Chapter 8 - Statistical Modeling and Inference

STAT 251

Lecture 25

Point Estimation for μ and σ

Unbiased estimators

Confidence Interval for μ

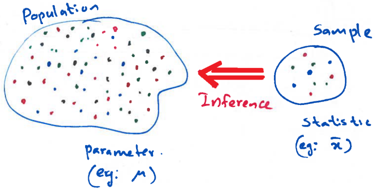
Dr. Lasantha Premarathna

Chapter 8 - Learning Outcomes

- Point Estimation for μ ans σ
- Bias of an estimator
- Confidence Interval for μ
- Testing of Hypotheses about μ
- One sample problems
- Two sample problems

Statistical Inference

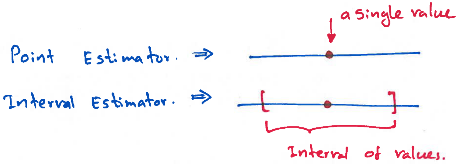
- Methods of making decisions or predictions about a population based on information obtained from a sample.



- The objective of estimation is to determine the approximate value of a population parameter on the basis of a sample statistic

  Ex: sample mean $\bar{x}$ is used to estimate the population mean μ.

Two Types of Estimators



- A **point estimator** draws inference about a population by estimating the value of an unknown parameter using a single value or point.
- An **interval estimator** draws inferences about a population by estimating the value of an unknown parameter using an interval
  - the population parameter of interest is between some lower and upper bounds
- A point estimate does not tell us how close the estimate is likely to be to the parameter.
- an interval estimate is more useful

Point Estimators

- A good estimator has a sampling distribution that is centered at the parameter (unbiasedness)
- An **unbiased** estimator of a population parameter is an estimator whose expected value is equal to that parameter.

$$E(\hat{\theta}) = \theta$$

θ is the parameter. $\hat{\theta}$ is a point estimator

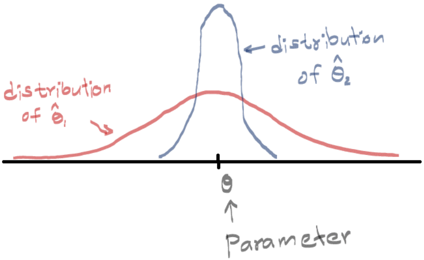
- **Bias of an estimator** is

$$bias(\hat{\theta}) = E(\hat{\theta}) - \theta$$

- An estimator is said to be unbiased if its bias is equal to zero

Point Estimators

- A good estimator has a **small standard error** compared to the other estimators.



- Let $\hat{\theta}_1$ and $\hat{\theta}_2$ are unbiased estimators for θ.

$$\text{i.e} \quad E(\hat{\theta}_1) = \theta \text{ and } E(\hat{\theta}_2) = \theta$$

- $\hat{\theta}_2$ is better than $\hat{\theta}_1$ because it has smaller standard error.

Example

Suppose that $X_1,\ X_2,\ \ldots,\ X_n$ is a random sample from a population with mean μ and variance σ^2.

- $\bar{x}$ is an unbiased estimator of μ
- s^2 is an unbiased estimator of σ^2

solution to this example is posted in a separate file

Recall: Sampling Distribution of Sample Mean

Consider a random sample $X_1,\ X_2,\ \ldots,\ X_n$ from a population with mean μ and variance σ^2

- Ch 7 $\Rightarrow$ If n is large, then by CLT, the sampling distribution of sample mean

$$\bar{X} \sim N(\mu,\ \frac{\sigma^2}{n})\ \ ; approximately$$

- Ch 5 $\Rightarrow$ If the underline population distribution is normal, then the sampling distribution of sample mean $(\bar{X})$ is normal. In this situation the sample size (n) need not to be large.

$$\bar{X} \sim N(\mu,\ \frac{\sigma^2}{n})$$

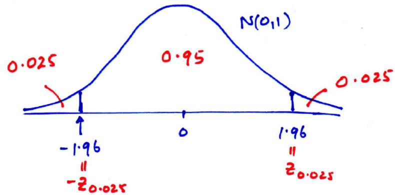
95% Confidence Interval for Population mean μ

Example: Let $X_1,\ X_2,\ \cdots,\ X_n$ is a random sample from a population with mean $= \mu$ and variance $= \sigma^2$. Consider either n is large or population is normal. then,

$$\bar{X} = \frac{X_1 + X_2 + \cdots + X_n}{n} \sim N\left(\mu,\ \frac{\sigma^2}{n}\right)$$

$$\Rightarrow \quad E[\bar{X}] = \mu \qquad Var(\bar{X}) = \frac{\sigma^2}{n}$$

$$\Rightarrow \quad Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \sim N(0,1)$$



95% Confidence Interval for Population mean μ

$$P\left(-1.96 \le \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} \le 1.96\right) = 0.95$$

$$P\left(-1.96\frac{\sigma}{\sqrt{n}} \le \bar{X}-\mu \le 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$P\left(-\bar{X}-1.96\frac{\sigma}{\sqrt{n}} \le -\mu \le -\bar{X}+1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$$P\left(\bar{X}-1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}+1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

$\Rightarrow$ 95% confidence interval for μ is $\left[\bar{X}-1.96\frac{\sigma}{\sqrt{n}}\ ,\ \bar{X}+1.96\frac{\sigma}{\sqrt{n}}\right]$

$$\Rightarrow \quad \left[\bar{X}-Z_{0.025}\frac{\sigma}{\sqrt{n}}\ ,\ \bar{X}+Z_{0.025}\frac{\sigma}{\sqrt{n}}\right]$$

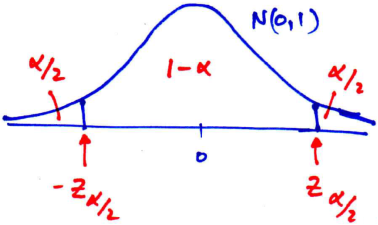
$(1-\alpha)100\%$ Confidence Interval (CI) for μ

In general, $(1-\alpha)100\%$ Confidence Interval (CI) for Population mean μ is

$$\left[\bar{X} - Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\ ,\ \bar{X} + Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}\right]$$

or

$$\bar{X} \pm Z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$



$(1-\alpha)100\%$ Confidence Interval (CI) for μ

Typical values of α are

$$\begin{array}{lllll} \alpha = 0.1 & \Rightarrow & 90\% \ \text{CI} & \Rightarrow & Z_{\frac{\alpha}{2}} = Z_{0.05} = 1.645 \\ \alpha = 0.05 & \Rightarrow & 95\% \ \text{CI} & \Rightarrow & Z_{\frac{\alpha}{2}} = Z_{0.025} = 1.96 \\ \alpha = 0.01 & \Rightarrow & 99\% \ \text{CI} & \Rightarrow & Z_{\frac{\alpha}{2}} = Z_{0.005} = 2.575 \end{array}$$

- $1-\alpha$ is called confidence level
- as confidence level increases, the width of the CI increases
- as n increases, the width of the CI decreases.

Confidence Interval

In general, a confidence interval takes the form

$$\text{point estimate} \pm \text{ margin of error}$$

$(1-\alpha)100\%$ CI for μ is

$$\bar{x} \pm Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$\bar{x}$ is the point estimator for μ

$Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$ is the margin of error

- The purpose of an interval estimate is to provide information about how close the point estimate is to the value of the parameter.

Before the next class ...

Visit the course website at canvas.ubc.ca

- Review Lecture 25 and related sections in the text book
- Topic of next class: **Chapter 8: Point and Interval estimation Examples, t-distribution**