



Data Science Notes by Sarowar Ahmed



Chapter: inferential statistics



Topic: Confidence Intervals



What are Confidence Intervals?

- Confidence intervals help us estimate the range of values for a population parameter (like a mean or proportion) based on a sample from that population. They express the uncertainty inherent in our estimates and provide a measure of how confident we are in our findings.



Formula:

The general formula for a confidence interval for a population mean (μ) with known standard deviation (σ) is:

$$CI = \bar{x} \pm z(\sigma/\sqrt{n})$$

Where:

- \bar{x} is the sample mean
- z is the z-score corresponding to the desired confidence level
- σ is the population standard deviation
- n is the sample size

 Example:

- Suppose we want to estimate the average height of students in a school. We take a random sample of 100 students and find the average height to be 65 inches with a standard deviation of 3 inches. Assuming a 95% confidence level (corresponding to a z-score of approximately 1.96 for large samples), the confidence interval would be:

$$CI = 65 \pm 1.96(3/\sqrt{100}) = 65 \pm 0.588 = (64.412, 65.588)$$

This means we are 95% confident that the true average height of all students in the school falls within the range of 64.412 inches to 65.588 inches.

💡 Why are Confidence Intervals Useful?

- Confidence intervals provide a way to quantify the uncertainty in our estimates and communicate the precision of our findings to others. They are essential in making informed decisions based on data, whether in scientific research, public policy, or business strategy.

🔍 Key Takeaway:

- Understanding confidence intervals empowers us to interpret statistical results more effectively and make sound decisions based on reliable evidence.

Got any questions on the Confidence Intervals? Feel free to ask me via LinkedIn! Let's keep learning together.

My LinkedIn

Date: 10/05/2024

