

Unit 5 *Confidence Intervals*

Introduction

Learning objectives

This unit brings together earlier work from the Normal Distribution and Sampling units in order to make estimates on the mean of a population based on the sample mean and the size of the sample. After studying this unit you should

- understand the concept of a confidence interval
- appreciate the meaning of the central limit theorem
- know how to calculate a confidence interval of a certain percentage probability, given the sample mean and sample size.

Notes

Confidence intervals give an interval in which the mean of a population lies, subject to a certain probability, e.g. 95% confidence intervals are based on the probability of being correct 95% of the time, that is, probability of being wrong is 5% or 1 in 20. The interval is based on both the sample mean and the size of the sample. The larger the sample, the more certain you can be of the confidence interval containing the population mean.

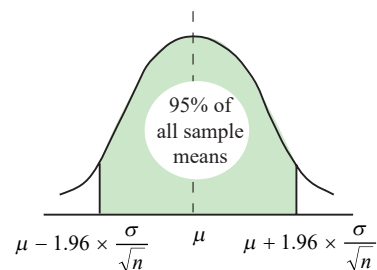
Key points

- The central limit theorem states that the distribution of $\bar{X}$, the sample mean of size n , is approximately $N\left(\mu, \frac{\sigma^2}{n}\right)$
- As the size of samples, n , increases, the variance tends to zero

The 95% confidence interval is given by

$$\mu \pm 1.96 \frac{\sigma}{\sqrt{n}}$$

where $\frac{\sigma}{\sqrt{n}}$ is known as the standard error



Distribution of all possible sample means for random samples of size n

Facts to remember

- $X \sim N\left(\mu, \frac{\sigma^2}{n}\right)$
- Standard error is $\frac{\sigma}{\sqrt{n}}$
- 95% confidence interval is given by $\left(\mu - 1.96 \frac{\sigma}{\sqrt{n}}, \mu + 1.96 \frac{\sigma}{\sqrt{n}}\right)$
- 95% confidence interval means that there is a 5% probability (that is, 1 in 20) of being wrong, i.e. the confidence interval does not include the population mean

Unit 5 *Confidence Intervals*Introduction

- When the populations variance is unknown, you estimate the variance by $\frac{ns^2}{(n-1)}$ where s^2 is the sample variance.

Glossary of terms

Confidence intervals: gives a range of values for which the mean of the population is expected to be within a certain percentage probability (e.g. 95% or 99%) based on the sample mean and sample size

Central limit theorem: the distribution of sample means, $\bar{X}$, of size n have approximate distribution

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

giving 95% confidence interval as $\left(\mu - 1.96\frac{\sigma}{\sqrt{n}}, \mu + 1.96\frac{\sigma}{\sqrt{n}}\right)$

Standard form: defined as $\frac{\sigma}{\sqrt{n}}$