$X_1, \dots, X_n \overset{ind}{\sim} f(x;\theta)$

$\uparrow$ parameter $\quad \theta \in \Theta.$

STAT 713 sp 2023 Lec 04 slides

Moment estimators, Maximum likelihood estimators

Karl B. Gregory

University of South Carolina

These slides are an instructional aid; their sole purpose is to display, during the lecture, definitions, plots, results, etc. which take too much time to write by hand on the blackboard. They are not intended to explain or expound on any material.

Moment estimator

Let $X_1, \dots, X_n$ be iid with a distribution characterized by θ.

1. The sample kth moment $\hat{m}_k = n^{-1}\sum_{i=1}^n X_i^k$ is an estimator of $m_k = \mathbb{E}X_1^k$.
2. If $\theta = g(m_1, \dots, m_k)$ then a *moment estimator* of θ is $\hat{\theta} = g(\hat{m}_1, \dots, \hat{m}_k)$.

The *method of moments* (MoMs) prescribes finding g by solving the system

$$\hat{m}_j = m_j(\theta), \quad j = 1, \dots, \text{as many as you need.}$$

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Normal}(\mu, \sigma^2)$.

1. Find the MoMs estimators of μ and σ^2.
2. Discuss whether better estimators might exist.

Have two unknown parameters, so try putting

$$\frac{1}{n}\sum_{i=1}^n x_i = \hat{m}_1 = \boxed{m_1 = \mu}$$
$$\frac{1}{n}\sum_{i=1}^n x_i^2 \quad \hat{m}_2 = \boxed{m_2 = \sigma^2 + \mu^2}$$

$\Big(\mathrm{Var}\, X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = m_2 - m_1^2$
$\sigma^2 = m_2 - \mu^2$

$$\boxed{\begin{aligned} \mu &= m_1 \\ \sigma^2 &= m_2 - m_1^2 \end{aligned}}$$

So the M.o.M. estimators of μ and σ^2 are

(Just plug in sample moments).

$$\hat{\mu} = \hat{m}_1 = \bar{x}_n$$
$$\hat{\sigma}^2 = \hat{m}_2 - \hat{m}_1^2 = \frac{1}{n}\sum_{i=1}^n x_i^2 - \left(\frac{1}{n}\sum_{i=1}^n x_i\right)^2$$
$$\vdots$$
$$= \frac{1}{n}\sum_{i=1}^n (x_i - \bar{x}_n)^2$$

$$\mathbb{E}X_1 = \alpha\beta$$
$$\text{Var}\, X_1 = \alpha\beta^2$$
$$\mathbb{E}X_1^2 = \alpha\beta^2 + (\alpha\beta)^2$$
$$= \beta^2\alpha(1+\alpha)$$

Exercise: Let $X_1, \ldots, X_n \overset{\text{ind}}{\sim} \text{Gamma}(\alpha, \beta)$.

1. Find the MoMs estimators of α and β.
2. Discuss whether better estimators might exist.
3. Compute MoMs estimators on birth data set from Davison (2003)

①

$$m_1 = \alpha\beta$$
$$m_2 = \beta^2\alpha(1+\alpha)$$

$\Leftrightarrow$

$$\alpha = \frac{m_1}{\beta}$$

$$m_2 = \beta^2 \frac{m_1}{\beta}\left(1 + \frac{m_1}{\beta}\right)$$
$$= \beta m_1 \left(1 + \frac{m_1}{\beta}\right)$$
$$\Rightarrow \quad = \beta m_1 + m_1^2$$

$$\boxed{\beta = \frac{m_2 - m_1^2}{m_1}}$$

$\alpha =$

$$m_1 = \alpha\left(\frac{m_2 - m_1^2}{m_1}\right)$$

$$\boxed{\alpha = \frac{m_1^2}{m_2 - m_1^2}}$$

The MoM estimators for α, β are

$$\hat{\alpha} = \frac{\hat{m}_1^2}{\hat{m}_2 - \hat{m}_1^2} \qquad \hat{\beta} = \frac{\hat{m}_2 - \hat{m}_1^2}{\hat{m}_1}.$$

$\hat{m}_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2$

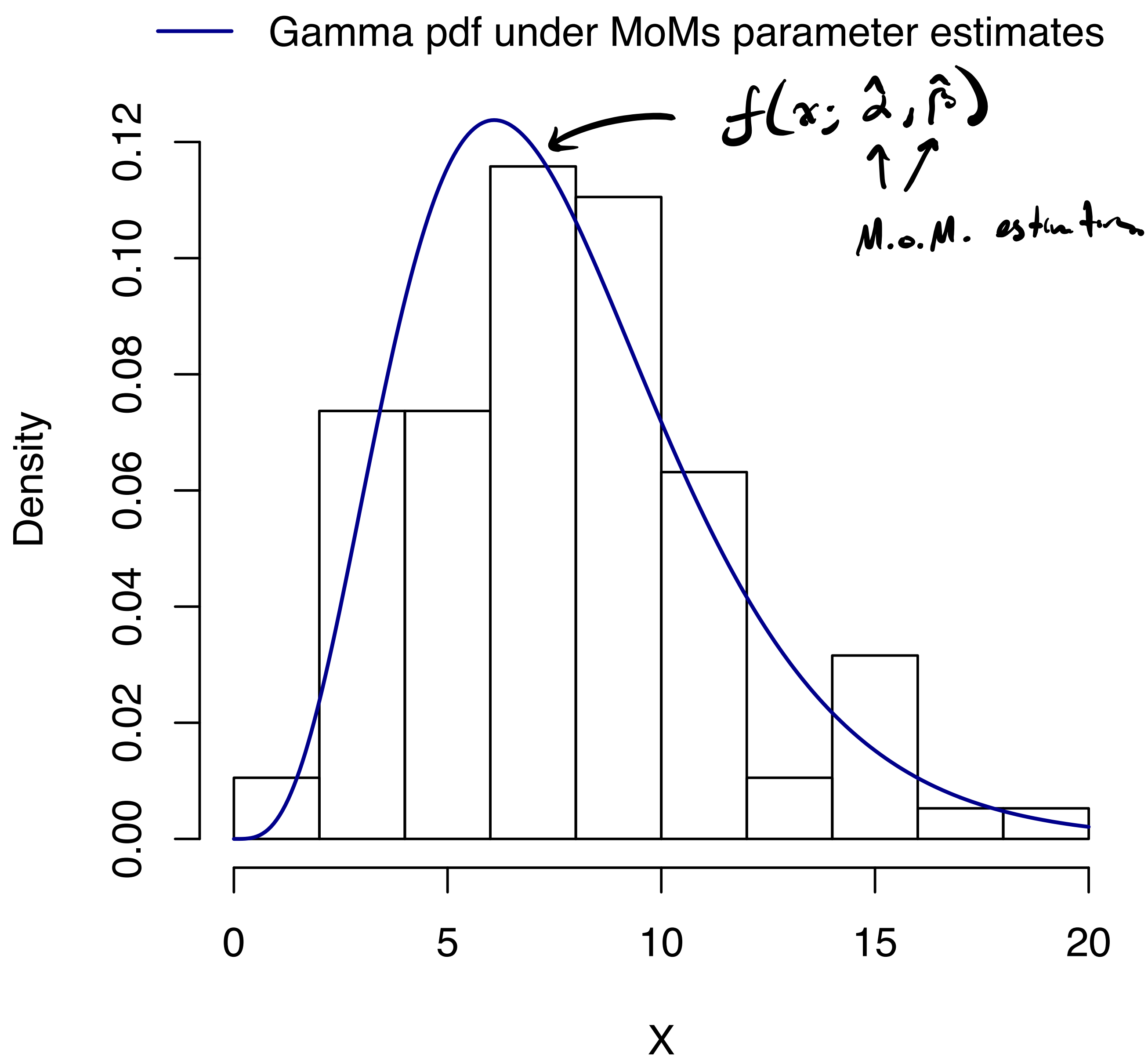
$\hat{m}_1 = \frac{1}{n}\sum_{i=1}^{n} X_i$

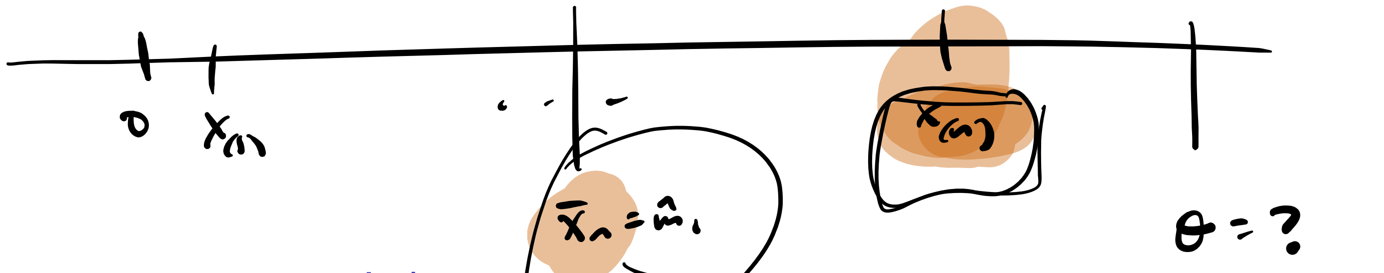
② $$f(x; \alpha, \beta) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-x/\beta} \mathbb{1}(x>0)$$

$$f(\underline{x}; \alpha, \beta) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^\alpha} x_i^{\alpha-1} e^{-x_i/\beta} \mathbb{1}(x>0)$$

$$= \left(\frac{1}{\Gamma(\alpha)\beta^\alpha}\right)^n \left(\prod x_i\right)^{\alpha-1} e^{-\sum x_i/\beta} \mathbb{1}(x>0)$$

$\therefore$ Suff stat is $T(\underline{X}) = \left(\prod X_i, \sum X_i\right)$


Gamma pdf under MoMs parameter estimates
$f(x; \hat{\alpha}, \hat{\beta})$
M.o.M. estimators
Density
0.00
0.02
0.04
0.06
0.08
0.10
0.12
0
5
10
15
20
X



Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Uniform}(0, \theta)$.

