Date / /
Page No.

Confidence Intervals . Introduction

Key Concepts

1) Confidence Interval (CI) - A range of values likely to contain the population parameter

2) Why CI is Not a Probability - A confidence interval does not mean there is a 95% probability that the interval contains the parameter

Example estimating Mean weight

- You want to estimate the mean weight of 10 year old girls in the UK
- A sample of 16 gives a mean weight of 90 pounds (point estimate)
- The point estimate alone lacks information about uncertainty
  - It does not tell how close it is to the true population mean

Confidence intervals provide More information

A 95% confidence Interval (CI) means that if we repeatedly took samples and calculated CIs, 95% of them would contain the true mean

Example of a 95% CI

$$72.85 < \mu < 107.15$$

Interpretation: The mean is likely within this range 5% of the time, a CI will not contain the the mean.

Date / /
Page No.

Why confidence intervals are not probabilities

- A 95% of CI does not mean that there is a 0.95 probability that the true mean is within the interval.

Key Issues
- Other prior information (eg previous studies showing means above 110) can affect interpretation
- Different methods can produce different CI, that contain the population mean 95% of the time.
- The standard CI method is commonly used because it is symetrical and contiguous around the estimat

Confidence Interval on the mean

Understanding confidence intervals for the mean

A confidence Interval (CI) on the mean estimates population mean (μ) based on a sample Mean (M). Since the true mean is unknown, a CI gives a range where the true mean likely falls.

Example known population standard deviation (σ)

Assume the weights of 10 year old children are normally distributed.

Population mean (μ) = 90
Population standard deviation (σ) = 36
Sample size $N = 9$

Step 1. compute standard error (SE)

$$\sigma_M = \frac{\sigma}{\sqrt{N}} = \frac{36}{\sqrt{9}} = 12$$

The standard error (σ_M) measures how much sample means vary from the population mean.

step2: Compute 95% Confidence Interval

For a normal distribution, 95% of values fall within 1.96 standard deviation of the mean

Lower Limit $= 90 - (1.96 \times 12) = 66.48$

Upper limit $= 90 + (1.96 \times 12) = 113.52$

Thus, the 95% confidence interval is $(66.48, 113.52)$

This means that if we repeatedly sample from the population 95% of computed CIs will contain the true mean.

Example: When σ is Unknown (Use t distribution)

When the population standard deviation (σ) is unknown we estimate it from the sample standard deviation (s) and use the t-distribution.

Step1: Compute sample mean and standard error

Given sample 2, 3, 5, 6, 9

Sample Mean $(M) = 5$

Sample Variance $(s^2) = 7.5$

Estimated standard error

$$s_m = \frac{s}{\sqrt{N}} = \frac{2.5}{\sqrt{5}} = 1.225$$

Step2: find t value for 95% CI $(df = N-1 = 4)$

From a t-table for $dF = 4$, the t value for 95% confidence is 2.776.

step3 : compute 95 % confidence interval.

lower limit $= 5 - (2.776 \times 1.225) = 1.60$

upper limit $= 5 + (2.776 \times 1.225) = 8.40$

Thus the 95 % confidence interval is (1.60, 8.40)

When to use t vs Normal distribution

- Use Normal distribution (Z-values) when σ is known.
- Use t distribution when σ is unknown and estimated from sample data.
- As sample size increases ($N \geq 100$), the t-distribution approaches the normal distribution

T-Distribution Key concepts

2 Difference Between the t-Distribution and Normal Distribution

- The t distribution is similar to the Normal distribution but has heavier tails (leptokurtic), meaning it has more extreme values
- The shape of the t-distribution depends on the degrees of freedom (df)
- As df increases, the t-distribution becomes more like the standard normal distribution (Z-distribution)

3 Finding t-values from the t-Table

- The t-table provides critical t-values for different confidence levels and df
- Example for 95 % confidence and df = 8, the t-value is 2.306 (from table I)

- for higher confidence levels (99%) the t values are larger since a wid interval is needed to capture the population mean

4) using a t-calculator

- A t-calculator helps determine the area in the tails of a t-distribution for any given score
- example: If df = 8, the probability that the sample mean falls within 1.96 standard error is 0.914 compared to 0.95 in the normal distribution
- This confirms that using an estimated standard deviation increases uncertainity, requiring wider confidence intervals.