Confidence Intervals

Noara Razzak

May 9, 2020

Confidence Interval (CI)

A **confidence interval** provides a range of plausible values for a population parameter (e.g., the mean) with a specified confidence level $(1 - \alpha)$.

General Formula For a population mean μ, the confidence interval is:

$$\text{CI} = \left[\bar{X} - E, \bar{X} + E\right]$$

where:

- $\bar{X}$ = sample mean,
- E = margin of error, which depends on the distribution used (normal or t).

Population Mean Using Normal Distribution

When the population standard deviation (σ) is known and the sample size (n) is large ($n \geq 30$), we use the **normal distribution (Z-distribution)**.

Formula

$$\text{CI} = \bar{X} \pm z_{\alpha/2} \left(\frac{\sigma}{\sqrt{n}} \right)$$

where:

- $z_{\alpha/2}$ = critical value from the standard normal distribution (e.g., 1.96 for 95% confidence),
- σ = population standard deviation,
- n = sample size.

Example

If $\bar{X} = 50$, $\sigma = 10$, $n = 100$, and a 95% confidence level ($z_{0.025} = 1.96$):

$$\text{CI} = 50 \pm 1.96 \left(\frac{10}{\sqrt{100}} \right) = 50 \pm 1.96 \times 1 = [48.04, 51.96]$$

Population Mean Using Student t-Distribution

When σ is unknown and the sample size is small ($n < 30$), we use the **Student t-distribution**.

Formula

$$\text{CI} = \bar{X} \pm t_{\alpha/2,\, n-1} \left(\frac{s}{\sqrt{n}} \right)$$

where:

- $t_{\alpha/2,\, n-1}$ = critical value from the t-distribution with $n - 1$ degrees of freedom,
- s = sample standard deviation.

Example

If $\bar{X} = 50$, $s = 10$, $n = 25$, and a 95% confidence level ($t_{0.025,\,24} \approx 2.064$):

$$\text{CI} = 50 \pm 2.064\left(\frac{10}{\sqrt{25}}\right) = 50 \pm 2.064 \times 2 = [45.872, 54.128]$$

Key Differences

Criterion	**Normal (Z)**	**Student t**
σ known?	Yes	No
Sample size (n)	Large (≥ 30)	Small (< 30)
Distribution	Standard normal	t-distribution
Critical value	$z_{\alpha/2}$	$t_{\alpha/2,\, n-1}$

Table 1: Comparison of Normal and t-Distribution for CI Estimation

Recap

- Use the **normal distribution** when σ is known or n is large.
- Use the ***t*-distribution** when σ is unknown and n is small.
- The confidence interval width depends on variability, sample size, and confidence level.

Next class we will cover the concept of expectations and variances of different distributions