

Math 362: Mathematical Statistics II

Le Chen

le.chen@emory.edu
chenle02@gmail.com

Emory University
Atlanta, GA

Last updated on Spring 2021
Last compiled on January 15, 2023

2021 Spring

Creative Commons License
(CC By-NC-SA)

Chapter 5. Estimation

§ 5.1 Introduction

§ 5.2 Estimating parameters: MLE and MME

§ 5.3 Interval Estimation

§ 5.4 Properties of Estimators

§ 5.5 Minimum-Variance Estimators: The Cramér-Rao Lower Bound

§ 5.6 Sufficient Estimators

§ 5.7 Consistency

§ 5.8 Bayesian Estimation

Plan

§ 5.1 Introduction

§ 5.2 Estimating parameters: MLE and MME

§ 5.3 Interval Estimation

§ 5.4 Properties of Estimators

§ 5.5 Minimum-Variance Estimators: The Cramér-Rao Lower Bound

§ 5.6 Sufficient Estimators

§ 5.7 Consistency

§ 5.8 Bayesian Estimation

Chapter 5. Estimation

§ 5.1 Introduction

§ 5.2 Estimating parameters: MLE and MME

§ 5.3 Interval Estimation

§ 5.4 Properties of Estimators

§ 5.5 Minimum-Variance Estimators: The Cramér-Rao Lower Bound

§ 5.6 Sufficient Estimators

§ 5.7 Consistency

§ 5.8 Bayesian Estimation

§ 5.5 MVE: The Cramér-Rao Lower Bound

Question: Can one identify the unbiased estimator having the *smallest* variance?

Short answer: In many cases, yes!

We are going to develop the theory to answer this question in details!

Regular Estimation/Condition: The set of y (resp. k) values, where $f_Y(y; \theta) \neq 0$ (resp. $p_X(k; \theta) \neq 0$), does not depend on θ .

i.e., the domain of the pdf does not depend on the parameter (so that one can differentiate under integration).

Definition. The **Fisher's Information** of a continuous (resp. discrete) random variable Y (resp. X) with pdf $f_Y(y; \theta)$ (resp. $p_X(k; \theta)$) is defined as

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial \ln f_Y(Y; \theta)}{\partial \theta} \right)^2 \right] \quad \left(\text{resp.} \quad \mathbb{E} \left[\left(\frac{\partial \ln p_X(X; \theta)}{\partial \theta} \right)^2 \right] \right).$$

Regular Estimation/Condition: The set of y (resp. k) values, where $f_Y(y; \theta) \neq 0$ (resp. $p_X(k; \theta) \neq 0$), does not depend on θ .

i.e., the domain of the pdf does not depend on the parameter (so that one can differentiate under integration).

Definition. The **Fisher's Information** of a continuous (resp. discrete) random variable Y (resp. X) with pdf $f_Y(y; \theta)$ (resp. $p_X(k; \theta)$) is defined as

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial \ln f_Y(Y; \theta)}{\partial \theta} \right)^2 \right] \quad \left(\text{resp.} \quad \mathbb{E} \left[\left(\frac{\partial \ln p_X(X; \theta)}{\partial \theta} \right)^2 \right] \right).$$

Lemma. Under regular condition, let Y_1, \dots, Y_n be a random sample of size n from the continuous population pdf $f_Y(y; \theta)$. Then the Fisher Information in the random sample Y_1, \dots, Y_n equals n times the Fisher information in X :

$$\mathbb{E} \left[\left(\frac{\partial \ln f_{Y_1, \dots, Y_n}(Y_1, \dots, Y_n; \theta)}{\partial \theta} \right)^2 \right] = n \mathbb{E} \left[\left(\frac{\partial \ln f_Y(Y; \theta)}{\partial \theta} \right)^2 \right] = n I(\theta). \quad (1)$$

(A similar statement holds for the discrete case $p_X(k; \theta)$).

Proof. Based on two observations:

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

$$\begin{aligned} \mathbb{E} \left(\frac{\partial}{\partial \theta} \ln f_{Y_i}(Y_i; \theta) \right) &= \int_{\mathbb{R}} \frac{\frac{\partial}{\partial \theta} f_Y(y; \theta)}{f_Y(y; \theta)} f_Y(y; \theta) dy = \int_{\mathbb{R}} \frac{\partial}{\partial \theta} f_Y(y; \theta) dy \\ &\stackrel{\text{R.C.}}{=} \frac{\partial}{\partial \theta} \int_{\mathbb{R}} f_Y(y; \theta) dy = \frac{\partial}{\partial \theta} 1 = 0. \end{aligned}$$

□

Lemma. Under regular condition, let Y_1, \dots, Y_n be a random sample of size n from the continuous population pdf $f_Y(y; \theta)$. Then the Fisher Information in the random sample Y_1, \dots, Y_n equals n times the Fisher information in X :

$$\mathbb{E} \left[\left(\frac{\partial \ln f_{Y_1, \dots, Y_n}(Y_1, \dots, Y_n; \theta)}{\partial \theta} \right)^2 \right] = n \mathbb{E} \left[\left(\frac{\partial \ln f_Y(Y; \theta)}{\partial \theta} \right)^2 \right] = n I(\theta). \quad (1)$$

(A similar statement holds for the discrete case $p_X(k; \theta)$).