1. Find the MoMs estimator of θ.
2. Discuss whether better a estimator might exist.

1. $m_1 = \frac{\theta}{2} \iff \theta = 2m_1$ $\qquad \hat{\theta} = 2\hat{m}_1$

Exercise: Let $X_1, \dots, X_n \stackrel{\text{ind}}{\sim}$ Weibull(a, b). The Weibull(a, b) pdf is given by

$$f_X(x; a, b) = \left(\frac{a}{b}\right)\left(\frac{x}{b}\right)^{a-1} \exp\left[-\left(\frac{x}{b}\right)^{a}\right] \mathbf{1}(x > 0).$$

1. Find the MoMs estimators of a and b.
2. Discuss whether better estimators might exist.
3. Compute MoMs estimators on the trees in Camden data set.

$$m_1 = b\,\Gamma\left(1 + \frac{1}{a}\right)$$

$$m_2 = b^2\,\Gamma\left(1 + \frac{2}{a}\right)$$

$$b = \frac{m_1}{\Gamma\left(1 + \frac{1}{a}\right)}$$

$$\frac{m_1^2}{m_2} = \frac{b^2\,\Gamma^2\left(1 + \frac{1}{a}\right)}{b^2\,\Gamma\left(1 + \frac{2}{a}\right)} = \frac{\Gamma^2\left(1 + \frac{1}{a}\right)}{\Gamma\left(1 + \frac{2}{a}\right)}$$

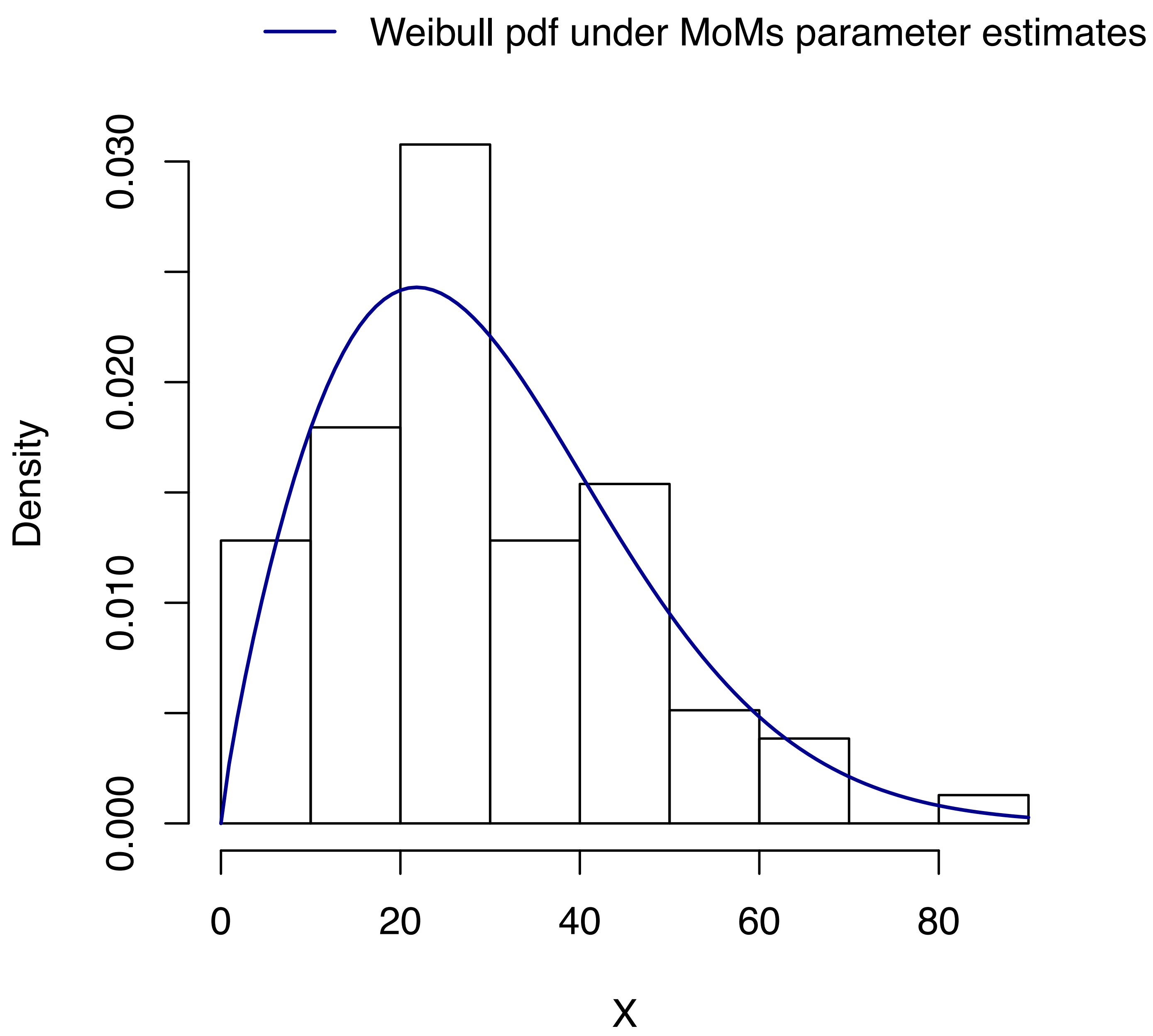
Need to find a which solves

$$\frac{m_1^2}{m_2} - \frac{\Gamma^2\left(1+\frac{1}{a}\right)}{\Gamma\left(1+\frac{2}{a}\right)} = 0$$

So MoMs estimators $\hat{a}, \hat{b}$ are solutions to

$$\hat{b} = \frac{\hat{m}_1}{\Gamma\left(1+\frac{1}{\hat{a}}\right)}$$

$$\frac{\hat{m}_1^2}{\hat{m}_2} - \frac{\Gamma^2\left(1+\frac{1}{\hat{a}}\right)}{\Gamma\left(1+\frac{1}{\hat{a}}\right)} = 0$$


Weibull pdf under MoMs parameter estimates
Density
0.000
0.010
0.020
0.030
0
20
40
60
80
X

The *maximum-likelihood* approach asks:

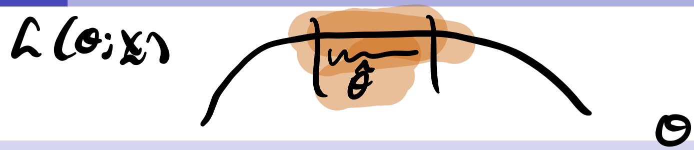
Under which value of the parameter are the observed data the most "likely"?

Likelihood and log-likelihood functions

Let $\mathbf{X}$ have joint pdf/pmf $\boxed{f(\mathbf{x};\theta)}$, where $\theta \in \Theta$.

1. The function $\mathcal{L}(\theta;\mathbf{X}) = \boxed{f(\mathbf{X};\theta)}$ is called the *likelihood function*.
2. The function $\ell(\theta;\mathbf{X}) = \log \mathcal{L}(\theta;\mathbf{X})$ is called the *log-likelihood function*.

The likelihood is just the joint pdf/pmf of the rvs in the random sample.



Maximum likelihood estimator (MLE)

The statistic $\hat{\theta} := \hat{\theta}(\mathbf{X})$ is an *MLE* for θ if $\mathcal{L}(\theta; \mathbf{X}) \leq \mathcal{L}(\hat{\theta}; \mathbf{X})$ for all $\theta \in \Theta$.

We obtain a/the MLE for θ by maximizing the likelihood function.

The maximizer may not be unique, but it is for most of our favorite settings.

We sometimes write the MLE $\hat{\theta}$ as

$$\hat{\theta} \in \operatorname{argmax}_{\theta \in \Theta} \mathcal{L}(\theta; \mathbf{X}) = \operatorname{argmax}_{\theta \in \Theta} \ell(\theta; \mathbf{X}).$$

If the maximizer of $\mathcal{L}(\theta; \mathbf{X})$ is unique, then "$\in$" changes to "$=$".

