad x > 0, \ \theta > 0$$

then:

$$f_{joint}(x,\theta) = \prod_{i=1}^{n} f(x_i,\theta) = \prod_{i=1}^{n} \frac{1}{\theta} e^{-x_i/\theta}$$

$$= \underbrace{\left(\frac{1}{\theta}\right)^n e^{-\frac{1}{n}\left(\sum x_i\right)}}_{\phi(y,\theta)} \cdot \underbrace{1}_{h(x_1,\dots,x_n)}$$

$y = \sum x_i$ or $y = \frac{1}{n}\sum x_i$

if Y is a sufficient statistic and v is an invertible function then $W = v(Y)$ is also sufficient

$X_1, X_2, \ldots, X_n \sim \text{Uniform}(0, \theta)$

↳ pdf $f(x, \theta) = \frac{1}{\theta}$, $0 < x < \theta$ [redacted]

MLE :

$\hat{\theta} = \boxed{\max\{X_1, \ldots, X_n\} = Y_n}$ 6.4

Y_n is unbiased 6.4

may show up on Questions 1 and 2

But not on the questions regarding sufficient statistics!