Proof. Based on two observations:

$$\begin{aligned} LHS &= \mathbb{E} \left[\left(\sum_{i=1}^n \frac{\partial}{\partial \theta} \ln f_{Y_i}(Y_i; \theta) \right)^2 \right] \\ \mathbb{E} \left(\frac{\partial}{\partial \theta} \ln f_{Y_i}(Y_i; \theta) \right) &= \int_{\mathbb{R}} \frac{\frac{\partial}{\partial \theta} f_Y(y; \theta)}{f_Y(y; \theta)} f_Y(y; \theta) dy = \int_{\mathbb{R}} \frac{\partial}{\partial \theta} f_Y(y; \theta) dy \\ &\stackrel{\text{R.C.}}{=} \frac{\partial}{\partial \theta} \int_{\mathbb{R}} f_Y(y; \theta) dy = \frac{\partial}{\partial \theta} 1 = 0. \end{aligned}$$

□

Lemma. Under regular condition, if $\ln f_Y(y; \theta)$ is twice differentiable in θ , then

$$I(\theta) = -\mathbb{E} \left[\frac{\partial^2}{\partial \theta^2} \ln f_Y(Y; \theta) \right]. \quad (2)$$

(A similar statement holds for the discrete case $p_X(k; \theta)$).

Proof. This is due to the two facts:

$$\begin{aligned} \frac{\partial^2}{\partial \theta^2} \ln f_Y(Y; \theta) &= \frac{\frac{\partial^2}{\partial \theta^2} f_Y(Y; \theta)}{f_Y(Y; \theta)} - \underbrace{\left(\frac{\frac{\partial}{\partial \theta} f_Y(Y; \theta)}{f_Y(Y; \theta)} \right)^2}_{= \left(\frac{\partial}{\partial \theta} \ln f_Y(Y; \theta) \right)^2} \\ &= \left(\frac{\partial}{\partial \theta} \ln f_Y(Y; \theta) \right)^2 \end{aligned}$$

$$\begin{aligned} \mathbb{E} \left(\frac{\frac{\partial^2}{\partial \theta^2} f_Y(Y; \theta)}{f_Y(Y; \theta)} \right) &= \int_{\mathbb{R}} \frac{\frac{\partial^2}{\partial \theta^2} f_Y(y; \theta)}{f_Y(y; \theta)} f_Y(y; \theta) dy = \int_{\mathbb{R}} \frac{\partial^2}{\partial \theta^2} f_Y(y; \theta) dy. \\ &\stackrel{R.C.}{=} \frac{\partial^2}{\partial \theta^2} \int_{\mathbb{R}} f_Y(y; \theta) dy = \frac{\partial^2}{\partial \theta^2} 1 = 0. \end{aligned}$$

□

Lemma. Under regular condition, if $\ln f_Y(y; \theta)$ is twice differentiable in θ , then

$$I(\theta) = -\mathbb{E} \left[\frac{\partial^2}{\partial \theta^2} \ln f_Y(Y; \theta) \right]. \quad (2)$$

(A similar statement holds for the discrete case $p_X(k; \theta)$).

Proof. This is due to the two facts:

$$\begin{aligned} \frac{\partial^2}{\partial \theta^2} \ln f_Y(Y; \theta) &= \frac{\frac{\partial^2}{\partial \theta^2} f_Y(Y; \theta)}{f_Y(Y; \theta)} - \underbrace{\left(\frac{\frac{\partial}{\partial \theta} f_Y(Y; \theta)}{f_Y(Y; \theta)} \right)^2}_{= \left(\frac{\partial}{\partial \theta} \ln f_Y(Y; \theta) \right)^2} \\ &= \left(\frac{\partial}{\partial \theta} \ln f_Y(Y; \theta) \right)^2 \end{aligned}$$

$$\begin{aligned} \mathbb{E} \left(\frac{\frac{\partial^2}{\partial \theta^2} f_Y(Y; \theta)}{f_Y(Y; \theta)} \right) &= \int_{\mathbb{R}} \frac{\frac{\partial^2}{\partial \theta^2} f_Y(y; \theta)}{f_Y(y; \theta)} f_Y(y; \theta) dy = \int_{\mathbb{R}} \frac{\partial^2}{\partial \theta^2} f_Y(y; \theta) dy. \\ &\stackrel{R.C.}{=} \frac{\partial^2}{\partial \theta^2} \int_{\mathbb{R}} f_Y(y; \theta) dy = \frac{\partial^2}{\partial \theta^2} 1 = 0. \end{aligned}$$

□

Theorem (Cramér-Rao Inequality) Under regular condition, let Y_1, \dots, Y_n be a random sample of size n from the continuous population pdf $f_Y(y; \theta)$. Let $\hat{\theta} = \hat{\theta}(Y_1, \dots, Y_n)$ be any unbiased estimator for θ . Then

$$\text{Var}(\hat{\theta}) \geq \frac{1}{n I(\theta)}.$$

(A similar statement holds for the discrete case $p_X(k; \theta)$).