In the above, $\operatorname{argmax}_{z \in A} g(z) = \{z \in A : g(z') \leq g(z) \text{ for all } z' \in A\}$.

We can often (not always) use calculus methods to find the MLEs...

Theorem (Finding MLEs with calculus)

Suppose $\ell(\theta; \mathbf{X})$ *is differentiable and has a single maximum in the interior of* Θ. *If* $\theta = (\theta_1, \dots, \theta_d)$, *then the MLE* $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_d)$ *is the solution to the system*

$$\frac{\partial}{\partial \theta_j} \ell(\theta; \mathbf{X}) = 0, \quad j = 1, \dots, d.$$

Steps for finding MLE via calculus:

1. Set 1st-order partial derivatives equal to zero and solve system of equations.
2. Check Hessian to see if a maximum (need neg-def: $\mathbf{a}^T \mathbf{H} \mathbf{a} < 0 \ \forall \mathbf{a}$).
3. Check for uniqueness or a solution on the boundary of Θ.

Most often easier to differentiate the log-likelihood than the likelihood.

$m_1 = \lambda$ $\hat{m}_1 = \bar{X}_n$

$\hat{\lambda}_{mom} = \bar{X}_n$

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Exponental}(\lambda)$.

$f(x;\lambda) = \frac{1}{\lambda} e^{-x/\lambda} \mathbb{1}(x > 0)$.

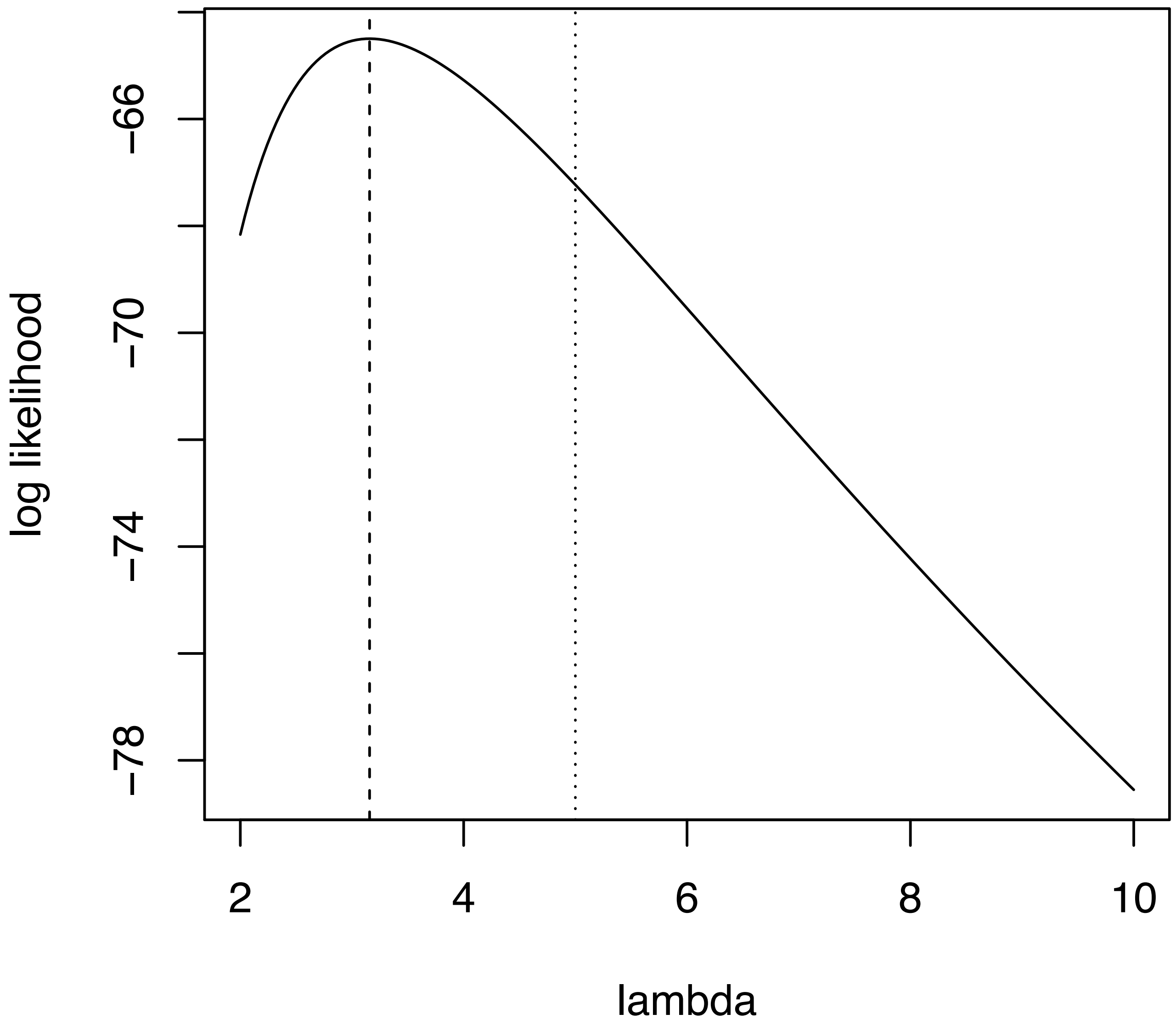
1. Find the MLE of λ.
2. Plot the log-likelihood function based on a simulated sample under $\lambda = 5$.

①

$$L(\lambda; \underset{\sim}{X}) = f(\underset{\sim}{X}; \lambda) = \prod_{i=1}^{n} \frac{1}{\lambda} e^{-X_i/\lambda} = \left(\frac{1}{\lambda}\right)^n e^{-\frac{\sum_{i=1}^n X_i}{\lambda}}$$

$$\ell(\lambda; \underset{\sim}{X}) = -n \log \lambda - \frac{\sum_{i=1}^n X_i}{\lambda}$$

$$\frac{\partial}{\partial \lambda} \ell(\lambda; \underset{\sim}{X}) = -\frac{n}{\lambda} + \frac{\sum_{i=1}^n X_i}{\lambda^2} \overset{\text{set}}{=} 0 \implies \hat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i = \bar{X}_n.$$



log likelihood
−66
−70
−74
−78
2
4
6
8
10
lambda

Exercise: Let $X_1, \ldots, X_n \stackrel{\text{ind}}{\sim} \text{Normal}(\mu, \boxed{\sigma^2})$. Find the MLEs of μ and $\boxed{\sigma^2}$.

$$f(x; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left[- \frac{(x-\mu)^2}{2\sigma^2} \right]$$

$$L(\mu, \sigma^2; \underset{\sim}{X}) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma} \exp\left[- \frac{(X_i - \mu)^2}{2\sigma^2} \right]$$

$$= (2\pi)^{-n/2} (\sigma^2)^{-n/2} \exp\left[- \frac{\sum (X_i - \mu)^2}{2\sigma^2} \right].$$

$$\ell(\mu, \sigma^2; \underset{\sim}{X}) = -\frac{n}{2} \log(2\pi) - \frac{n}{2} \log \sigma^2 - \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{2\sigma^2}$$

$\Leftrightarrow \quad \sum x_i - n\mu = 0$

$$\frac{\partial}{\partial \mu} \ell(\mu, \sigma^2; \underset{\sim}{x}) = \frac{1}{\sigma^2} \sum (x_i - \mu) = 0 \Leftrightarrow \boxed{\mu = \frac{1}{n} \sum x_i = \bar{x}_n}$$

