ISyE6739 – Statistical Methods

Point Estimation – Concepts and Properties (Ch. 7)

Instructor: Kamran Paynabar
H. Milton Stewart School of
Industrial and Systems Engineering
Georgia Tech

Kamran.paynabar@isye.gatech.edu
Office: Groseclose 436

List of Topics

- Point Estimation and Estimators
- Point Estimation Properties
  - Unbiased estimators
  - Variance and standard error of estimation
  - Minimum variance unbiased estimator
  - Relative efficiency

Summary of Sampling Distributions

		One population	Two populations
Sample Mean	Known Variance	$\frac{\overline{X}-\mu}{\sigma/\sqrt{n}} \sim N(0,1)$	$\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}} \sim NID(0,1)$
	Unknown Variance	$\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t(n-1)$	$\frac{(\overline{X}_1-\overline{X}_2)-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t(n_1+n_2-2)$
Sample Variance		$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$	$\frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1-1, n_2-1)$

Estimator

Suppose X is a random variable with $f(x;\theta)$ as the pdf. If $X_1, X_2, \ldots X_n$ is a random sample of size n from X, the statistic

$$\hat{\Theta} = h(X_1, X_2, \ldots, X_n)$$

Is called a **point estimator** of θ.

After the sample has been selected, $\hat{\Theta}$ takes on a particular numerical value called the **point estimate** of θ.

Parameter: μ **Estimator:** $\hat{\mu} = \overline{X} = \frac{\sum_{i=}^{n} X_i}{n}$ **Estimate:** $\overline{x} = \frac{25+30+29+31}{4} = 28.75$

Note that $\hat{\Theta}$ is a random variable because it is a statistic (function of random variables)

Examples of Point Estimators

Reasonable point estimates of these parameters are as follows:

- For μ, the estimate is $\hat{\mu} = \bar{x}$, the sample mean.
- For σ^2, the estimate is $\hat{\sigma}^2 = s^2$, the sample variance.
- For p, the estimate is $\hat{p} = x/n$, the sample proportion, where x is the number of items in a random sample of size n that belong to the class of interest.
- For $\mu_1 - \mu_2$, the estimate is $\hat{\mu}_1 - \hat{\mu}_2 = \bar{x}_1 - \bar{x}_2$, the difference between the sample means of two independent random samples.
- For $p_1 - p_2$, the estimate is $\hat{p}_1 - \hat{p}_2$, the difference between two sample proportions computed from two independent random samples.

General Concepts of Point Estimation: Unbiased Estimator

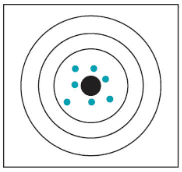
The point estimator $\hat{\Theta}$ is an **unbiased estimator** for the parameter θ if

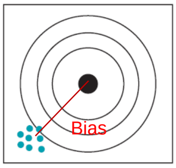
$$E(\hat{\Theta}) = \theta \qquad (7\text{-}5)$$

If the estimator is not unbiased, then the difference

$$E(\hat{\Theta}) - \theta \qquad (7\text{-}6)$$

is called the **bias** of the estimator $\hat{\Theta}$.





Example

Suppose that X is a random variable with mean μ and variance σ^2. Let $X_1, X_2, \ldots, X_n$ be a random sample of size n from the population represented by X. Show that the sample mean $\overline{X}$ and sample variance S^2 are unbiased estimators of μ and σ^2, respectively.

Example

Suppose the mean of a population is known (μ), show that the following estimator is a biased estimator for the variance

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n}(X_i - \mu)^2}{n-1}$$

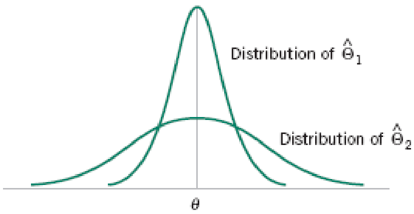
Modify this estimator such that it becomes an unbiased estimator

General Concepts of Point Estimation: Variance and Standard Error of a Point Estimator

If two estimators are unbiased, the one with smaller variance is preferred.

$$\text{var}(\hat{\Theta}_1) < \text{var}(\hat{\Theta}_2)$$

e.g., $\text{var}(\overline{X}) < \text{var}(\tilde{X})$



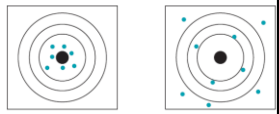


Figure 7-5 The sampling distributions of two unbiased estimators

The **standard error** of an estimator $\hat{\Theta}$ is its standard deviation, given by $\sigma_{\hat{\Theta}} = \sqrt{V(\hat{\Theta})}$. If the standard error involves unknown parameters that can be estimated, substitution of those values into $\sigma_{\hat{\Theta}}$ produces an **estimated standard error**, denoted by $\hat{\sigma}_{\hat{\Theta}}$.