Proof. If $n = 1$, then by Cauchy-Schwartz inequality,

$$\mathbb{E} \left[(\hat{\theta} - \theta) \frac{\partial}{\partial \theta} \ln f_Y(Y; \theta) \right] \leq \sqrt{\text{Var}(\hat{\theta}) \times I(\theta)}$$

On the other hand,

$$\begin{aligned} \mathbb{E} \left[(\hat{\theta} - \theta) \frac{\partial}{\partial \theta} \ln f_Y(Y; \theta) \right] &= \int_{\mathbb{R}} (\hat{\theta} - \theta) \frac{\frac{\partial}{\partial \theta} f_Y(y; \theta)}{f_Y(y; \theta)} f_Y(y; \theta) dy \\ &= \int_{\mathbb{R}} (\hat{\theta} - \theta) \frac{\partial}{\partial \theta} f_Y(y; \theta) dy \\ &= \frac{\partial}{\partial \theta} \underbrace{\int_{\mathbb{R}} (\hat{\theta} - \theta) f_Y(y; \theta) dy}_{= \mathbb{E}(\hat{\theta} - \theta) = 0} + 1 = 1. \end{aligned}$$

For general n , apply for (1).

□.

Theorem (Cramér-Rao Inequality) Under regular condition, let Y_1, \dots, Y_n be a random sample of size n from the continuous population pdf $f_Y(y; \theta)$. Let $\hat{\theta} = \hat{\theta}(Y_1, \dots, Y_n)$ be any unbiased estimator for θ . Then

$$\text{Var}(\hat{\theta}) \geq \frac{1}{n I(\theta)}.$$

(A similar statement holds for the discrete case $p_X(k; \theta)$).

Proof. If $n = 1$, then by Cauchy-Schwartz inequality,

$$\mathbb{E} \left[(\hat{\theta} - \theta) \frac{\partial}{\partial \theta} \ln f_Y(Y; \theta) \right] \leq \sqrt{\text{Var}(\hat{\theta}) \times I(\theta)}$$

On the other hand,

$$\begin{aligned} \mathbb{E} \left[(\hat{\theta} - \theta) \frac{\partial}{\partial \theta} \ln f_Y(Y; \theta) \right] &= \int_{\mathbb{R}} (\hat{\theta} - \theta) \frac{\frac{\partial}{\partial \theta} f_Y(y; \theta)}{f_Y(y; \theta)} f_Y(y; \theta) dy \\ &= \int_{\mathbb{R}} (\hat{\theta} - \theta) \frac{\partial}{\partial \theta} f_Y(y; \theta) dy \\ &= \frac{\partial}{\partial \theta} \underbrace{\int_{\mathbb{R}} (\hat{\theta} - \theta) f_Y(y; \theta) dy}_{= \mathbb{E}(\hat{\theta} - \theta) = 0} + 1 = 1. \end{aligned}$$

For general n , apply for (1).

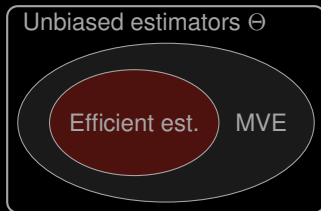
□.

Definition. Let Θ be the set of all estimators $\hat{\theta}$ that are unbiased for the parameter θ . We say that $\hat{\theta}^*$ is a **best** or **minimum-variance** estimator (MVE) if $\hat{\theta}^* \in \Theta$ and

$$\text{Var}(\hat{\theta}^*) \leq \text{Var}(\hat{\theta}) \quad \text{for all } \hat{\theta} \in \Theta.$$

Definition. An unbiased estimator $\hat{\theta}$ is **efficient** if $\text{Var}(\hat{\theta})$ is equal to the Cramér-Rao lower bound, i.e., $\text{Var}\hat{\theta} = (n I(\theta))^{-1}$.

The **efficiency** of an unbiased estimator $\hat{\theta}$ is defined to be $\left(n I(\theta) \text{Var}(\hat{\theta}) \right)^{-1}$.



E.g. 1. $X \sim \text{Bernoulli}(p)$. Check whether $\hat{p} = \bar{X}$ is efficient?

Step 1. Compute Fisher's Information:

$$p_X(k; p) = p^k (1 - p)^{1-k}.$$

E.g. 1. $X \sim \text{Bernoulli}(p)$. Check whether $\hat{p} = \bar{X}$ is efficient?

