

# Computational Simplifications

$$I_n(\boldsymbol{\theta}) := E \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{X}; \boldsymbol{\theta}) \right)^2 \right]$$

**1**

$$E \left[ \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{X}; \boldsymbol{\theta}) \right] = 0$$

**Proof:**

$$E \left[ \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{X}; \boldsymbol{\theta}) \right] = \int \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{x}; \boldsymbol{\theta}) \right) f(\vec{x}; \boldsymbol{\theta}) d\mathbf{x}$$

$$E \left[ \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{X}; \boldsymbol{\theta}) \right] = \int \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{x}; \boldsymbol{\theta}) \right) f(\vec{x}; \boldsymbol{\theta}) d\mathbf{x}$$

$$= \int \left( \frac{\frac{\partial}{\partial \boldsymbol{\theta}} f(\vec{x}; \boldsymbol{\theta})}{f(\vec{x}; \boldsymbol{\theta})} \right) f(\vec{x}; \boldsymbol{\theta}) d\vec{x}$$

$$= \int \frac{\partial}{\partial \boldsymbol{\theta}} f(\vec{x}; \boldsymbol{\theta}) d\vec{x} = \frac{\partial}{\partial \boldsymbol{\theta}} \int f(\vec{x}; \boldsymbol{\theta}) d\vec{x}$$

$$= \frac{\partial}{\partial \boldsymbol{\theta}} 1 = 0$$



# Computational Simplifications

$$I_n(\boldsymbol{\theta}) := E \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{X}; \boldsymbol{\theta}) \right)^2 \right]$$

2

$$I_n(\boldsymbol{\theta}) = - E \left[ \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln f(\vec{X}; \boldsymbol{\theta}) \right]$$

Proof: From 1, we have that

$$\int \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{x}; \boldsymbol{\theta}) \right) f(\vec{x}; \boldsymbol{\theta}) d\vec{x} = 0$$

Take the derivative on both sides WRT  $\boldsymbol{\theta}$ .

$$\int \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{x}; \boldsymbol{\theta}) \right) f(\vec{x}; \boldsymbol{\theta}) d\vec{x} = 0$$

---


$$\int \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{x}; \boldsymbol{\theta}) \right) \frac{\partial}{\partial \boldsymbol{\theta}} f(\vec{x}; \boldsymbol{\theta}) d\vec{x} + \int \left( \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln f(\vec{x}; \boldsymbol{\theta}) \right) f(\vec{x}; \boldsymbol{\theta}) d\vec{x} = 0$$

$\frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{x}; \boldsymbol{\theta})$   
 $\cdot f(\vec{x}; \boldsymbol{\theta})$

---


$$\underbrace{E \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{X}; \boldsymbol{\theta}) \right)^2 \right]}_{I_n(\boldsymbol{\theta})} + E \left[ \frac{\partial^2}{\partial \boldsymbol{\theta}^2} \ln f(\vec{X}; \boldsymbol{\theta}) \right] = 0$$

✓

# Computational Simplifications

$$I_n(\boldsymbol{\theta}) := E \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{X}; \boldsymbol{\theta}) \right)^2 \right]$$

**3**

If  $X_1, X_2, \dots, X_n$  are iid, then:

$$I_n(\boldsymbol{\theta}) = n I_1(\boldsymbol{\theta})$$

**Proof:**

$$\begin{aligned} I_n(\boldsymbol{\theta}) &:= E \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{X}; \boldsymbol{\theta}) \right)^2 \right] \\ &= E \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln \prod_{i=1}^n f(X_i; \boldsymbol{\theta}) \right)^2 \right] \end{aligned}$$

$$= \mathbb{E} \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \sum_{i=1}^n \ln f(X_i; \boldsymbol{\theta}) \right)^2 \right]$$

$$= \mathbb{E} \left[ \left( \sum_{i=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(X_i; \boldsymbol{\theta}) \right)^2 \right]$$

$$= \mathbb{E} \left[ \left( \sum_{i=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(X_i; \boldsymbol{\theta}) \right) \left( \sum_{j=1}^n \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(X_j; \boldsymbol{\theta}) \right) \right]$$

Now pull the “j sum” into the “i sum”.

$$= E \left[ \sum_{i=1}^n \sum_{j=1}^n \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(X_i; \boldsymbol{\theta}) \right) \left( \frac{\partial}{\partial \boldsymbol{\theta}} f(X_j; \boldsymbol{\theta}) \right) \right]$$

$$= \sum_{i=1}^n \sum_{j=1}^n E \left[ \underbrace{\left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(X_i; \boldsymbol{\theta}) \right)}_{\text{independent if } j \neq i} \underbrace{\left( \frac{\partial}{\partial \boldsymbol{\theta}} f(X_j; \boldsymbol{\theta}) \right)}_{\text{independent if } j \neq i} \right]$$

independent if  $j \neq i$

In this case, the expectation factors and both are zero by

1

$$E \left[ \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(\vec{X}; \boldsymbol{\theta}) \right] = 0$$