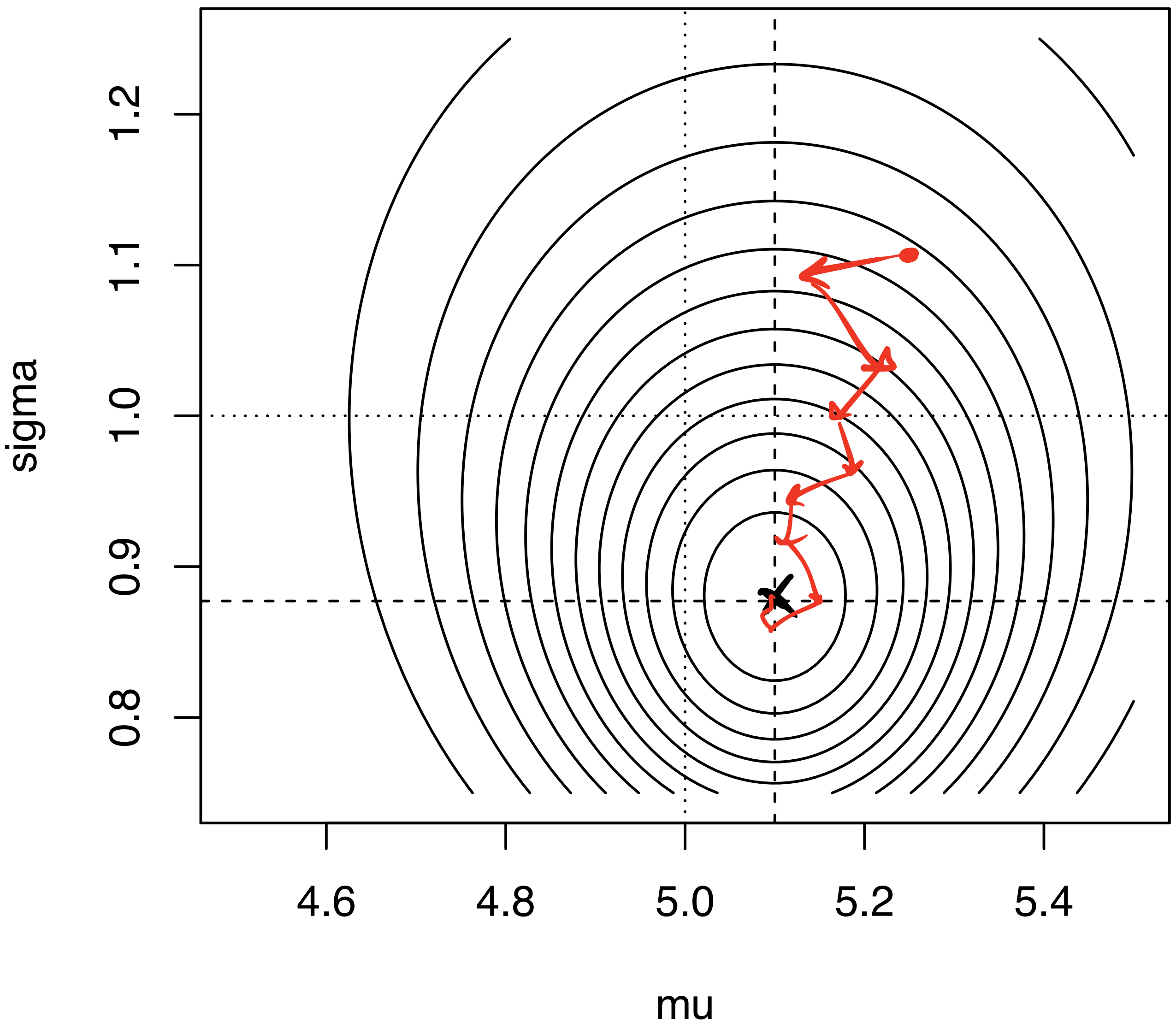
$$\frac{\partial}{\partial \sigma^2} \ell(\mu, \sigma^2; \underset{\sim}{x}) = \underbrace{-\frac{n}{2\sigma^2} + \frac{\sum_{i=1}^n (x_i - \mu)^2}{2(\sigma^2)^2}} = 0$$

$$\boxed{\sigma^2 = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2}$$

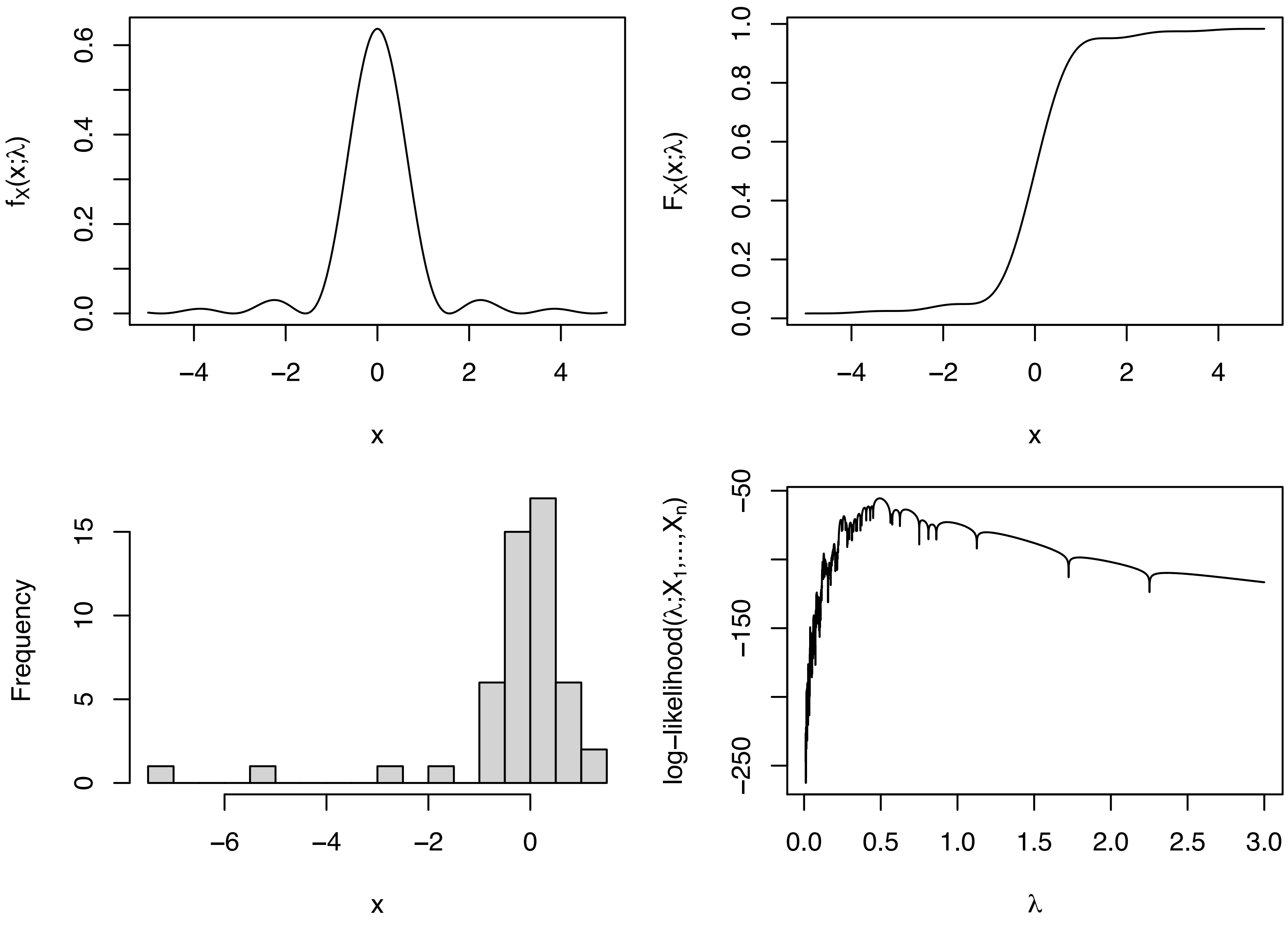
MLEs.

$$\boxed{\begin{aligned} \hat{\mu} &= \bar{x}_n \\ \hat{\sigma}^2 &= \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x}_n)^2 \end{aligned}}$$

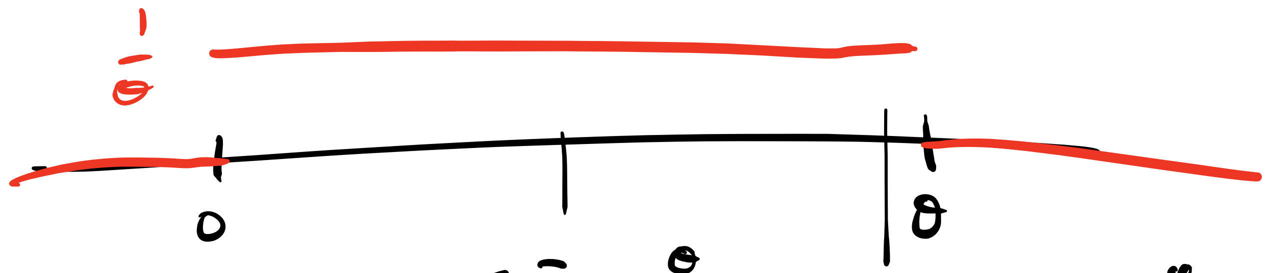
$$S_n^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x}_n)^2$$


sigma
1.2
1.1
1.0
0.9
0.8
4.6
4.8
5.0
5.2
5.4
mu

Example: Let $X_1, \dots, X_n \sim f_X(x;\lambda) = \frac{\lambda}{\pi}\frac{\sin^2(x/\lambda)}{x^2}$. Consider finding the MLE of λ.



$\hat{\theta}_{MOM} = 2\bar{x}_n$



$X_{(n)}$ is min suff.

