Math 362: Mathematical Statistics II

Le Chen
le.chen@emory.edu

Emory University
Atlanta, GA

Last updated on April 13, 2021

2021 Spring

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

Chapter 5. Estimation

§ 5.4 Properties of Estimators

Question: Estimators are not in general unique (MLE or MME ...). How to select one estimator?

Recall: For a random sample of size n from the population with given pdf, we have $X_1, \cdots, X_n$, which are i.i.d. r.v.'s. The estimator $\hat{\theta}$ is a function of $X_i's$:

$$\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n).$$

Criterions:

1. Unbiased. (Mean)
2. Efficiency, the minimum-variance estimator. (Variance)
3. Sufficency.
4. Consistency. (Asymptotic behavior)

§ 5.4 Properties of Estimators

Question: Estimators are not in general unique (MLE or MME ...). How to select one estimator?

Recall: For a random sample of size n from the population with given pdf, we have $X_1, \cdots, X_n$, which are i.i.d. r.v.'s. The estimator $\hat{\theta}$ is a function of $X_i's$:

$$\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n).$$

Criterions:

1. Unbiased. (Mean)
2. Efficiency, the minimum-variance estimator. (Variance)
3. Sufficency.
4. Consistency. (Asymptotic behavior)

§ 5.4 Properties of Estimators

Question: Estimators are not in general unique (MLE or MME ...). How to select one estimator?

Recall: For a random sample of size n from the population with given pdf, we have $X_1, \cdots, X_n$, which are i.i.d. r.v.'s. The estimator $\hat{\theta}$ is a function of $X_i's$:

$$\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n).$$

Criterions:

1. Unbiased. (Mean)
2. Efficiency, the minimum-variance estimator. (Variance)
3. Sufficency.
4. Consistency. (Asymptotic behavior)

§ 5.4 Properties of Estimators

Question: Estimators are not in general unique (MLE or MME ...). How to select one estimator?

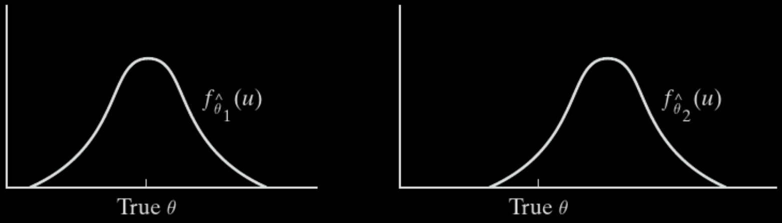
Recall: For a random sample of size n from the population with given pdf, we have $X_1, \cdots, X_n$, which are i.i.d. r.v.'s. The estimator $\hat{\theta}$ is a function of $X_i's$:

$$\hat{\theta} = \hat{\theta}(X_1, \cdots, X_n).$$

Criterions:

1. Unbiased. (Mean)
2. Efficiency, the minimum-variance estimator. (Variance)
3. Sufficency.
4. Consistency. (Asymptotic behavior)

Unbiasedness



Definition 5.4.1. Given a random sample of size n whose population distribution dependes on an unknown parameter θ, let $\hat{\theta}$ be an estimator of θ.

Then $\hat{\theta}$ is called **unbiased** if $\mathbb{E}(\hat{\theta}) = \theta$;

and $\hat{\theta}$ is called **asymptotically unbiased** if $\lim_{n\to\infty} \mathbb{E}(\hat{\theta}) = \theta$.

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_2 = Y_{max}$.
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$. Show that for any constants a_i's,

$$\hat{\theta} = \sum_{i=1}^{n} a_i X_i \quad \text{is unbiased} \quad \Longleftrightarrow \quad \sum_{i=1}^{n} a_i = 1.$$

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_2 = Y_{max}$.
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$. Show that for any constants a_i's,

$$\hat{\theta} = \sum_{i=1}^{n} a_i X_i \quad \text{is unbiased} \quad \Longleftrightarrow \quad \sum_{i=1}^{n} a_i = 1.$$

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_2 = Y_{max}$.
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$. Show that for any constants a_i's,

$$\hat{\theta} = \sum_{i=1}^{n} a_i X_i \quad \text{is unbiased} \quad \Longleftrightarrow \quad \sum_{i=1}^{n} a_i = 1.$$

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_2 = Y_{max}$.
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$. Show that for any constants a_i's,

$$\hat{\theta} = \sum_{i=1}^{n} a_i X_i \quad \text{is unbiased} \quad \Longleftrightarrow \quad \sum_{i=1}^{n} a_i = 1.$$

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$.

