Confidence Intervals

Statistical Inference and Estimation

Yolymatics Tutorials

www.yolymaticstutorials.com

November 3, 2025

Professional Mathematics Education

Overview

What is a Confidence Interval?

Definition

A confidence interval is a range of values, computed from sample data, that is likely to contain the true population parameter with a specified level of confidence.

General Form:

$$\text{Point Estimate} \pm \text{Margin of Error}$$

Key Components:

- **Point Estimate**: Sample statistic (e.g., $\bar{x}$, $\hat{p}$)
- **Margin of Error**: Uncertainty measure based on standard error and confidence level
- **Confidence Level**: Probability that interval contains true parameter (e.g., 90%, 95%, 99%)

Interpretation of Confidence Intervals

Correct Interpretation:
If we construct many $(1-\alpha)100\%$ confidence intervals from repeated samples, approximately $(1-\alpha)100\%$ of them will contain the true population parameter.

Common Confidence Levels:

- 90% confidence level: $\alpha = 0.10$, $z_{\alpha/2} = 1.645$
- 95% confidence level: $\alpha = 0.05$, $z_{\alpha/2} = 1.96$
- 99% confidence level: $\alpha = 0.01$, $z_{\alpha/2} = 2.576$

Trade-offs:

- Higher confidence $\Rightarrow$ Wider interval
- Larger sample size $\Rightarrow$ Narrower interval

CI for Mean: Known Population Variance (σ^2)

When σ is known (rare in practice):

$$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

Where:

- $\bar{x}$ = sample mean
- $z_{\alpha/2}$ = critical value from standard normal distribution
- σ = population standard deviation (known)
- n = sample size

Assumptions:

- Random sample
- Population is normally distributed OR $n \geq 30$ (CLT)
- σ is known

CI for Mean: Unknown Population Variance (σ^2)

When σ is unknown (most common):

$$\bar{x} \pm t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}}$$

Where:

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

Assumptions:

- Random sample
- Population is approximately normally distributed (especially for small n)
- σ is unknown

t-Distribution Critical Values

Common critical values for t-distribution:

df	90% CI ($t_{0.05}$)	95% CI ($t_{0.025}$)	99% CI ($t_{0.005}$)
5	2.015	2.571	4.032
10	1.812	2.228	3.169
15	1.753	2.131	2.947
20	1.725	2.086	2.845
25	1.708	2.060	2.787
30	1.697	2.042	2.750
∞ (z)	1.645	1.96	2.576

Note: As df $\to \infty$, t-distribution $\to$ standard normal (z)

Practice Problem 1: Battery Life (Known σ)

A manufacturer claims their batteries have a mean life of 500 hours. A random sample of 36 batteries has a mean life of 485 hours. The population standard deviation is known to be 40 hours.

Construct a 95% confidence interval for the true mean battery life.

Problem 1: Working Space

Practice Problem 2: Student Heights (Unknown σ)

A random sample of 25 university students has a mean height of 170 cm with a standard deviation of 8 cm.

Construct a 90% confidence interval for the mean height of all university students.

Problem 2: Working Space

Practice Problem 3: Fuel Efficiency

An automotive engineer tests 15 cars and finds the mean fuel efficiency is 32.5 mpg with a standard deviation of 3.2 mpg.

Calculate a 99% confidence interval for the true mean fuel efficiency.

Problem 3: Working Space

Practice Problem 4: Test Scores

A teacher records test scores for 20 students. The sample mean is 78.4 and the sample standard deviation is 12.6.

Construct a 95% confidence interval for the true mean test score.

Problem 4: Working Space

Practice Problem 5: Manufacturing Process

A quality control inspector measures 40 components. The mean diameter is 5.02 mm with a standard deviation of 0.15 mm.

Find a 90% confidence interval for the true mean diameter.

Problem 5: Working Space

CI for Population Proportion

For large samples:

$$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Where:

- $\hat{p} = \frac{x}{n}$ = sample proportion
- x = number of successes in sample
- n = sample size
- $z_{\alpha/2}$ = critical value from standard normal distribution

Conditions:

- Random sample
- $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$ (large sample approximation)
- Population size $\geq 10n$ (or use finite population correction)

Practice Problem 6: Voter Preference

In a survey of 500 registered voters, 285 indicated they support a particular candidate.

Construct a 95% confidence interval for the true proportion of voters who support this candidate.

Problem 6: Working Space

Practice Problem 7: Product Defect Rate

A quality inspector examines 200 products and finds 12 are defective.

Calculate a 90% confidence interval for the true defect rate.