Example

Suppose we are sampling from a normal distribution with mean μ and variance σ^2. Now the distribution of $\overline{X}$ is normal with mean μ and variance σ^2/n, so the standard error of $\overline{X}$ is

$$\sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

If we did not know σ but substituted the sample standard deviation S into the above equation, the estimated standard error of $\overline{X}$ would be

$$\hat{\sigma}_{\overline{X}} = \frac{S}{\sqrt{n}}$$

Example

An article in the *Journal of Heat Transfer* (Trans. ASME, Sec. C, 96, 1974, p. 59) described a new method of measuring the thermal conductivity of Armco iron. Using a temperature of 100°F and a power input of 550 watts, the following 10 measurements of thermal conductivity (in Btu/hr-ft-°F) were obtained:

$$41.60, 41.48, 42.34, 41.95, 41.86,$$
$$42.18, 41.72, 42.26, 41.81, 42.04$$

A point estimate of the mean thermal conductivity at 100°F and 550 watts is the sample mean or

$$\bar{x} = 41.924 \text{ Btu/hr-ft-°F}$$

The standard error of the sample mean is $\sigma_{\bar{X}} = \sigma/\sqrt{n}$, and since σ is unknown, we may replace it by the sample standard deviation $s = 0.284$ to obtain the estimated standard error of $\bar{X}$ as

$$\hat{\sigma}_{\bar{X}} = \frac{s}{\sqrt{n}} = \frac{0.284}{\sqrt{10}} = 0.0898$$

General Concepts of Point Estimation: Minimum Variance Unbiased Estimator (MVUE)

If we consider all unbiased estimators of θ, the one with the smallest variance is called the **minimum variance unbiased estimator** (MVUE).

$\hat{\Theta}$ is an MVUE if:

a) It is unbiased estimator of θ,
b) It satisfies the following equality,

$$\operatorname{var}(\hat{\Theta}) = \frac{1}{nE\left[\left(\frac{\partial \ln f(x)}{\partial \theta}\right)^2\right]}$$ **Cramér-Rao Bound**

General Concepts of Point Estimation: Minimum Variance Unbiased Estimator (MVUE)

If $X_1, X_2, \ldots, X_n$ is a random sample of size n from a normal distribution with mean μ and variance σ^2, the sample mean $\overline{X}$ is the MVUE for μ.

General Concepts of Point Estimation: Mean Square Error (MSE) and Relative Efficiency

The **mean squared error** of an estimator $\hat{\Theta}$ of the parameter θ is defined as

$$\mathrm{MSE}(\hat{\Theta}) = E(\hat{\Theta} - \theta)^2 \tag{7-7}$$

$$MSE(\hat{\Theta}) = E(\hat{\Theta} - \Theta)^2 = \left[E(\hat{\Theta} - \Theta)\right]^2 + \mathrm{var}(\hat{\Theta} - \Theta)$$

$$MSE(\hat{\Theta}) = \left[\mathrm{Bias}(\hat{\Theta})\right]^2 + \mathrm{var}(\hat{\Theta})$$

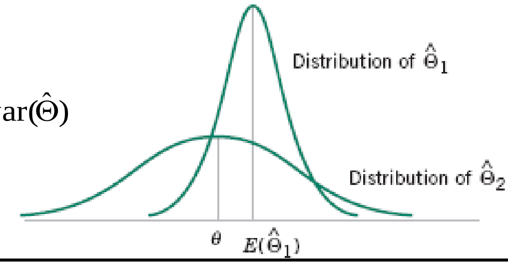
General Concepts of Point Estimation: Mean Square Error (MSE) and Relative Efficiency

The mean squared error is an important criterion for comparing two estimators. Let $\hat{\Theta}_1$ and $\hat{\Theta}_2$ be two estimators of the parameter θ, and let MSE $(\hat{\Theta}_1)$ and MSE $(\hat{\Theta}_2)$ be the mean squared errors of $\hat{\Theta}_1$ and $\hat{\Theta}_2$. Then the **relative efficiency** of $\hat{\Theta}_2$ to $\hat{\Theta}_1$ is defined as

$$\frac{\mathrm{MSE}(\hat{\Theta}_1)}{\mathrm{MSE}(\hat{\Theta}_2)} \tag{7-8}$$

If this relative efficiency is less than 1, we would conclude that $\hat{\Theta}_1$ is a more efficient estimator of θ than $\hat{\Theta}_2$, in the sense that it has a smaller mean square error.

$$MSE(\hat{\Theta}) = \left[\mathrm{Bias}(\hat{\Theta})\right]^2 + \mathrm{var}(\hat{\Theta})$$



Example

Let $X_1, X_2, \ldots X_7$ denote a random sample from a population with mean μ and variance σ^2. Calculate the relative efficiency of the following estimators of μ.

$$\hat{\Theta}_1 = \frac{\sum_{i=1}^{7} X_i}{7}$$

$$\hat{\Theta}_2 = \frac{2X_1 - X_6 + X_4}{2}$$

Example

Suppose $X \sim Uniform(\theta, 3\theta)$

- Show that $\frac{\bar{X}}{2}$ is an unbiased estimator of θ

- Calculate the relative efficiency of $\frac{\bar{X}}{2}$ and $\bar{X}$

Example

If *X* is normally distributed, which of the following estimators are more efficient estimator for the variance.

$$S_1^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$$

$$S_2^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$$