- $\hat{\theta}_1 = \dfrac{3}{2}\overline{Y}$
- $\hat{\theta}_2 = Y_{max}$.
- $\hat{\theta}_3 = \dfrac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$. Show that for any constants a_i's,

$$\hat{\theta} = \sum_{i=1}^{n} a_i X_i \quad \text{is unbiased} \quad \Longleftrightarrow \quad \sum_{i=1}^{n} a_i = 1.$$

E.g. 3. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\sigma^2 = \mathrm{Var}(X)$.

- $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

- $S^2 = \text{Sample Variance} = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

- $S = \text{Sample Standard Deviation} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}.$ (Biased for σ!)

E.g. 3. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\sigma^2 = \mathrm{Var}(X)$.

- $\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2$

- $S^2 = \text{Sample Variance} = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2$

- $S = \text{Sample Standard Deviation} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2}.$ (Biased for σ!)

E.g. 3. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\sigma^2 = \mathrm{Var}(X)$.

- $\widehat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2$

- $S^2 = \text{Sample Variance} = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2$

- $S = \text{Sample Standard Deviation} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \overline{X} \right)^2}$. (Biased for σ!)

E.g. 3. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\sigma^2 = \mathrm{Var}(X)$.

- $\widehat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

- $S^2 = \text{Sample Variance} = \frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2$

- $S = \text{Sample Standard Deviation} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i - \overline{X}\right)^2}.$ (Biased for σ!)

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^{n} Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\begin{aligned}
\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) &= n \int_0^\infty \frac{1}{y} \frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y \\
&= \frac{n\lambda}{n-1} \int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y \\
&= \frac{n}{n-1}\lambda.
\end{aligned}$$

Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(Y_i) = \frac{1}{n}\sum_{i=1}^{n} \theta = \theta.$$

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^{n} Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\begin{aligned}
\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) &= n \int_0^\infty \frac{1}{y} \frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y \\
&= \frac{n\lambda}{n-1} \int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y \\
&= \frac{n}{n-1}\lambda.
\end{aligned}$$

Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(Y_i) = \frac{1}{n}\sum_{i=1}^{n} \theta = \theta.$$

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^{n} Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\begin{aligned}
\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) &= n \int_0^\infty \frac{1}{y} \frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y \\
&= \frac{n\lambda}{n-1} \int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y \\
&= \frac{n}{n-1}\lambda.
\end{aligned}$$

Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(Y_i) = \frac{1}{n}\sum_{i=1}^{n} \theta = \theta.$$

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^{n} Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) = n\int_0^\infty \frac{1}{y}\frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y$$
$$= \frac{n\lambda}{n-1}\int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y$$
$$= \frac{n}{n-1}\lambda.$$

Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(Y_i) = \frac{1}{n}\sum_{i=1}^{n} \theta = \theta.$$

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^{n} Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) = n\int_0^\infty \frac{1}{y}\frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y$$

$$= \frac{n\lambda}{n-1}\int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y$$

$$= \frac{n}{n-1}\lambda.$$

Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(Y_i) = \frac{1}{n}\sum_{i=1}^{n} \theta = \theta.$$

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^{n} Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) = n\int_0^{\infty} \frac{1}{y}\frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y$$

$$= \frac{n\lambda}{n-1}\int_0^{\infty} \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y$$

$$= \frac{n}{n-1}\lambda.$$

Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(Y_i) = \frac{1}{n}\sum_{i=1}^{n} \theta = \theta$$

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^{n} Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\begin{aligned}
\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) &= n \int_0^\infty \frac{1}{y} \frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y \\
&= \frac{n\lambda}{n-1} \int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y \\
&= \frac{n}{n-1}\lambda.
\end{aligned}$$

Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{E}(Y_i) = \frac{1}{n} \sum_{i=1}^{n} \theta = \theta$$

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^{n} Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) = n\int_0^\infty \frac{1}{y}\frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y$$
$$= \frac{n\lambda}{n-1}\int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y$$
$$= \frac{n}{n-1}\lambda.$$

Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n}\mathbb{E}(Y_i) = \frac{1}{n}\sum_{i=1}^{n}\theta = \theta.$$

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^n Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\begin{aligned}
\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) &= n\int_0^\infty \frac{1}{y}\frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y \\
&= \frac{n\lambda}{n-1}\int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y \\
&= \frac{n}{n-1}\lambda.
\end{aligned}$$

Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n}\sum_{i=1}^n \mathbb{E}(Y_i) = \frac{1}{n}\sum_{i=1}^n \theta = \theta.$$

E.g. 4. Exponential distr.: $f_Y(y;\lambda) = \lambda e^{-\lambda y}$ for $y \geq 0$. $\widehat{\lambda} = 1/\overline{Y}$ is biased.