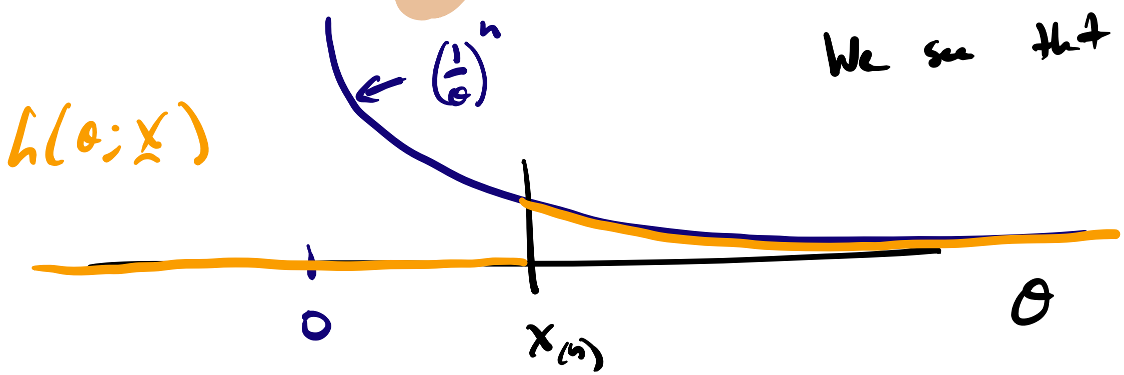
Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Uniform}(0, \theta)$. Find the MLE of θ. $\in \Theta = (0, \infty)$

$$f(x;\theta) = \frac{1}{\theta}\, \mathbb{1}(0 \le x \le \theta)$$

$$L(\theta; \underset{\sim}{x}) = \prod_{i=1}^{n} \frac{1}{\theta}\, \mathbb{1}(0 \le x_i \le \theta)$$

$$= \left(\frac{1}{\theta}\right)^n \mathbb{1}(0 \le x_{(1)})\, \mathbb{1}(x_{(n)} \le \theta)$$

We see that $\hat{\theta}_{mle} = X_{(n)}$.



Theorem (MLEs functions of minimal sufficient statistics)

If the MLE is unique, it is a function of a minimal sufficient statistic.

So MLEs use all the information in the sample about the target parameter.

Exercise: Prove sufficiency part of above result.

$$L(\theta; \underset{\sim}{X}) = f(\underset{\sim}{X}; \theta) = g(T(\underset{\sim}{X}); \theta)\, \cancel{h(\underset{\sim}{X})}$$

$\Rightarrow \underset{\theta \in \Theta}{\arg\max}\, L(\theta; \underset{\sim}{X})$ depends on the data only via $T(\underset{\sim}{X})$.

Let $X_1, \ldots, X_n \sim f(x;\beta) = \beta x^{-(\beta+1)} \mathbb{1}(x>1)$.

Find the MLE for β?

$$L(\beta;\underset{\sim}{X}) = \prod_{i=1}^{n} \beta X_i^{-(\beta+1)} = \beta^n \left(\prod_{i=1}^{n} X_i\right)^{-(\beta+1)}$$

$$\ell(\beta;\underset{\sim}{X}) = n \log \beta - (\beta+1) \sum_{i=1}^{n} \log X_i$$

$$\frac{\partial}{\partial\beta} \ell(\beta;\underset{\sim}{X}) = \frac{n}{\beta} - \sum_{i=1}^{n} \log X_i \overset{\text{set}}{=} 0$$

$$\boxed{\hat{\beta}_{MLE} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} \log X_i}}$$

How to find MoM estimator

$$m_1 = \mathbb{E} X_1 = \int_1^\infty x \cdot \beta x^{-(\beta+1)} \, dx = \frac{\beta}{\beta-1} \quad (\text{Require } \beta > 1)$$

$$f(x;\alpha) = \frac{1}{\Gamma(\alpha)2^{\alpha}} x^{\alpha-1} e^{-x/2} \mathbb{1}(x>0)$$

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Gamma}(\alpha, 2)$.

1. Find the MLE $\hat{\alpha}$ of α. ✓
2. Plot many values of the pair $(\hat{\alpha}, \prod_{i=1}^n X_i)$ for simulated data under $\alpha = 3$.
3. Compare the MSE of $\hat{\alpha}$ to that of the MoMs estimator $\bar{\alpha}$.

①

$$L(\alpha; \underset{\sim}{X}) = \prod_{i=1}^n \frac{1}{\Gamma(\alpha)2^{\alpha}} X_i^{\alpha-1} e^{-X_i/2} \mathbb{1}(x>0)$$

$$= \left(\frac{1}{\Gamma(\alpha)2^{\alpha}}\right)^n \left(\prod_{i=1}^n X_i\right)^{\alpha-1} e^{-\frac{\sum_{i=1}^n X_i}{2}} \prod_{i=1}^n \mathbb{1}(X_i>0)$$

Does not involve α

$$\ell(\alpha; \underset{\sim}{X}) = -n\log\Gamma(\alpha) - n\alpha\log 2 + (\alpha-1)\sum_{i=1}^n \log X_i - \frac{\sum_{i=1}^n X_i}{2}.$$

$$\frac{\partial}{\partial \alpha} \ell(\alpha; \underset{\sim}{X}) = -n \underbrace{\frac{\Gamma'(\alpha)}{\Gamma(\alpha)}}_{} - n \log 2 + \sum_{i=1}^{n} \log X_i \overset{\text{set}}{=} 0.$$

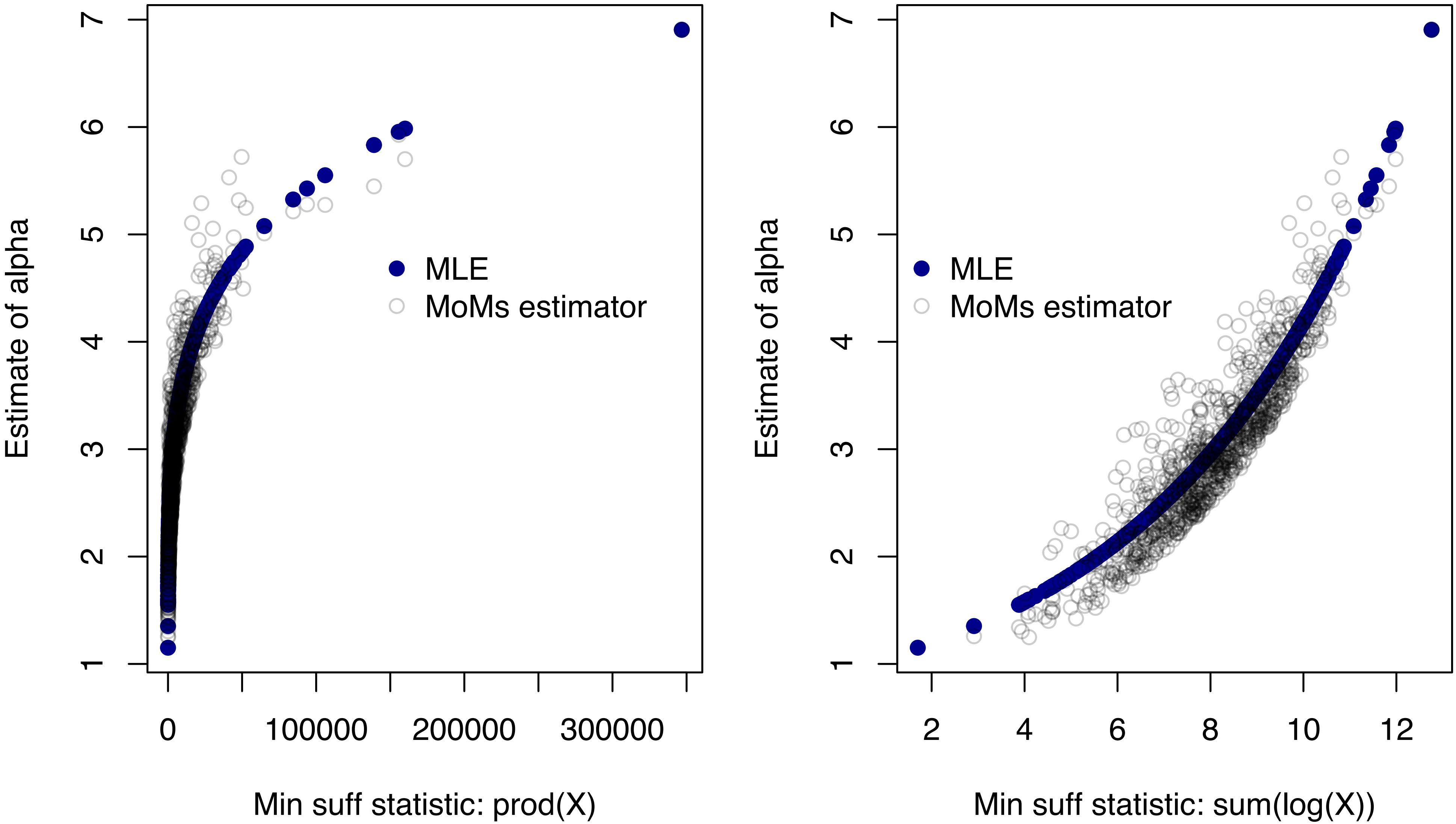
$$\frac{\Gamma'(\alpha)}{\Gamma(\alpha)} = \psi(\alpha) = \text{"digamma" function}$$

$$\hat{\alpha} = g\Big(\underbrace{\prod_{i=1}^{n} X_i}_{} \Big)$$

min suff stat

(3) $m_1 = \alpha 2 \iff \alpha = \dfrac{m_1}{2}$

$$\hat{\alpha}_{MoM} = \frac{\bar{X}_n}{2}$$


Estimate of alpha
1
2
3
4
5
6
7
0
100000
200000
300000
Min suff statistic: prod(X)
MLE
MoMs estimator
Estimate of alpha
2
4
6
8
10
12
Min suff statistic: sum(log(X))
MLE
MoMs estimator