Step 1. Compute Fisher's Information:

$$p_X(k; p) = p^k (1 - p)^{1-k}.$$

$$\ln p_X(k; p) = k \ln p + (1 - k) \ln(1 - p)$$

$$\frac{\partial}{\partial p} \ln p_X(k; p) = \frac{k}{p} - \frac{1 - k}{1 - p}$$

$$-\frac{\partial^2}{\partial^2 p} \ln p_X(k; p) = \frac{k}{p^2} + \frac{1 - k}{(1 - p)^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 p} \ln p_X(X; p) \right] = \mathbb{E} \left[\frac{X}{p^2} + \frac{1 - X}{(1 - p)^2} \right] = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{pq}.$$

$$I(p) = \frac{1}{pq}, \quad q = 1 - p.$$

Step 2. Compute the variance of $\hat{p} = \bar{X}$ and compare it with $1/I(p)$.

E.g. 1. $X \sim \text{Bernoulli}(p)$. Check whether $\hat{p} = \bar{X}$ is efficient?

Step 1. Compute Fisher's Information:

$$p_X(k; p) = p^k (1 - p)^{1-k}.$$

$$\ln p_X(k; p) = k \ln p + (1 - k) \ln(1 - p)$$

$$\frac{\partial}{\partial p} \ln p_X(k; p) = \frac{k}{p} - \frac{1 - k}{1 - p}$$

$$-\frac{\partial^2}{\partial^2 p} \ln p_X(k; p) = \frac{k}{p^2} + \frac{1 - k}{(1 - p)^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 p} \ln p_X(X; p) \right] = \mathbb{E} \left[\frac{X}{p^2} + \frac{1 - X}{(1 - p)^2} \right] = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{pq}.$$

$$I(p) = \frac{1}{pq}, \quad q = 1 - p.$$

Step 2. Compute the variance of the estimator $\hat{p} = \bar{X}$ and compare it with the Cramér-Rao lower bound.

E.g. 1. $X \sim \text{Bernoulli}(p)$. Check whether $\hat{p} = \bar{X}$ is efficient?

Step 1. Compute Fisher's Information:

$$p_X(k; p) = p^k (1 - p)^{1-k}.$$

$$\ln p_X(k; p) = k \ln p + (1 - k) \ln(1 - p)$$

$$\frac{\partial}{\partial p} \ln p_X(k; p) = \frac{k}{p} - \frac{1 - k}{1 - p}$$

$$-\frac{\partial^2}{\partial^2 p} \ln p_X(k; p) = \frac{k}{p^2} + \frac{1 - k}{(1 - p)^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 p} \ln p_X(X; p) \right] = \mathbb{E} \left[\frac{X}{p^2} + \frac{1 - X}{(1 - p)^2} \right] = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{pq}.$$

$$I(p) = \frac{1}{pq}, \quad q = 1 - p.$$

Step 2. Compute the variance of the estimator $\hat{p} = \bar{X}$ and compare it with the Cramér-Rao lower bound.

E.g. 1. $X \sim \text{Bernoulli}(p)$. Check whether $\hat{p} = \bar{X}$ is efficient?

Step 1. Compute Fisher's Information:

$$p_X(k; p) = p^k (1 - p)^{1-k}.$$

$$\ln p_X(k; p) = k \ln p + (1 - k) \ln(1 - p)$$

$$\frac{\partial}{\partial p} \ln p_X(k; p) = \frac{k}{p} - \frac{1 - k}{1 - p}$$

$$-\frac{\partial^2}{\partial^2 p} \ln p_X(k; p) = \frac{k}{p^2} + \frac{1 - k}{(1 - p)^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 p} \ln p_X(X; p) \right] = \mathbb{E} \left[\frac{X}{p^2} + \frac{1 - X}{(1 - p)^2} \right] = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{pq}.$$

$$I(p) = \frac{1}{pq}, \quad q = 1 - p.$$

Step 2. Compute the variance of $\hat{p} = \bar{X}$ and compare it with $1/I(p)$.

E.g. 1. $X \sim \text{Bernoulli}(p)$. Check whether $\hat{p} = \bar{X}$ is efficient?

Step 1. Compute Fisher's Information:

$$p_X(k; p) = p^k (1 - p)^{1-k}.$$

$$\ln p_X(k; p) = k \ln p + (1 - k) \ln(1 - p)$$

$$\frac{\partial}{\partial p} \ln p_X(k; p) = \frac{k}{p} - \frac{1 - k}{1 - p}$$

$$-\frac{\partial^2}{\partial^2 p} \ln p_X(k; p) = \frac{k}{p^2} + \frac{1 - k}{(1 - p)^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 p} \ln p_X(X; p) \right] = \mathbb{E} \left[\frac{X}{p^2} + \frac{1 - X}{(1 - p)^2} \right] = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{pq}.$$

$$I(p) = \frac{1}{pq}, \quad q = 1 - p.$$

Efficiency: $\text{Var}(\hat{p}) = \frac{1}{n} \frac{1}{pq} = \frac{1}{n} \frac{1}{p(1-p)}$

E.g. 1. $X \sim \text{Bernoulli}(p)$. Check whether $\hat{p} = \bar{X}$ is efficient?