$n\overline{Y} = \sum_{i=1}^{n} Y_i \sim$ Gamma distribution(n, λ). Hence,

$$\mathbb{E}\left(\widehat{\lambda}\right) = \mathbb{E}\left(1/\overline{Y}\right) = n\int_0^\infty \frac{1}{y}\frac{\lambda^n}{\Gamma(n)} y^{n-1} e^{-\lambda y} \mathrm{d}y$$
$$= \frac{n\lambda}{n-1}\int_0^\infty \underbrace{\frac{\lambda^{n-1}}{\Gamma(n-1)} y^{(n-1)-1} e^{-\lambda y}}_{\text{pdf for Gamma distr. } (n-1,\lambda)} \mathrm{d}y$$
$$= \frac{n}{n-1}\lambda.$$

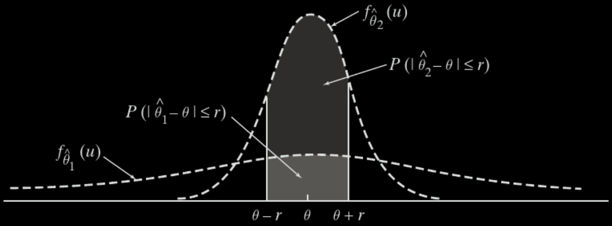
Biased! But $\mathbb{E}(\widehat{\lambda}) = \frac{n}{n-1}\lambda \to \lambda$ as $n \to \infty$. (Asymptotically unbiased.)

Note: $\widehat{\lambda}^* = \frac{n-1}{n\overline{Y}}$ is unbiased.

E.g. 4'. Exponential distr.: $f_Y(y;\theta) = \frac{1}{\theta} e^{-y/\theta}$ for $y \geq 0$. $\widehat{\theta} = \overline{Y}$ is unbiased.

$$\mathbb{E}\left(\widehat{\theta}\right) = \frac{1}{n}\sum_{i=1}^{n} \mathbb{E}(Y_i) = \frac{1}{n}\sum_{i=1}^{n} \theta = \theta.$$

Efficiency



Definition 5.4.2. Let $\widehat{\theta}_1$ and $\widehat{\theta}_2$ be two unbiased estimators for a parameter θ. If $\text{Var}(\widehat{\theta}_1) < \text{Var}(\widehat{\theta}_2)$, then we say that $\widehat{\theta}_1$ is **more efficient** than $\widehat{\theta}_2$. The **relative efficiency** of $\widehat{\theta}_1$ w.r.t. $\widehat{\theta}_2$ is the ratio $\text{Var}(\widehat{\theta}_1)/\text{Var}(\widehat{\theta}_2)$.

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$. Which is more efficient? Find the relative efficiency of $\widehat{\theta}_1$ w.r.t. $\widehat{\theta}_3$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$ (suppose $\sigma^2 = \mathrm{Var}(X) < \infty$).

Among all possible unbiased estimators $\hat{\theta} = \sum_{i=1}^n a_i X_i$ with $\sum_{i=1}^n a_i = 1$. Find the most efficient one.

Sol:

$$\mathrm{Var}(\hat{\theta}) = \sum_{i=1}^n a_i^2 \mathrm{Var}(X) = \sigma^2 \sum_{i=1}^n a_i^2 \geq \sigma^2 \frac{1}{n}\left(\sum_{i=1}^n a_i\right)^2 = \frac{1}{n}\sigma^2.$$

with equality iff $a_1 = \cdots = a_n = 1/n$.

Hence, the most efficient one is the sample mean $\hat{\theta} = \overline{X}$. □

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$. Which is more efficient? Find the relative efficiency of $\widehat{\theta}_1$ w.r.t. $\widehat{\theta}_3$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$ (suppose $\sigma^2 = \mathrm{Var}(X) < \infty$).

Among all possible unbiased estimators $\hat{\theta} = \sum_{i=1}^n a_i X_i$ with $\sum_{i=1}^n a_i = 1$. Find the most efficient one.

Sol:

$$\mathrm{Var}(\widehat{\theta}) = \sum_{i=1}^n a_i^2 \mathrm{Var}(X) = \sigma^2 \sum_{i=1}^n a_i^2 \geq \sigma^2 \frac{1}{n}\left(\sum_{i=1}^n a_i\right)^2 = \frac{1}{n}\sigma^2,$$

with equality iff $a_1 = \cdots = a_n = 1/n$.

Hence, the most efficient one is the sample mean $\widehat{\theta} = \overline{X}$. □

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$. Which is more efficient? Find the relative efficiency of $\widehat{\theta}_1$ w.r.t. $\widehat{\theta}_3$.

- $\hat{\theta}_1 = \dfrac{3}{2}\overline{Y}$
- $\hat{\theta}_3 = \dfrac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$ (suppose $\sigma^2 = \mathrm{Var}(X) < \infty$).

Among all possible unbiased estimators $\hat{\theta} = \sum_{i=1}^n a_i X_i$ with $\sum_{i=1}^n a_i = 1$. Find the most efficient one.

