# Properties of Estimators

## Efficiency

Bruce M. Boghosian

Tufts UNIVERSITY | School of Arts and Sciences

Department of Mathematics

Tufts University

# Outline

Bruce M. Boghosian

# Efficiency: A tale of two estimators



- Three iid random variables $X_1, X_2, X_3$
- Assume $E(X_j) = \mu$ and $\text{Var}(X_j) = \sigma^2$ for $j = 1, \ldots, 3$
- Consider the two estimators,

$$\hat{\mu}_1(\vec{X}) := \frac{1}{4}X_1 + \frac{1}{2}X_2 + \frac{1}{4}X_3$$
$$\hat{\mu}_2(\vec{X}) := \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$$

- Linear and unbiased since $E(\hat{\mu}_1) = E(\hat{\mu}_2) = \mu$, and

$$\text{Var}\left(\hat{\mu}_1(\vec{X})\right) := \frac{1}{16}\sigma^2 + \frac{1}{4}\sigma^2 + \frac{1}{16}\sigma^2 = \frac{3}{8}\sigma^2$$
$$\text{Var}\left(\hat{\mu}_2(\vec{X})\right) := \frac{1}{9}\sigma^2 + \frac{1}{9}\sigma^2 + \frac{1}{9}\sigma^2 = \frac{3}{9}\sigma^2$$

# The sequel: Relative efficiency



- We continue our analysis of the two unbiased estimators,

$$\hat{\mu}_1(\vec{X}) := \frac{1}{4}X_1 + \frac{1}{2}X_2 + \frac{1}{4}X_3$$
$$\hat{\mu}_2(\vec{X}) := \frac{1}{3}X_1 + \frac{1}{3}X_2 + \frac{1}{3}X_3$$

- We have found that $\hat{\mu}_2$ is more *efficient* since

$$\text{Var}\left(\hat{\mu}_2(\vec{X})\right) = \frac{3}{9}\sigma^2 < \frac{3}{8}\sigma^2 = \text{Var}\left(\hat{\mu}_1(\vec{X})\right)$$

- The *relative efficiency* of $\hat{\mu}_2$ with respect to $\hat{\mu}_1$ is

$$\frac{\text{Var}\left(\hat{\mu}_1\right)}{\text{Var}\left(\hat{\mu}_2\right)} = \frac{\frac{3}{8}\sigma^2}{\frac{3}{9}\sigma^2} = \frac{9}{8}$$

# Statement of the Cramér-Rao bound



- Let $f_Y(y;\theta)$ be a continuous PDF with continuous first and second derivatives
- Suppose that $\{y \mid f_Y(y) \neq 0\}$ does not depend on $\theta$
- We are given $n$ samples $\vec{Y} = \{Y_1, Y_2, \ldots, Y_n\}$
- Let $\hat{\theta}(\vec{Y})$ be an unbiased estimator of $\theta$
- Two expressions for the lower bound of $\mathrm{Var}(\hat{\theta})$

$$\mathrm{Var}(\hat{\theta}) \geq \left\{ n\, E\left[ \left( \frac{\partial \ln f_Y(Y;\theta)}{\partial \theta} \right)^2 \right] \right\}^{-1} = \left\{ -n\, E\left[ \frac{\partial^2 \ln f_Y(Y;\theta)}{\partial \theta^2} \right] \right\}^{-1}$$

- This gives us an upper bound on the efficiency of any unbiased estimator.
- The *absolute efficiency* of an unbiased estimator $\hat{\theta}$ is the ratio of the Cramér-Rao lower bound to the variance of $\hat{\theta}$.

# Example: Bernoulli and binomial distributions



- Suppose $p_X(k;p) = p^k(1-p)^{1-k}$ where $k \in \{0,1\}$
- Flip coin $n$ times, and define $X = X_1 + X_2 + \cdots + X_n$ where $X_j \in \{0,1\}$.
- Define the unbiased estimator $\hat{p} = X/n$
- The variance of the result is

$$\mathrm{Var}(\hat{p}) = \mathrm{Var}\left(\frac{X}{n}\right) = \frac{1}{n^2}\mathrm{Var}(X) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$$

# Cramér-Rao bound for binomial distribution I



- To calculate the Cramér-Rao bound, note

$$\ln p_{X_j}(X_j; p) = X_j \ln p + (1 - X_j) \ln(1 - p)$$

- Taking derivatives,

$$\frac{\partial \ln p_{X_j}(X_j; p)}{\partial p} = \frac{X_j}{p} - \frac{1 - X_j}{1 - p}$$

$$\frac{\partial^2 \ln p_{X_j}(X_j; p)}{\partial p^2} = -\frac{X_j}{p^2} - \frac{1 - X_j}{(1 - p)^2}$$

- Taking the expectation value

$$\left\{ -n\, E\left[ \frac{\partial^2 \ln p_{X_j}(X_j; p)}{\partial p^2} \right] \right\}^{-1} = \left\{ n \left( \frac{p}{p^2} + \frac{1 - p}{(1 - p)^2} \right) \right\}^{-1} = \frac{p(1 - p)}{n}$$

- $\mathrm{Var}(\hat{p})$ achieves Cramér-Rao bound – maximally efficient.

# Cramér-Rao bound for binomial distribution II



- Recall that the first derivative was

$$\frac{\partial \ln p_{X_j}(X_j;p)}{\partial p} = \frac{X_j}{p} - \frac{1-X_j}{1-p}$$

- Squaring the first derivative yields

$$\left(\frac{\partial \ln p_{X_j}(X_j;p)}{\partial p}\right)^2 = \frac{X_j^2}{p^2} - 2\frac{X_j}{p}\frac{1-X_j}{1-p} + \frac{(1-X_j)^2}{(1-p)^2} = \frac{(X_j-p)^2}{p^2(1-p)^2}$$

- Taking the expectation value then yields

$$\left\{n\, E\left[\frac{(X_j-p)^2}{p^2(1-p)^2}\right]\right\}^{-1} = \left\{n\left(\frac{p(1-p)}{p^2(1-p)^2}\right)\right\}^{-1} = \frac{p(1-p)}{n}$$

- $\mathrm{Var}(\hat{p})$ achieves Cramér-Rao bound – maximally efficient.

# Summary



- We have studied the concepts of *efficiency* and *relative efficiency* and presented examples.
- We have learned the statement of the *Cramér-Rao bound*, and the notation of *absolute efficiency*.
- Proof of *Cramér-Rao bound* given in another module.