Chapter 9: Properties of Point Estimators and Methods of Estimation

- We have studied two "good" properties of an estimator $\hat{\theta}$ of a target parameter θ.
- Both are related to the sampling distribution of $\hat{\theta}$:

- Now we will study several more good characteristics of estimators.
- These help us to answer:
  - Which estimator of θ is "better"?
  - Is there a single "best" estimator of θ?

9.2 Relative Efficiency

- Used to compare two unbiased estimators.
- If $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased for θ, if $V(\hat{\theta}_2) > V(\hat{\theta}_1)$, then:

The <u>efficiency</u> of $\hat{\theta}_1$ relative to $\hat{\theta}_2$ is the ratio:

- If $\text{eff}(\hat{\theta}_1, \hat{\theta}_2) > 1$, then:

<u>Example 1:</u> $Y_1, \ldots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$. Let $\bar{Y}$ be the sample mean, and let $\tilde{Y}$ be the sample median. Both are unbiased estimators of μ. Which is more efficient, for large samples?

<u>Example 2:</u> $Y_1, \ldots, Y_n \overset{iid}{\sim} \text{Unif}(0, \theta)$. Both $\hat{\theta}_1 = 2\bar{Y}$ and $\hat{\theta}_2 = \left(\frac{n+1}{n}\right) Y_{(n)}$ are unbiased for θ. What is the efficiency of $\hat{\theta}_1$ relative to $\hat{\theta}_2$?

Note:

Recall the density of $Y_{(n)}$ is:

Cramer-Rao Lower Bound (CRLB)

- Under general conditions, if $Y_1, \dots, Y_n$ are iid from some pdf $f(y|\theta)$, and $\hat{\theta}$ is an <u>unbiased</u> <u>estimator</u> of θ, then:

$$\mathrm{var}(\hat{\theta}) \geq \qquad\qquad\qquad\qquad = \text{CRLB}$$

- <u>Note</u>: $I(\theta)$ is called the (Fisher) <u>information</u> <u>number</u> of the sample.
- Any unbiased estimator whose variance <u>equals</u> this "Cramer-Rao Lower Bound" is simply called ____________.

<u>Example 1:</u> If $Y_1, \dots, Y_n \overset{iid}{\sim} N(\mu, \sigma^2)$, show that $\overline{Y}$ is an <u>efficient</u> estimator of μ.

- We know $\overline{Y}$ is

<u>Find CRLB</u>:

Example 2: If $Y_1, \ldots, Y_n \overset{iid}{\sim}$ Bernoulli (p), show that $\hat{p} = \frac{\sum Y_i}{n}$ is an efficient estimator of p.

- We know $\hat{p}$ is unbiased for p and $var(\hat{p}) =$

Find CRLB:

9.4 Sufficiency

- Another property of a "good" estimator of θ is that it is based on a statistic that summarizes all the information about θ contained in the whole sample.
- Such a statistic is called a sufficient statistic.
- "Good" estimators are generally functions of a sufficient statistic.

Definition: Let $Y_1, \ldots, Y_n$ be a random sample from a distribution with unknown parameter θ. The statistic $U = g(Y_1, \ldots, Y_n)$ is sufficient for θ if the conditional distribution of $Y_1, \ldots, Y_n$ given U does not depend on θ.

- This implies that if U is sufficient, then if we know U, nothing else in the sample can give us any more information about θ.
  - U itself contains all the available information about θ.

Likelihood Function

- Suppose $Y_1, \dots, Y_n$ are a random sample from pdf $f_Y(y)$.
- Then we call

the joint density function. It is a function of $y_1, y_2, \dots, y_n$ for a fixed (unknown) value of θ.

- We could also view $\prod_{i=1}^{n} f_Y(y_i; \theta)$ as a function of θ for a given data set $y_1, y_2, \dots, y_n$. In that case we write

and we call $L(\theta | y_1, \dots, y_n)$ the likelihood function.

- The joint density and the likelihood are the same function, but the first is treated as a function of the data, and the second as a function of the parameter θ.

Factorization Theorem: Let U be a statistic based on the random sample $Y_1, \dots, Y_n$. Then U is a sufficient statistic for θ if and only if the likelihood $L(\theta \mid \underset{\sim}{y})$ can be factored as:

Example 1: Let $Y_1, \dots, Y_n \overset{iid}{\sim} \text{Pois}(\lambda)$. Show $\sum_{i=1}^{n} Y_i$ is sufficient for λ.

<u>Example 2:</u> $Y_1, \dots, Y_n \overset{iid}{\sim} \text{Expon}(\beta)$. Show $\sum_{i=1}^{n} Y_i$ is sufficient for β.

<u>Note:</u> There are often many choices of sufficient statistics for θ.

- In Example 2, it is easy to see $\bar{Y}$ is sufficient for β.
- If U is sufficient, any one-to-one function $g(U)$ is also sufficient.
- The set of order statistics $\underset{\sim}{U} = (Y_{(1)}, Y_{(2)}, \dots, Y_{(n)})$ is always sufficient for θ.
- In choosing the "best" estimator for θ, we will often look at estimators which are functions of a sufficient statistic.