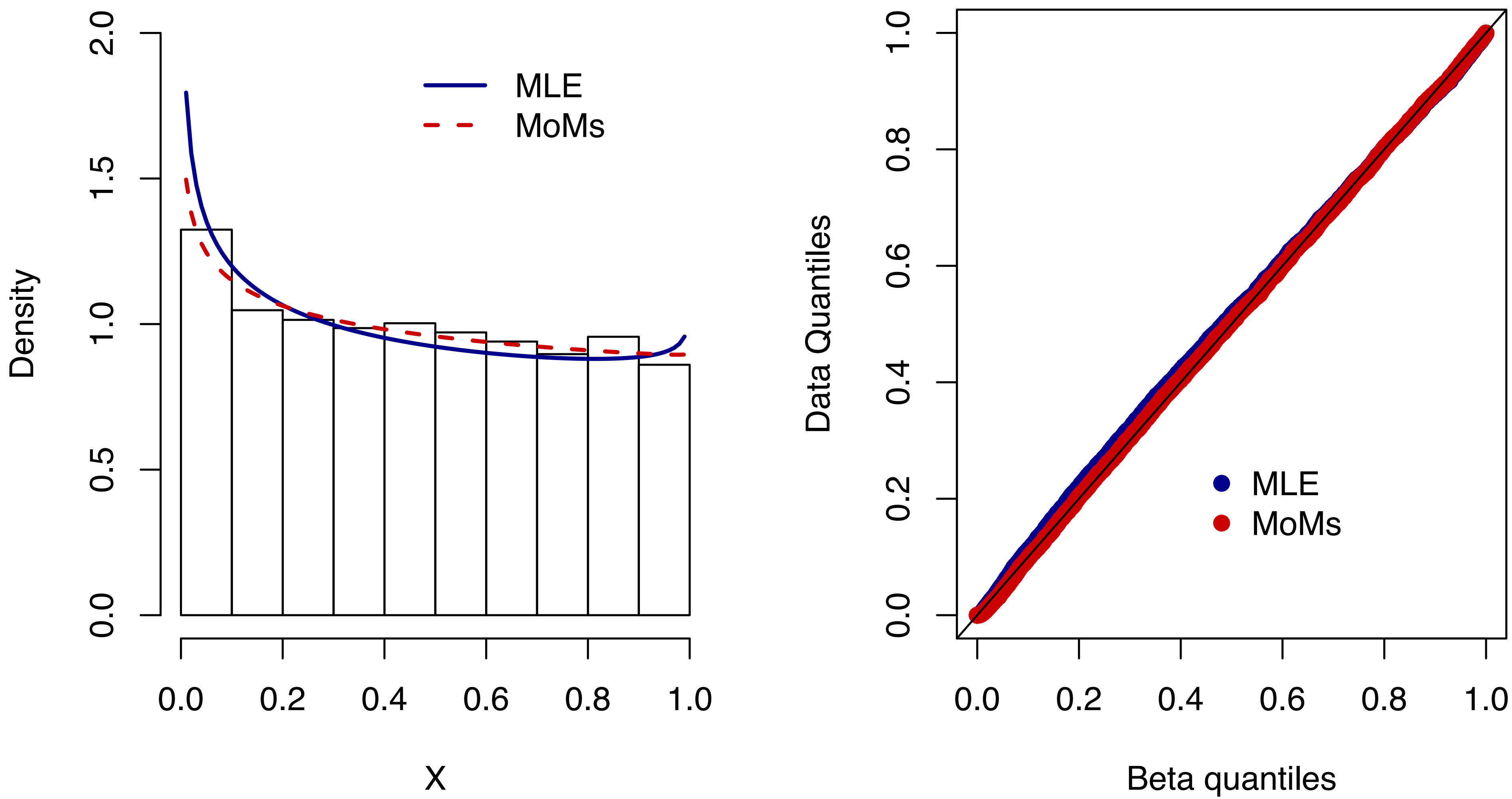
$$f(x;\alpha,\beta)=\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1}\mathbb{1}(0<x<1)$$

$$L(\alpha,\beta;X)=\prod_{i=1}^{n}\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}X_i^{\alpha-1}(1-X_i)^{\beta-1}=\left(\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right)^n\left(\prod_{i=1}^{n}X_i\right)^{\alpha-1}\left(\prod_{i=1}^{n}(1-X_i)\right)^{\beta-1}$$

Exercise: Let $X_1,\dots,X_n \overset{\text{ind}}{\sim} \text{Beta}(\alpha,\beta)$.

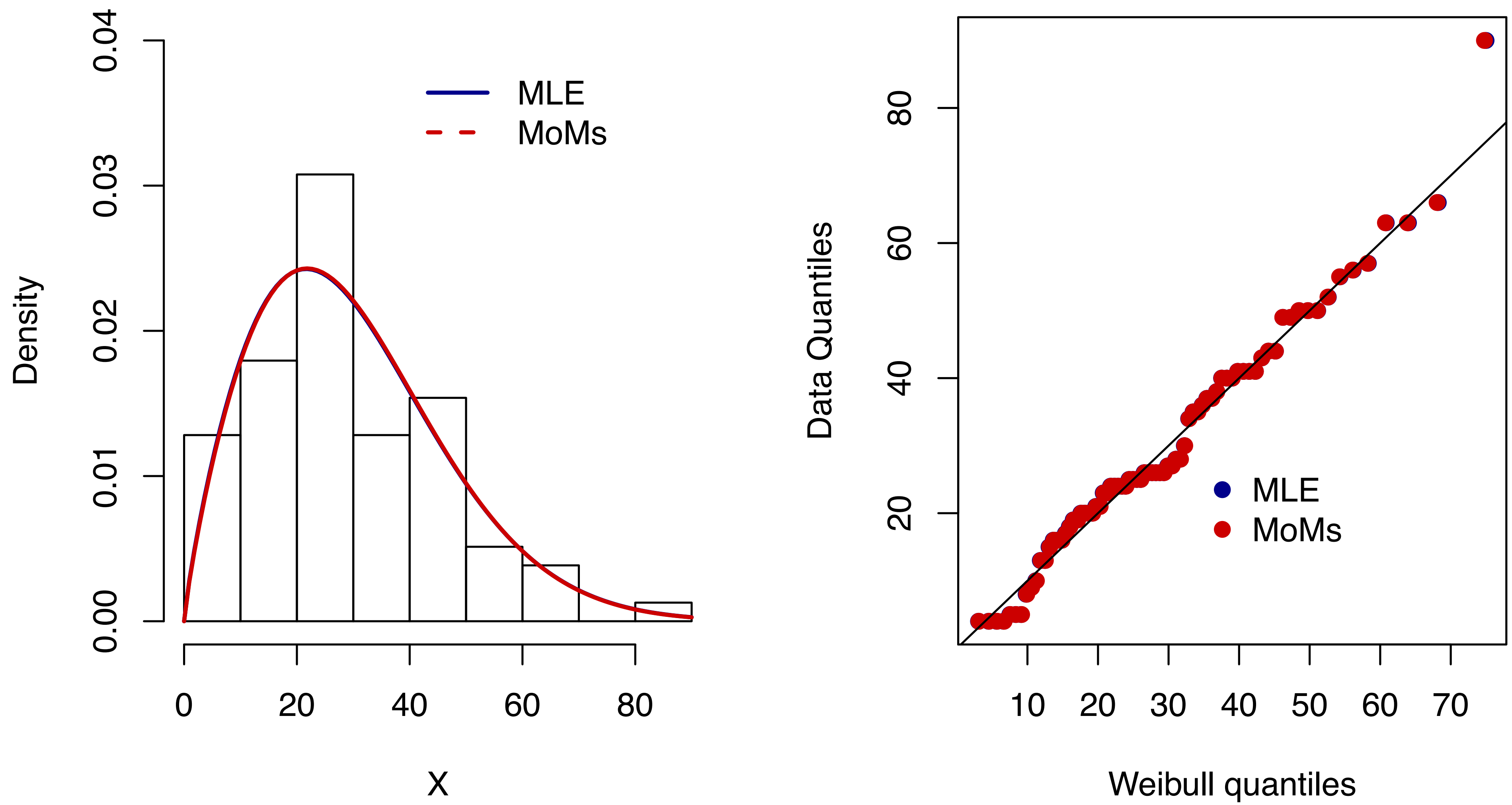
1. Find the MLEs $\hat{\alpha}$ and $\hat{\beta}$ of α and β.
2. Compute MLEs on prostate cancer p-values data from Efron (2012).
3. Compare to MoMs estimates.

$$\ell(\alpha,\beta;X)=n\log\Gamma(\alpha+\beta)-n\log\Gamma(\alpha)-n\log\Gamma(\beta)$$
$$+(\alpha-1)\sum_{i=1}^{n}\log X_i+(\beta-1)\sum_{i=1}^{n}\log(1-X_i)$$


MLE
MoMs
Density
X
0.0
0.5
1.0
1.5
2.0
0.0
0.2
0.4
0.6
0.8
1.0
Data Quantiles
Beta quantiles
MLE
MoMs

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} \text{Weibull}(a, b)$.

