Confidence Intervals (CI)

Confidence Intervals (CI) in Statistics

A **confidence interval (CI)** is a range of values used to estimate an unknown population parameter (like a mean or proportion) with a certain level of confidence. It gives a measure of uncertainty around an estimate.

---

1. What Does a Confidence Interval Represent?

A **confidence interval** provides a range within which the true population parameter is likely to fall.

For example, if you compute a **95% confidence interval** for the **mean height** of students in a school, you might get:

$$(160 \text{ cm}, 170 \text{ cm})$$

This means:

- We are **95% confident** that the true average height of all students is between **160 cm and 170 cm**.
- If we were to take multiple samples and compute a confidence interval each time, **95% of those intervals would contain the true mean**.

---

2. General Formula for Confidence Intervals

A confidence interval is typically written as:

$$\text{Estimate} \pm (\text{Critical Value}) \times (\text{Standard Error})$$

or,

$$\left(\hat{\theta} - z^* \cdot SE, \hat{\theta} + z^* \cdot SE\right)$$

Where:

- $\hat{\theta}$ = Sample estimate (e.g., sample mean $\bar{X}$ or sample proportion $\hat{p}$)
- z^* = **Critical value** from the standard normal (Z) table or t-distribution table, based on the confidence level

- SE = **Standard Error**, which measures how much the sample estimate varies

---

3. Confidence Interval for a Population Mean

When estimating a population mean μ using a sample mean $\bar{X}$, the confidence interval is:

$\bar{X} \pm z^* \cdot \frac{\sigma}{\sqrt{n}}$

Where:

- $\bar{X}$ = Sample mean
- z^* = Critical value from Z-table for the desired confidence level
- σ= Population standard deviation (or sample standard deviation if unknown)
- n = Sample size

Common Z Critical Values

Confidence Level	z^* Value
90%	1.645
95%	1.96
99%	2.576

---

4. Example: Confidence Interval for Mean

Suppose a sample of **n = 50** students has a mean test score of **75** and a standard deviation of **10**. Find a **95% confidence interval** for the population mean.

Using the formula:

$\bar{X} \pm z^* \cdot \frac{\sigma}{\sqrt{n}}$

Substituting values:

$75 \pm 1.96 \times \frac{10}{\sqrt{50}}$

$75 \pm 1.96 \times 1.41$

75 ± 2.76

(72.24, 77.76)

Interpretation: We are **95% confident** that the true average test score of all students is between **72.24 and 77.76**.

---

5. Confidence Interval for a Population Proportion

When estimating a **population proportion** p, the confidence interval is:

D

Where:

- $\hat{p}$ = Sample proportion ($\frac{x}{n}$, where x is the number of successes)
- z^* = Critical value from the Z-table
- n = Sample size

Example: Confidence Interval for Proportion

Suppose **200** people are surveyed, and **60%** favor a new policy. Find a **95% confidence interval**.

$$0.6 \pm 1.96 \times \sqrt{\frac{0.6(1-0.6)}{200}}$$

$$0.6 \pm 1.96 \times \sqrt{\frac{0.24}{200}}$$

$$0.6 \pm 1.96 \times 0.0346$$

$$0.6 \pm 0.0679$$

(0.532, 0.668)

Interpretation: We are **95% confident** that the true proportion of people who favor the policy is between **53.2% and 66.8%**.

---

6. Key Takeaways

✅ A **confidence interval** provides a **range** within which the true population parameter is likely to fall.

✅ The **wider the interval**, the **more uncertainty** exists in the estimate.

✅ **Higher confidence levels (e.g., 99%)** produce **wider** intervals, meaning more certainty but less precision.

✅ **Larger sample sizes** reduce the margin of error, making the interval **narrower** (more precise).
