M378K: April 4^{th}, 2025.

M378K Introduction to Mathematical Statistics
Problem Set #17
Relative efficiency.

Definition 17.1. *Given two unbiased estimators $\hat{\theta}_1$ and $\hat{\theta}_2$, the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ is defined as*

$$\boxed{\mathit{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{\mathrm{Var}[\hat{\theta}_2]}{\mathrm{Var}[\hat{\theta}_1]}.}$$

Problem 17.1. *Let Y_1, Y_2 be a random sample from the exponential distribution with the unknown parameter θ.*

$Y_i \sim E(\tau = \theta)$

(i) *The estimator $\hat{\theta}_1 = (Y_1 + Y_2)/2$ for θ is proposed. What is its variance?*

(ii) *The estimator $\boxed{\hat{\theta}_2 = cY_{(1)}}$ for θ is proposed. Find the constant c such that $\hat{\theta}_2$ is an unbiased estimator of θ. What is its variance?*

(iii) *Calculate $\boxed{\text{the efficiency of } \hat{\theta}_1 \text{ relative to } \hat{\theta}_2}$.*

i. $\mathrm{Var}[\hat{\theta}_1] = \mathrm{Var}\left[\frac{1}{2}(Y_1+Y_2)\right] = \frac{\mathrm{Var}[Y_1]}{2} = \frac{\theta^2}{2}$

ii. $Y_{(1)} \sim E\left(\frac{\theta}{2}\right)$

$\mathbb{E}[\hat{\theta}_2] = \theta \quad \Rightarrow \quad \boxed{c = 2}$

$\mathrm{Var}[\hat{\theta}_2] = \mathrm{Var}[2 \cdot Y_{(1)}] = 4 \cdot \mathrm{Var}[Y_{(1)}] = 4 \cdot \frac{\theta^2}{4} = \theta^2$

iii. $\mathit{eff}(\hat{\theta}_1, \hat{\theta}_2) = \frac{\theta^2}{\frac{\theta^2}{2}} = 2$

$\square$

M378K Introduction to Mathematical Statistics
Problem Set #18
Consistency.

Definition 18.1. *$\hat{\theta}_n$ is said to be a* consistent estimator *of θ if*

$$\hat{\theta}_n \to \theta \quad \text{in probability as } n \to \infty,$$

i.e., if for any $\varepsilon > 0$,

$$\lim_{n\to\infty} \boxed{\mathbb{P}\left[|\hat{\theta}_n - \theta| > \varepsilon\right]} = 0.$$

$\boxed{\text{Markov Inequality} \quad \text{Chebyshev Inequality}}$

Theorem 18.2. *Let $\hat{\theta}_n$ be* $\boxed{\text{unbiased}}$ *and such that*

$$\boxed{\text{Var}\left[\hat{\theta}_n\right] \xrightarrow{n\to\infty} 0.}$$

Then, $\hat{\theta}_n$ is a **consistent estimator**.

Problem 18.1. *Let $Y_1, Y_2, \dots, Y_n$ be a random sample from any distribution with finite first and second moments. Propose a consistent estimator for the population mean μ and* **prove** *that it is, indeed, consistent.*

→ : Propose: $\bar{Y}_n = \dfrac{Y_1 + \dots + Y_n}{n}$

Criteria for consistency ✓
- unbiased ? ✓
- $\text{Var}[\bar{Y}_n] \xrightarrow[n]{?} 0$

$$= \frac{\boxed{\text{Var}[Y_1]}}{n} \xrightarrow[n\to\infty]{} 0 \quad ✓$$

$$\text{Var}[Y_1] = \underbrace{\underbrace{\mathbb{E}[Y_1^2]}_{<\infty} - \underbrace{(\mathbb{E}[Y_1])^2}_{<\infty}}_{<\infty}$$

□

$\boxed{\text{Google}}$ Cauchy Dist'n

Problem 18.2. *Consider a random sample $Y_1, Y_2, \ldots, Y_n$ from a power distribution with the density of the form*

$$f_Y(y) = \theta y^{\theta-1}\mathbf{1}_{(0,1)}(y).$$

What is a consistent estimator for $\frac{\theta}{\theta+1}$? **Prove** *that your choice is indeed consistent.*

$$\mathbb{E}[Y_1] = \int_0^1 \theta y^{\theta-1} \cdot y \, dy = \theta \int_0^1 y^{\theta} dy = \theta \cdot \frac{y^{\theta+1}}{\theta+1}\Big|_{y=0}^{1}$$

$$\boxed{\mathbb{E}[Y_1] = \frac{\theta}{\theta+1}}$$

We propose $\bar{Y}_n$:

$$\boxed{\bar{Y}_n = \frac{1}{n}(Y_1 + \cdots + Y_n)}$$

- unbiased ✓
- $\mathrm{Var}[\bar{Y}_n] \xrightarrow{?} 0$ ✓

If $\mathrm{Var}[Y_1] < \infty$, we're done :)

$$\mathrm{Var}[Y_1] = \boxed{\mathbb{E}[Y_1^2]} - \underbrace{(\mathbb{E}[Y_1])^2}_{<\infty} \quad < \infty$$

$$\mathbb{E}[Y_1^2] = \int_0^1 y^2 \theta y^{\theta-1} dy = \theta \int_0^1 y^{\theta+1} dy = \frac{\theta}{\theta+2} < \infty$$

$\square$

Example. $Y_1, \dots, Y_n$ random sample from $E(\tau = \theta)$.

$$\hat{\theta}_n = \underline{n}\, \underbrace{Y_{(1)}}_{\sim E\left(\frac{\tau}{n}\right)} \quad \text{unbiased for all } n$$

Consider $Var[\hat{\theta}_n] = Var[n \cdot Y_{(1)}] =$

$$= n^2 \cdot Var[\underbrace{Y_{(1)}}_{\sim E\left(\frac{\tau}{n}\right)}] = n^2 \cdot \left(\frac{\theta}{n}\right)^2 = \theta^2 \cancel{\rightarrow} 0$$

The criterion is not satisfied.

We would still have to explore whether

$\hat{\theta}_n$ is consistent.

Task: Find an unbiased estimator for θ in

$$Y_1, \dots, Y_n \sim U(0, \theta)$$

Is it consistent?