1. Find the MLEs $\hat{a}$ and $\hat{b}$ of a and b.
2. Compute MLEs on the trees in Camden data set.
3. Compare to MoMs estimates.



Density
0.00
0.01
0.02
0.03
0.04
MLE
MoMs
0
20
40
60
80
X
Data Quantiles
20
40
60
80
10
20
30
40
50
60
70
Weibull quantiles

Suppose we are interested in a function $\eta = \tau(\theta)$ of the parameter(s).

Induced likelihood or profile likelihood

eta

For $\mathbf{X}$ with likelihood $\mathcal{L}(\theta; \mathbf{X})$ and a new parameter $\eta = \tau(\theta)$, the function

$$\mathcal{L}^*(\eta; \mathbf{X}) = \sup_{\{\theta \in \Theta : \tau(\theta) = \eta\}} \mathcal{L}(\theta; \mathbf{X}).$$

is called the *induced likelihood* for η.

Exercise: Give the induced likelihood for η in the following settings:

1. $\mathbf{X} \overset{\text{ind}}{\sim} \text{Uniform}(a, b)$, with $\eta = b - a$.
2. $\mathbf{X} \overset{\text{ind}}{\sim} \text{Gamma}(\alpha, \beta)$ with $\eta = \alpha$ (discuss *profile likelihood*, *nuisance param*).

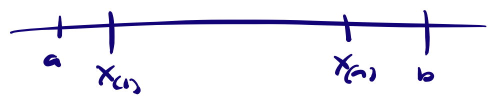
(a) $X_1,\dots,X_n \sim f(x;a,b) = \frac{1}{b-a}\,\mathbb{1}(a<x<b)$

$$\eta = \tau(a,b) = b-a$$

$$L(a,b;\underset{\sim}{x}) = \left(\frac{1}{b-a}\right)^n \mathbb{1}(a<x_{(1)})\,\mathbb{1}(x_{(n)}<b)$$

Induced likelihood for $\eta = b-a$ is

$$L^*(\eta;\underset{\sim}{x}) = \sup_{a<b:\; b-a=\eta} L(a,b;\underset{\sim}{x})$$



$$= \sup_{a<b;\; b-a=\eta} \left(\frac{1}{b-a}\right)^n \underbrace{\mathbb{1}(a<x_{(1)})}_{\le 1}\,\mathbb{1}(x_{(n)}-x_{(1)}<b-a)\,\underbrace{\mathbb{1}(x_{(n)}<b)}_{\le 1}$$

$$= \left(\frac{1}{\eta}\right)^n \mathbb{1}(x_{(n)}-x_{(1)}<\eta)$$

(2) $X_1, \ldots, X_n \overset{ind}{\sim} \text{Gamma}(\alpha, \beta)$, $\eta = \tau(\alpha,\beta) = \alpha$

$$L(\alpha,\beta;\underset{\sim}{X}) = \prod_{i=1}^{n} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} X_i^{\alpha-1} e^{-X_i/\beta}$$

$$= \left[\frac{1}{\Gamma(\alpha)\beta^{\alpha}}\right]^n \left(\prod_{i=1}^{n} X_i\right)^{\alpha-1} e^{-\frac{\sum_{i=1}^{n} X_i}{\beta}}.$$

$$\ell(\alpha,\beta;\underset{\sim}{X}) = -n\log\Gamma(\alpha) - n\alpha\log\beta + (\alpha-1)\sum_{i=1}^{n}\log X_i - \frac{1}{\beta}\sum_{i=1}^{n} X_i$$

The "induced" or "profile" likelihood is

$$L(\eta;\underset{\sim}{X}) = \sup_{\{\alpha,\beta\,:\,\alpha=\eta\}} L(\alpha,\beta;\underset{\sim}{X})$$

$$\mathcal{L}^*(\eta;\underset{\sim}{\mathbf{X}}) = \sup_{\{\theta\in\Theta:\tau(\theta)=\eta\}} \mathcal{L}(\theta;\mathbf{X}).$$

$$= \sup_{\beta} L(\eta,\beta;\underset{\sim}{X})$$

Find β which maximizes $L(\eta,\beta;\underset{\sim}{X})$ and plug it in.

$$\ell(\eta,\beta;\underset{\sim}{X}) = -n\log\Gamma(\eta) - n\eta\log\beta + (\eta-1)\sum_{i=1}^{n}\log X_i - \frac{1}{\beta}\sum_{i=1}^{n} X_i$$

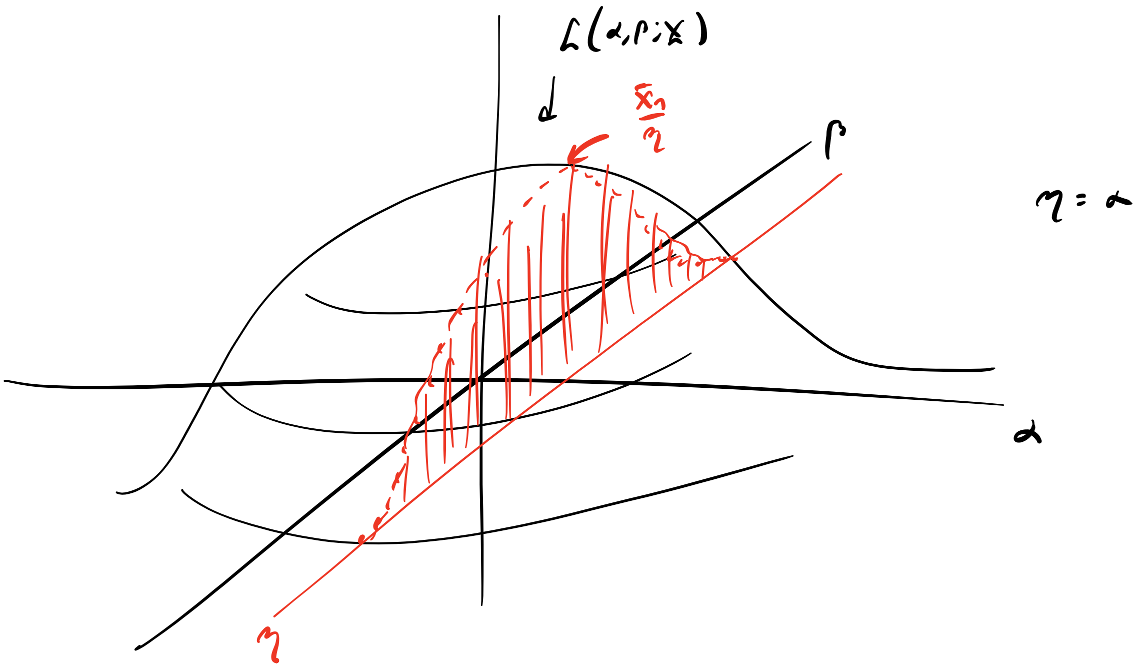
$$\frac{\partial}{\partial\beta}\ell(\eta,\beta;\underset{\sim}{X}) = \frac{-n\eta}{\beta} + \frac{1}{\beta^2}\sum_{i=1}^{n} X_i = 0$$

$$\iff \beta = \frac{\sum_{i=1}^{n} X_i}{n\eta} = \frac{\bar{X}_n}{\eta}$$

Now for $\eta = \alpha$:

$$L^*(\eta; \underset{\sim}{X}) = \sup_{\beta} L(\eta, \beta; \underset{\sim}{X})$$

$$= \left[\frac{1}{\Gamma(\eta)\left(\frac{\bar{X}_n}{\eta}\right)^{\eta}}\right]^n \left(\prod_{i=1}^n X_i\right)^{\eta-1} e^{-\frac{\sum_{i=1}^n X_i}{\bar{X}_n/\eta}}$$



Theorem (Invariance property of MLEs, cf. Thm 7.2.10 of CB)

Let $\hat{\theta}$ *be an MLE for* θ *and suppose* $\eta = \tau(\theta)$ *induces the likelihood* $\mathcal{L}^*(\eta; \mathbf{X})$. *Then* $\hat{\eta} = \tau(\hat{\theta})$ *is a maximizer of* $\mathcal{L}^*(\eta; \mathbf{X})$.

Says if $\hat{\theta}$ is an MLE for θ, then $\hat{\eta} = \tau(\hat{\theta})$ is an MLE for $\eta = \tau(\theta)$.

a maximum of $L(\cdot; \underset{\sim}{X})$

Exercise: Prove the result.