Sol:

$$\mathrm{Var}(\widehat{\theta}) = \sum_{i=1}^n a_i^2 \mathrm{Var}(X) = \sigma^2 \sum_{i=1}^n a_i^2 \geq \sigma^2 \frac{1}{n}\left(\sum_{i=1}^n a_i\right)^2 = \frac{1}{n}\sigma^2,$$

with equality iff $a_1 = \cdots = a_n = 1/n$.

Hence, the most efficient one is the sample mean $\widehat{\theta} = \overline{X}$. □

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0, \theta]$. Which is more efficient? Find the relative efficiency of $\widehat{\theta}_1$ w.r.t. $\widehat{\theta}_3$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$ (suppose $\sigma^2 = \mathrm{Var}(X) < \infty$).

Among all possible unbiased estimators $\hat{\theta} = \sum_{i=1}^n a_i X_i$ with $\sum_{i=1}^n a_i = 1$. Find the most efficient one.

Sol:

$$\mathrm{Var}(\widehat{\theta}) = \sum_{i=1}^n a_i^2 \mathrm{Var}(X) = \sigma^2 \sum_{i=1}^n a_i^2 \geq \sigma^2 \frac{1}{n} \left(\sum_{i=1}^n a_i \right)^2 = \frac{1}{n}\sigma^2,$$

with equality iff $a_1 = \cdots = a_n = 1/n$.

Hence, the most efficient one is the sample mean $\widehat{\theta} = \overline{X}$. □

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$. Which is more efficient? Find the relative efficiency of $\widehat{\theta}_1$ w.r.t. $\widehat{\theta}_3$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$ (suppose $\sigma^2 = \mathrm{Var}(X) < \infty$).

Among all possible unbiased estimators $\hat{\theta} = \sum_{i=1}^n a_i X_i$ with $\sum_{i=1}^n a_i = 1$. Find the most efficient one.

Sol:

$$\mathrm{Var}(\widehat{\theta}) = \sum_{i=1}^n a_i^2 \mathrm{Var}(X) = \sigma^2 \sum_{i=1}^n a_i^2 \geq \sigma^2 \frac{1}{n}\left(\sum_{i=1}^n a_i\right)^2 = \frac{1}{n}\sigma^2,$$

with equality iff $a_1 = \cdots = a_n = 1/n$.

Hence, the most efficient one is the sample mean $\widehat{\theta} = \overline{X}$. □

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$. Which is more efficient? Find the relative efficiency of $\widehat{\theta}_1$ w.r.t. $\widehat{\theta}_3$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$ (suppose $\sigma^2 = \mathrm{Var}(X) < \infty$).

Among all possible unbiased estimators $\hat{\theta} = \sum_{i=1}^n a_i X_i$ with $\sum_{i=1}^n a_i = 1$. Find the most efficient one.

Sol:

$$\mathrm{Var}(\widehat{\theta}) = \sum_{i=1}^n a_i^2 \mathrm{Var}(X) = \sigma^2 \sum_{i=1}^n a_i^2 \ge \sigma^2 \frac{1}{n}\left(\sum_{i=1}^n a_i\right)^2 = \frac{1}{n}\sigma^2,$$

with equality iff $a_1 = \cdots = a_n = 1/n$.

Hence, the most efficient one is the sample mean $\widehat{\theta} = \overline{X}$. □

E.g. 1. $f_Y(y;\theta) = \frac{2y}{\theta^2}$ if $y \in [0,\theta]$. Which is more efficient? Find the relative efficiency of $\widehat{\theta}_1$ w.r.t. $\widehat{\theta}_3$.

- $\hat{\theta}_1 = \frac{3}{2}\overline{Y}$
- $\hat{\theta}_3 = \frac{2n+1}{2n} Y_{max}$.

E.g. 2. Let $X_1, \cdots, X_n$ be a random sample of size n with the unknown parameter $\theta = \mathbb{E}(X)$ (suppose $\sigma^2 = \mathrm{Var}(X) < \infty$).

Among all possible unbiased estimators $\hat{\theta} = \sum_{i=1}^n a_i X_i$ with $\sum_{i=1}^n a_i = 1$. Find the most efficient one.

Sol:

$$\mathrm{Var}(\widehat{\theta}) = \sum_{i=1}^n a_i^2 \mathrm{Var}(X) = \sigma^2 \sum_{i=1}^n a_i^2 \geq \sigma^2 \frac{1}{n}\left(\sum_{i=1}^n a_i\right)^2 = \frac{1}{n}\sigma^2,$$

with equality iff $a_1 = \cdots = a_n = 1/n$.

Hence, the most efficient one is the sample mean $\widehat{\theta} = \overline{X}$. □