Step 1. Compute Fisher's Information:

$$p_X(k; p) = p^k (1 - p)^{1-k}.$$

$$\ln p_X(k; p) = k \ln p + (1 - k) \ln(1 - p)$$

$$\frac{\partial}{\partial p} \ln p_X(k; p) = \frac{k}{p} - \frac{1 - k}{1 - p}$$

$$-\frac{\partial^2}{\partial^2 p} \ln p_X(k; p) = \frac{k}{p^2} + \frac{1 - k}{(1 - p)^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 p} \ln p_X(X; p) \right] = \mathbb{E} \left[\frac{X}{p^2} + \frac{1 - X}{(1 - p)^2} \right] = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{pq}.$$

$$I(p) = \frac{1}{pq}, \quad q = 1 - p.$$

E.g. 1. $X \sim \text{Bernoulli}(p)$. Check whether $\hat{p} = \bar{X}$ is efficient?

Step 1. Compute Fisher's Information:

$$p_X(k; p) = p^k (1 - p)^{1-k}.$$

$$\ln p_X(k; p) = k \ln p + (1 - k) \ln(1 - p)$$

$$\frac{\partial}{\partial p} \ln p_X(k; p) = \frac{k}{p} - \frac{1 - k}{1 - p}$$

$$-\frac{\partial^2}{\partial^2 p} \ln p_X(k; p) = \frac{k}{p^2} + \frac{1 - k}{(1 - p)^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 p} \ln p_X(X; p) \right] = \mathbb{E} \left[\frac{X}{p^2} + \frac{1 - X}{(1 - p)^2} \right] = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{pq}.$$

$$I(p) = \frac{1}{pq}, \quad q = 1 - p.$$

Step 2. Compute $\text{Var}(\hat{p})$.

$$\text{Var}(\hat{p}) = \frac{1}{n^2} \text{Var} \left(\sum_{i=1}^n X_i \right) = \frac{1}{n^2} npq = \frac{pq}{n}$$

Conclusion Because \hat{p} is unbiased and $\text{Var}(\hat{p}) = (nI(p))^{-1}$, \hat{p} is efficient.

E.g. 1. $X \sim \text{Bernoulli}(p)$. Check whether $\hat{p} = \bar{X}$ is efficient?

Step 1. Compute Fisher's Information:

$$p_X(k; p) = p^k (1 - p)^{1-k}.$$

$$\ln p_X(k; p) = k \ln p + (1 - k) \ln(1 - p)$$

$$\frac{\partial}{\partial p} \ln p_X(k; p) = \frac{k}{p} - \frac{1 - k}{1 - p}$$

$$-\frac{\partial^2}{\partial^2 p} \ln p_X(k; p) = \frac{k}{p^2} + \frac{1 - k}{(1 - p)^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 p} \ln p_X(X; p) \right] = \mathbb{E} \left[\frac{X}{p^2} + \frac{1 - X}{(1 - p)^2} \right] = \frac{1}{p} + \frac{1}{1 - p} = \frac{1}{pq}.$$

$$I(p) = \frac{1}{pq}, \quad q = 1 - p.$$

Step 2. Compute $\text{Var}(\hat{p})$.

$$\text{Var}(\hat{p}) = \frac{1}{n^2} \text{Var} \left(\sum_{i=1}^n X_i \right) = \frac{1}{n^2} npq = \frac{pq}{n}$$

Conclusion Because \hat{p} is unbiased and $\text{Var}(\hat{p}) = (nI(p))^{-1}$, \hat{p} is efficient.

E.g. 2. Exponential distr.: $f_Y(y; \lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. Is $\hat{\lambda} = 1/\bar{Y}$ efficient?

Answer: No, because $\hat{\lambda}$ is biased. Nevertheless, we can still compute Fisher's Information as follows

Fisher's Inf.

$$\ln f_Y(y; \lambda) = \ln \lambda - \lambda y$$

Try: $\lambda = 1/\bar{Y}$. It is unbiased. Is it efficient?

E.g. 2. Exponential distr.: $f_Y(y; \lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. Is $\hat{\lambda} = 1/\bar{Y}$ efficient?

Answer No, because $\hat{\lambda}$ is biased. Nevertheless, we can still compute Fisher's Information as follows

Fisher's Inf.

$$\ln f_Y(y; \lambda) = \ln \lambda - \lambda y$$

Try: $\lambda = 1/\bar{Y}$. It is unbiased. Is it efficient?

E.g. 2. Exponential distr.: $f_Y(y; \lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. Is $\hat{\lambda} = 1/\bar{Y}$ efficient?