Problem 7: Working Space

Practice Problem 8: Medical Treatment Success

In a clinical trial, 145 out of 180 patients showed improvement after treatment.

Construct a 99% confidence interval for the true success rate of the treatment.

Problem 8: Working Space

Practice Problem 9: Student Satisfaction

A university surveys 350 students, and 273 report being satisfied with campus facilities.

Find a 95% confidence interval for the proportion of all students satisfied with campus facilities.

Problem 9: Working Space

Practice Problem 10: Market Share

A market research firm surveys 600 consumers. 228 currently use Brand X.

Construct a 90% confidence interval for Brand X's true market share.

Problem 10: Working Space

Determining Sample Size

For estimating a mean (known σ):

$$n = \left(\frac{z_{\alpha/2} \cdot \sigma}{E}\right)^2$$

where E = desired margin of error

For estimating a proportion:

$$n = \hat{p}(1 - \hat{p}) \left(\frac{z_{\alpha/2}}{E}\right)^2$$

If no prior estimate of $\hat{p}$, use $\hat{p} = 0.5$ (most conservative)

Always round up to next integer!

Practice Problem 11: Required Sample Size for Mean

A researcher wants to estimate the mean weight of adults with a margin of error of 2 kg at 95% confidence. Previous studies suggest $\sigma = 15$ kg.

How many adults should be sampled?

Problem 11: Working Space

Practice Problem 12: Required Sample Size for Proportion

A political pollster wants to estimate voter support with a margin of error of 3% at 95% confidence. No prior estimate is available.

What sample size is required?

Problem 12: Working Space

CI for Difference of Two Means (Independent Samples)

When both σ_1 and σ_2 are unknown:

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2, df} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Degrees of freedom (Welch approximation):

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}$$

Assumptions:

- Independent random samples
- Both populations approximately normal (especially for small samples)

Practice Problem 13: Comparing Two Teaching Methods

Method A: $n_1 = 30$, $\bar{x}_1 = 82.5$, $s_1 = 8.2$
Method B: $n_2 = 25$, $\bar{x}_2 = 78.3$, $s_2 = 9.5$

Construct a 95% confidence interval for the difference in mean scores between the two methods.

Problem 13: Working Space

Practice Problem 14: Drug Efficacy Comparison

Drug A: $n_1 = 40$, $\bar{x}_1 = 15.2$ days, $s_1 = 3.8$ days
Drug B: $n_2 = 35$, $\bar{x}_2 = 18.5$ days, $s_2 = 4.2$ days

Find a 90% confidence interval for the difference in mean recovery time.

Problem 14: Working Space

CI for Difference of Two Proportions

For large samples:

$$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Where:

- $\hat{p}_1 = \frac{x_1}{n_1}$ and $\hat{p}_2 = \frac{x_2}{n_2}$ = sample proportions
- n_1, n_2 = sample sizes

Conditions:

- Independent random samples
- $n_1\hat{p}_1 \geq 10,\ n_1(1-\hat{p}_1) \geq 10$
- $n_2\hat{p}_2 \geq 10,\ n_2(1-\hat{p}_2) \geq 10$

Practice Problem 15: Gender Wage Gap

Sample 1 (Men): $n_1 = 200$, 156 earn above median
Sample 2 (Women): $n_2 = 180$, 126 earn above median

Construct a 95% confidence interval for the difference in proportions.

Problem 15: Working Space

Practice Problem 16: Treatment Comparison

Treatment A: 82 successes out of 120 patients
Treatment B: 65 successes out of 100 patients

Find a 99% confidence interval for the difference in success rates.

Problem 16: Working Space

Confidence Intervals: Summary Table

Parameter	Confidence Interval Formula
Mean (σ known)	$\bar{x} \pm z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$
Mean (σ unknown)	$\bar{x} \pm t_{\alpha/2,n-1} \cdot \frac{s}{\sqrt{n}}$
Proportion	$\hat{p} \pm z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$
Difference of Means	$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2,df} \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
Difference of Proportions	$(\hat{p}_1 - \hat{p}_2) \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$

Key Points to Remember

- Use **z-distribution** when σ is known or for proportions
- Use **t-distribution** when σ is unknown (most common for means)
- Check **assumptions** before constructing intervals
- **Wider intervals** $=$ higher confidence level
- **Narrower intervals** $=$ larger sample size
- Always verify **normality** or **large sample conditions**
- Round sample sizes **up** to ensure desired precision

Thank You

Questions?

Yolymatics Tutorials
www.yolymaticstutorials.com

Professional Mathematics Education