The surviving terms are

$$= \sum_{i=1}^n E \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(X_i; \boldsymbol{\theta}) \right) \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(X_i; \boldsymbol{\theta}) \right) \right]$$

$$= \sum_{i=1}^n E \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(X_i; \boldsymbol{\theta}) \right)^2 \right]$$

Because the  $X_i$  are iid, these expectations are all the same!

$$= n \cdot E \left[ \left( \frac{\partial}{\partial \boldsymbol{\theta}} \ln f(X_i; \boldsymbol{\theta}) \right)^2 \right] = n I_1(\boldsymbol{\theta})$$





## Example:

$$X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} \exp(\text{rate} = \lambda)$$

Find the Cramér-Rao lower bound of the variance of all unbiased estimators of  $\lambda$ .

$$\tau(\lambda) = \lambda \qquad \text{Var}[T] \geq \frac{[\tau'(\lambda)]^2}{I_n(\lambda)}$$

# The Fisher Information:

$$I_n(\boldsymbol{\lambda}) := E \left[ \left( \frac{\partial}{\partial \boldsymbol{p}} \ln f(\vec{X}; \boldsymbol{\lambda}) \right)^2 \right]$$

pdf:  $f(x; \lambda) = \lambda e^{-\lambda x} I_{(0, \infty)}(x)$

joint pdf:

$$f(\vec{x}; \lambda) = \lambda^n e^{-\lambda \sum_{i=1}^n x_i} \prod_{i=1}^n I_{(0, \infty)}(x_i)$$

Take the log:

$$\ln f(\vec{x}; \lambda) = n \ln \lambda - \lambda \sum_{i=1}^n x_i$$

Take the derivative:

$$\frac{\partial}{\partial \lambda} \ln f(\vec{x}; \lambda) = \frac{n}{\lambda} - \sum_{i=1}^n x_i$$

Put the random variables in, square,  
and take the expectation.

$$I_n(\lambda) := E \left[ \left( \frac{\partial}{\partial \lambda} \ln f(\vec{X}; \lambda) \right)^2 \right]$$

$$= E \left[ \left( \sum_{i=1}^n x_i - \frac{n}{\lambda} \right)^2 \right]$$

Let  $Y = \sum_{i=1}^n x_i$ . We know that  $Y \sim \Gamma(n, \lambda)$ .  
 We know that  $E[Y] = n/\lambda$ .

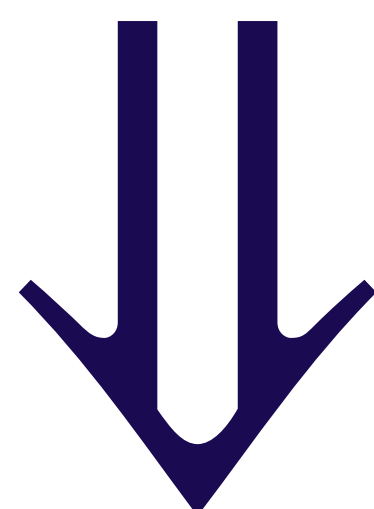
$$I_n(\lambda) := E \left[ \left( Y - \frac{n}{\lambda} \right)^2 \right] = \text{Var}[Y] = \frac{n}{\lambda^2}$$

$$\text{CRLB}_\lambda \geq \frac{[1]^2}{\text{I}_n(\lambda)} = \frac{1}{n/\lambda^2} = \frac{\lambda^2}{n}$$

---

Alternatively:

$$\frac{\partial}{\partial \lambda} \ln f(\vec{x}; \lambda) = \frac{n}{\lambda} - \sum_{i=1}^n x_i$$



$$\frac{\partial^2}{\partial \lambda^2} \ln f(\vec{x}; \lambda) = -\frac{n}{\lambda^2}$$

$$\frac{\partial^2}{\partial \lambda^2} \ln f(\vec{X}; \lambda) = -\frac{n}{\lambda^2}$$

$$I_n(\lambda) = -E \left[ \frac{\partial^2}{\partial \lambda^2} \ln f(\vec{X}; \lambda) \right]$$

$$= -E \left[ -\frac{n}{\lambda^2} \right] = \frac{n}{\lambda^2}$$

**which is the same thing we got with the single derivative expression.**

Either of these methods could have been done with a single  $X_1$  as opposed to the vector  $\vec{X}$ .

---

Example:

$$X_1, X_2, \dots, X_n \stackrel{\text{iid}}{\sim} \exp(\text{rate} = \lambda)$$

Find the Cramér-Rao lower bound of the variance of all unbiased estimators of  $\exp[-\lambda]$ .



We already did all of the work!

$$\tau(\lambda) = e^{-\lambda} \quad \text{Var}[T] \geq \frac{[\tau'(\lambda)]^2}{I_n(\lambda)}$$

The CRLB on the variance of all unbiased estimators of  $e^{-\lambda}$  is

$$\frac{[-\lambda e^{-\lambda}]^2}{n/\lambda^2} = \boxed{\frac{\lambda^4 e^{-2\lambda}}{n}}$$