Answer No, because $\hat{\lambda}$ is biased. Nevertheless, we can still compute Fisher's Information as follows

Fisher's Inf.

$$\ln f_Y(y; \lambda) = \ln \lambda - \lambda y$$

$$\frac{\partial}{\partial \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda} - y$$

$$-\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(Y; \lambda) \right] = \mathbb{E} \left[\frac{1}{\lambda^2} \right] = \frac{1}{\lambda^2}.$$

$$I(\lambda) = \lambda^{-2}$$

Try: $\lambda = 1/\bar{Y}$. It is unbiased. Is it efficient?

E.g. 2. Exponential distr.: $f_Y(y; \lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. Is $\hat{\lambda} = 1/\bar{Y}$ efficient?

Answer No, because $\hat{\lambda}$ is biased. Nevertheless, we can still compute Fisher's Information as follows

Fisher's Inf.

$$\ln f_Y(y; \lambda) = \ln \lambda - \lambda y$$

$$\frac{\partial}{\partial \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda} - y$$

$$-\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(Y; \lambda) \right] = \mathbb{E} \left[\frac{1}{\lambda^2} \right] = \frac{1}{\lambda^2}.$$

$$I(\lambda) = \lambda^{-2}$$

Try: $\lambda = 1/\bar{Y}$. It is unbiased. Is it efficient?

E.g. 2. Exponential distr.: $f_Y(y; \lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. Is $\hat{\lambda} = 1/\bar{Y}$ efficient?

Answer No, because $\hat{\lambda}$ is biased. Nevertheless, we can still compute Fisher's Information as follows

Fisher's Inf.

$$\ln f_Y(y; \lambda) = \ln \lambda - \lambda y$$

$$\frac{\partial}{\partial \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda} - y$$

$$-\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(Y; \lambda) \right] = \mathbb{E} \left[\frac{1}{\lambda^2} \right] = \frac{1}{\lambda^2}.$$

$$I(\lambda) = \lambda^{-2}$$

Try: $\lambda = 1/\bar{Y}$. It is unbiased. Is it efficient?

E.g. 2. Exponential distr.: $f_Y(y; \lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. Is $\hat{\lambda} = 1/\bar{Y}$ efficient?

Answer No, because $\hat{\lambda}$ is biased. Nevertheless, we can still compute Fisher's Information as follows

Fisher's Inf.

$$\ln f_Y(y; \lambda) = \ln \lambda - \lambda y$$

$$\frac{\partial}{\partial \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda} - y$$

$$-\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(Y; \lambda) \right] = \mathbb{E} \left[\frac{1}{\lambda^2} \right] = \frac{1}{\lambda^2}.$$

$$I(\lambda) = \lambda^{-2}$$

Try: $\lambda = 1/\bar{Y}$. It is unbiased. Is it efficient?

E.g. 2. Exponential distr.: $f_Y(y; \lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. Is $\hat{\lambda} = 1/\bar{Y}$ efficient?

Answer No, because $\hat{\lambda}$ is biased. Nevertheless, we can still compute Fisher's Information as follows

Fisher's Inf.

$$\ln f_Y(y; \lambda) = \ln \lambda - \lambda y$$

$$\frac{\partial}{\partial \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda} - y$$

$$-\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(Y; \lambda) \right] = \mathbb{E} \left[\frac{1}{\lambda^2} \right] = \frac{1}{\lambda^2}.$$

$$I(\lambda) = \lambda^{-2}$$

Try: $\lambda = 1/\bar{Y}$. It is unbiased. Is it efficient?

E.g. 2. Exponential distr.: $f_Y(y; \lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. Is $\hat{\lambda} = 1/\bar{Y}$ efficient?

Answer No, because $\hat{\lambda}$ is biased. Nevertheless, we can still compute Fisher's Information as follows

Fisher's Inf.

$$\ln f_Y(y; \lambda) = \ln \lambda - \lambda y$$

$$\frac{\partial}{\partial \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda} - y$$

$$-\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(y; \lambda) = \frac{1}{\lambda^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \lambda} \ln f_Y(Y; \lambda) \right] = \mathbb{E} \left[\frac{1}{\lambda^2} \right] = \frac{1}{\lambda^2}.$$

$$I(\lambda) = \lambda^{-2}$$

Try: $\hat{\lambda}^* := \frac{n-1}{n} \frac{1}{\bar{Y}}$. It is unbiased. Is it efficient?