For any η, $L^*(\eta; \underset{\sim}{X}) = \sup_{\{\theta \in \Theta : \tau(\theta) = \eta\}} L(\theta; \underset{\sim}{X}) \leq L(\hat{\theta}_n; \underset{\sim}{X})$

Also, for $\hat{\eta}$, $L^*(\hat{\eta}; \underset{\sim}{X}) = \sup_{\{\theta \in \Theta ; \tau(\theta) = \hat{\eta}\}} L(\theta; \underset{\sim}{X}) \geq L(\hat{\theta}_n; \underset{\sim}{X})$

$$\Rightarrow \quad L^+(\eta ; \underset{\sim}{X}) \le L^+(\hat{\eta} ; \underset{\sim}{X}),$$

$\hat{\eta}$ is a maximizer of $L^+(\eta ; \underset{\sim}{X})$.

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim}$ Exponential(λ).

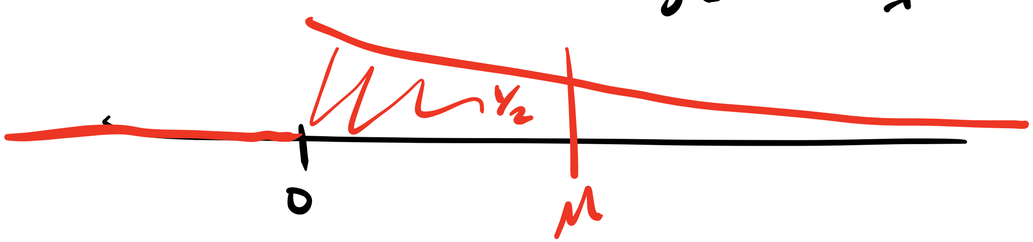
1. Find the MLE of $\eta = 1/\lambda$.
2. Find the MLE of the median of the Exponential(λ) distribution.

(1) $\hat{\eta}_{mle} = \dfrac{1}{\hat{\lambda}_{mle}}$, (by the invariance property of MLEs).

$\hat{\lambda}_{mle} = \bar{X}_n$. So $\boxed{\hat{\eta}_{mle} = \dfrac{1}{\bar{X}_n}}$

$f(x;\lambda) = \frac{1}{\lambda} e^{-x/\lambda} \mathbb{1}(x>0)$

(2)



$$\int_0^M \frac{1}{\lambda} e^{-x/\lambda}\, dx = \frac{1}{2}$$

$$F_X(x;\lambda) = \begin{cases} \int_0^x \frac{1}{\lambda} e^{-t/\lambda}\, dt = 1 - e^{-x/\lambda} & \text{for } x > 0 \\ 0, & \text{for } x \le 0 \end{cases}$$

Median is M such that $F_X(M;\lambda) = 1/2$

$$\iff \frac{1}{2} = 1 - e^{-M/\lambda}$$

$$e^{-M/\lambda} = 1/2$$

$$-\frac{M}{\lambda} = \log\left(\frac{1}{2}\right)$$

$$M = \lambda \log(2).$$

$\hat{M}_{mle} = \hat{\lambda}_{mle} \cdot \log(2) = \bar{X}_n \cdot \log(2)$ by inv. of MLE.

Table of Contents

Quantile estimators

Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} F(x; \theta)$, with pop. quantiles $\tau_1, \dots, \tau_k$ denoted by $\xi_{\tau_1}, \dots, \xi_{\tau_k}$.

If $\theta = g(\xi_{\tau_1}, \dots, \xi_{\tau_k})$ for some function g, a *quantile estimator* for θ is

$$\hat{\theta} = g(\xi_{\tau_1}, \dots, \xi_{\tau_k}), \quad \text{where } (\hat{\xi}_{\tau_1}, \dots, \hat{\xi}_{\tau_k}) = (X_{(\lceil \tau_1 n \rceil)}, \dots, X_{(\lceil \tau_k n \rceil)}).$$

One can define various quantile estimators for a parameter in a given setting.

Exercise: Let $X_1, \dots, X_n \overset{\text{ind}}{\sim} F(x; \lambda) = 1 - e^{-x/\lambda}$ for $x > 0$, for some $\lambda > 0$.

1. Give a two quantile estimators for λ. $\hat{\lambda}_{mle} = \bar{X}_n$
2. Run a simulation to compare their performance.

(1) Get Quantile function: Set $\tau = 1 - e^{-x/\lambda}$ and solve for x.

$$\tau = 1 - e^{-x/\lambda}$$

$$\Leftrightarrow \quad 1 - \tau = e^{-x/\lambda}$$

$$\Leftrightarrow \quad \log(1-\tau) = -x/\lambda$$

$$\Leftrightarrow \quad -\lambda \log(1-\tau) = x \quad .$$

The τ quantile ξ_τ is given by $\xi_\tau = -\lambda \log(1-\tau)$.

x_i

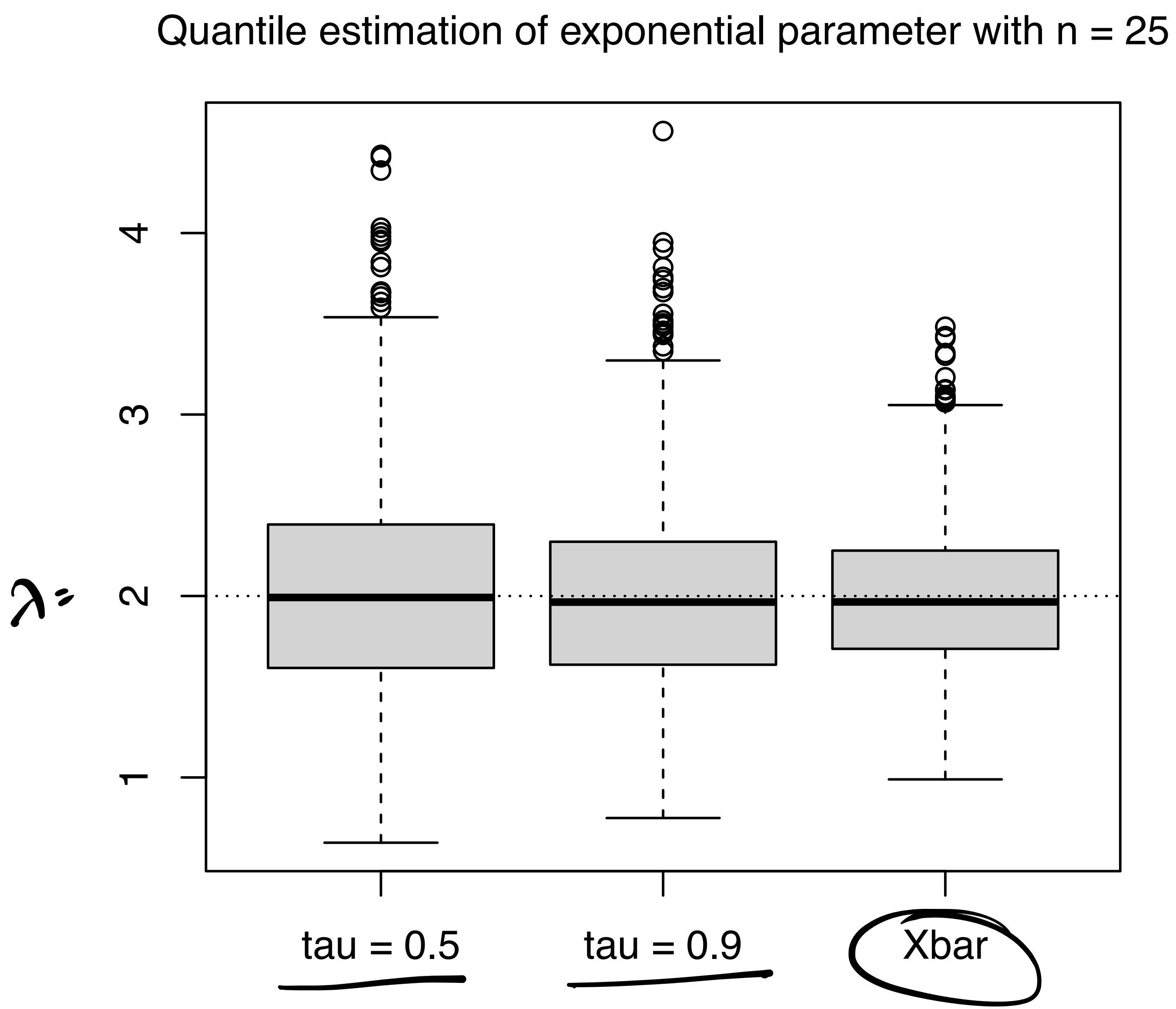
So, solving for λ gives

$$\lambda = \frac{\xi_\tau}{-\log(1-\tau)} \quad .$$

So a quantile estimator for λ is

$$\hat{\lambda}_\tau = \frac{\xi_\tau}{-\log(1-\tau)} \quad .$$

Can pick any $\tau \in (0,1)$.


Quantile estimation of exponential parameter with n = 25
λ =
1
2
3
4
tau = 0.5
tau = 0.9
Xbar

Anthony Christopher Davison. *Statistical models*, volume 11. Cambridge University Press, 2003.

Bradley Efron. *Large-scale inference: empirical Bayes methods for estimation, testing, and prediction*, volume 1. Cambridge University Press, 2012.