E.g. 2'. Exponential distr.: $f_Y(y; \theta) = \theta^{-1} e^{-y/\theta}$ for $y \geq 0$. $\hat{\theta} = \bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = -\ln \theta - \frac{y}{\theta}$$

E.g. 2'. Exponential distr.: $f_Y(y; \theta) = \theta^{-1} e^{-y/\theta}$ for $y \geq 0$. $\hat{\theta} = \bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = -\ln \theta - \frac{y}{\theta}$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta} + \frac{y}{\theta^2}$$

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta^2} + \frac{2y}{\theta^3}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{1}{\theta^2} + \frac{2Y}{\theta^3} \right] = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3} = \theta^{-2}.$$

$$I(\theta) = \theta^{-2}$$

E.g. 2'. Exponential distr.: $f_Y(y; \theta) = \theta^{-1} e^{-y/\theta}$ for $y \geq 0$. $\hat{\theta} = \bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = -\ln \theta - \frac{y}{\theta}$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta} + \frac{y}{\theta^2}$$

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta^2} + \frac{2y}{\theta^3}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{1}{\theta^2} + \frac{2Y}{\theta^3} \right] = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3} = \theta^{-2}.$$

$$I(\theta) = \theta^{-2}$$

E.g. 2'. Exponential distr.: $f_Y(y; \theta) = \theta^{-1} e^{-y/\theta}$ for $y \geq 0$. $\hat{\theta} = \bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = -\ln \theta - \frac{y}{\theta}$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta} + \frac{y}{\theta^2}$$

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta^2} + \frac{2y}{\theta^3}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{1}{\theta^2} + \frac{2Y}{\theta^3} \right] = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3} = \theta^{-2}.$$

$$I(\theta) = \theta^{-2}$$

E.g. 2'. Exponential distr.: $f_Y(y; \theta) = \theta^{-1} e^{-y/\theta}$ for $y \geq 0$. $\hat{\theta} = \bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = -\ln \theta - \frac{y}{\theta}$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta} + \frac{y}{\theta^2}$$

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta^2} + \frac{2y}{\theta^3}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{1}{\theta^2} + \frac{2Y}{\theta^3} \right] = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3} = \theta^{-2}.$$

$$I(\theta) = \theta^{-2}$$

E.g. 2'. Exponential distr.: $f_Y(y; \theta) = \theta^{-1} e^{-y/\theta}$ for $y \geq 0$. $\hat{\theta} = \bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = -\ln \theta - \frac{y}{\theta}$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta} + \frac{y}{\theta^2}$$

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta^2} + \frac{2y}{\theta^3}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{1}{\theta^2} + \frac{2Y}{\theta^3} \right] = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3} = \theta^{-2}.$$

$$I(\theta) = \theta^{-2}$$

E.g. 2'. Exponential distr.: $f_Y(y; \theta) = \theta^{-1} e^{-y/\theta}$ for $y \geq 0$. $\hat{\theta} = \bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = -\ln \theta - \frac{y}{\theta}$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta} + \frac{y}{\theta^2}$$

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta^2} + \frac{2y}{\theta^3}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{1}{\theta^2} + \frac{2Y}{\theta^3} \right] = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3} = \theta^{-2}.$$

$$\boxed{I(\theta) = \theta^{-2}}$$

Step 2. Compute $\text{Var}(\hat{\theta})$:

$$\text{Var}(\bar{Y}) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(Y_i) = \frac{1}{n^2} n\theta^2 = \frac{\theta^2}{n}.$$

Conclusion. Because $\hat{\theta}$ is unbiased and $\text{Var}(\hat{\theta}) = (nI(\theta))^{-1}$, $\hat{\theta}$ is efficient.

E.g. 2'. Exponential distr.: $f_Y(y; \theta) = \theta^{-1} e^{-y/\theta}$ for $y \geq 0$. $\hat{\theta} = \bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = -\ln \theta - \frac{y}{\theta}$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta} + \frac{y}{\theta^2}$$

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{1}{\theta^2} + \frac{2y}{\theta^3}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{1}{\theta^2} + \frac{2Y}{\theta^3} \right] = -\frac{1}{\theta^2} + \frac{2\theta}{\theta^3} = \theta^{-2}.$$

$$\boxed{I(\theta) = \theta^{-2}}$$

Step 2. Compute $\text{Var}(\hat{\theta})$:

$$\text{Var}(\bar{Y}) = \frac{1}{n^2} \sum_{i=1}^n \text{Var}(Y_i) = \frac{1}{n^2} n\theta^2 = \frac{\theta^2}{n}.$$

Conclusion. Because $\hat{\theta}$ is unbiased and $\text{Var}(\hat{\theta}) = (nI(\theta))^{-1}$, $\hat{\theta}$ is efficient.

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

Example 3: Fisher's Information

Example 3: Fisher's Information

Example 3: Fisher's Information

Example 3: Fisher's Information

Example 3: Fisher's Information

Example 3: Fisher's Information

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

Why is this? The answer is that the regularity conditions for the Cramér-Rao inequality are not satisfied. In particular, the support of the distribution depends on the parameter θ . The regularity conditions require that the support of the distribution does not depend on the parameter θ . In this case, the support is $[0, \theta]$, which depends on θ . Therefore, the Cramér-Rao inequality does not apply, and the estimator $\hat{\theta} = \frac{3}{2}\bar{Y}$ is not efficient.

Exercise 3.1.10. Show that $\hat{\theta} = \frac{3}{2}\bar{Y}$ is efficient.

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

Why is this? The answer is that the regularity conditions for the Cramér-Rao inequality are not satisfied. In particular, the support of the distribution depends on the parameter θ . The regularity conditions require that the support of the distribution does not depend on the parameter θ . In this case, the support is $[0, \theta]$, which depends on θ . Therefore, the Cramér-Rao inequality does not apply, and the estimator $\hat{\theta} = \frac{3}{2}\bar{Y}$ is not efficient.

Exercise 3.1.10. Show that $\hat{\theta} = \frac{3}{2}\bar{Y}$ is efficient.

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

Why is this? The answer is that the support of the distribution depends on the parameter θ . In this case, the support is $[0, \theta]$. The usual derivation of the Fisher information formula assumes that the support of the distribution does not depend on the parameter. This is not the case here, so the formula does not apply.

Exercise: Show that $\hat{\theta} = \frac{3}{2}\bar{Y}$ is efficient.

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

Why is this? The answer is that the regularity conditions for the Fisher information formula are not satisfied. In particular, the support of the distribution depends on the parameter θ . This is why the formula for Fisher information based on the second derivative of the log-likelihood function is not applicable in this case.

Exercise 3.1.10. Show that $\hat{\theta} = \frac{3}{2}\bar{Y}$ is efficient.

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

Step 2. Compute $\text{Var}(\hat{\theta})$:

$$\text{Var}(\hat{\theta}) = \frac{9}{4n} \text{Var}(Y) = \frac{9}{4n} \frac{\theta^2}{18} = \frac{\theta^2}{8n}.$$

Discussion. Even though $\hat{\theta}$ is unbiased, we have two discrepancies: (\dagger) and

$$\text{Var}(\hat{\theta}) = \frac{\theta^2}{8n} \leq \frac{\theta^2}{4n} = \frac{1}{nI(\theta)}$$

⇒ $\hat{\theta}$ is not efficient (it is not UMVUE).

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

Step 2. Compute $\text{Var}(\hat{\theta})$:

$$\text{Var}(\hat{\theta}) = \frac{9}{4n} \text{Var}(Y) = \frac{9}{4n} \frac{\theta^2}{18} = \frac{\theta^2}{8n}.$$

Discussion. Even though $\hat{\theta}$ is unbiased, we have two discrepancies: (\dagger) and

$$\text{Var}(\hat{\theta}) = \frac{\theta^2}{8n} \leq \frac{\theta^2}{4n} = \frac{1}{nI(\theta)}$$

This is because this is not a regular estimation!

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

Step 2. Compute $\text{Var}(\hat{\theta})$:

$$\text{Var}(\hat{\theta}) = \frac{9}{4n} \text{Var}(Y) = \frac{9}{4n} \frac{\theta^2}{18} = \frac{\theta^2}{8n}.$$

Discussion. Even though $\hat{\theta}$ is unbiased, we have two discrepancies: (\dagger) and

$$\text{Var}(\hat{\theta}) = \frac{\theta^2}{8n} \leq \frac{\theta^2}{4n} = \frac{1}{nI(\theta)}$$

This is because this is not a regular estimation!

E.g. 3. $f_Y(y; \theta) = 2y/\theta^2$ for $y \in [0, \theta]$. $\hat{\theta} = \frac{3}{2}\bar{Y}$ efficient?

Step. 1. Compute Fisher's Information:

$$\ln f_Y(y; \theta) = \ln(2y) - 2 \ln \theta$$

$$\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta}$$

By the definition of Fisher's information,

$$I(\theta) = \mathbb{E} \left[\left(\frac{\partial}{\partial \theta} \ln f_Y(y; \theta) \right)^2 \right] = \mathbb{E} \left[\left(-\frac{2}{\theta} \right)^2 \right] = \frac{4}{\theta^2}.$$

However, if we compute

$$-\frac{\partial^2}{\partial^2 \theta} \ln f_Y(y; \theta) = -\frac{2}{\theta^2}$$

$$-\mathbb{E} \left[\frac{\partial^2}{\partial^2 \theta} \ln f_Y(Y; \theta) \right] = \mathbb{E} \left[-\frac{2}{\theta^2} \right] = -\frac{2}{\theta^2} \neq \frac{4}{\theta^2} = I(\theta). \quad (\dagger)$$

Step 2. Compute $\text{Var}(\hat{\theta})$:

$$\text{Var}(\hat{\theta}) = \frac{9}{4n} \text{Var}(Y) = \frac{9}{4n} \frac{\theta^2}{18} = \frac{\theta^2}{8n}.$$

Discussion. Even though $\hat{\theta}$ is unbiased, we have two discrepancies: (\dagger) and

$$\text{Var}(\hat{\theta}) = \frac{\theta^2}{8n} \leq \frac{\theta^2}{4n} = \frac{1}{nI(\theta)}$$

This is because this is not a regular estimation!