Quantitative Reasoning – LAB 413

Confidence Intervals & Hypothesis Testing

Week 7 Recap ▪ Lab 6

Instructor: **Subin Na**

October 15, 2025

Today's Plan

1. **Quick Recap**
   - Overview
   - Confidence Interval
   - How CIs Behave
   - Hypothesis Testing
   - p-Values & Significance

2. **RStudio Practice**
   - Regression CIs on the `Prestige` data

01 Overview

- **Goal:** Quantify the uncertainty behind point estimates using probability.
- Two types of statistical inference:

Confidence Intervals

estimate **population parameters**

Hypothesis Tests

assess **evidence about claims** concerning the population

02 Confidence Intervals (1)

Key Concepts

- **Point estimate:** a single best guess of the parameter (often incorrect because of sampling variability).
- **Confidence interval (CI):** a range of plausible values for the parameter.

Formula

$$\text{CI} = \text{point estimate} \pm \underbrace{\text{margin of error}}_{\text{random variation from sampling}}$$

02 Confidence Intervals (2)

$$\text{CI} = \bar{x} \pm z^* \cdot \frac{\sigma}{\sqrt{n}}$$

- $\bar{x}$: sample mean
- z^* : critical value based on our confidence level
- $\frac{\sigma}{\sqrt{n}}$: standard error — how much the sample mean varies from sample to sample

02 Confidence Intervals (3)

Interpretation

A **95% CI** means: if we repeated our sampling process many times, about **95% of those intervals** would contain the true population parameter.

$$\mathrm{CI}_{90} = \bar{x} \pm 1.645 \cdot \mathrm{SE}$$

$$\mathrm{CI}_{95} = \bar{x} \pm 1.96 \cdot \mathrm{SE}$$

$$\mathrm{CI}_{99.7} = \bar{x} \pm 2.576 \cdot \mathrm{SE}$$

03 How CIs Behave

Key Concepts

The width of a confidence interval is driven by the **margin of error**.

The CI gets smaller when:

- the confidence level decreases (smaller z^*),
- the population SD decreases (smaller σ),
- the sample size increases (larger n).

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

Worked Example: Sleep Hours

Scenario

The director of public health in Philadelphia wants to estimate the average number of hours adults sleep per night. A survey of $n = 250$ randomly selected residents gives a sample mean of **6.8 hours** with a standard deviation of **1.4 hours**.

Construct a **95% confidence interval** for the true mean number of sleep hours among all adults in Philadelphia. What interval would you report?

Worked Example: Solution

Given

$$\bar{x} = 6.8, \qquad s = 1.4, \qquad n = 250$$

Compute the 95% CI

$$\bar{x} \pm z^* \cdot \frac{\sigma}{\sqrt{n}} = 6.8 \pm 1.96 \cdot \frac{1.4}{\sqrt{250}}$$

$$= 6.8 \pm 1.96 \cdot 0.0886 = 6.8 \pm 0.174$$

$$\Rightarrow \qquad (6.63,\ 6.97) \text{ hours}$$

The Setup

Null hypothesis H_0

Status quo, “no effect” — what we try to disprove.

Alternative hypothesis H_1

What we hope to prove.

Test statistic

$$z = \frac{\text{estimate} - \text{hypothesized value}}{\text{standard deviation of the estimate}} = \frac{\text{observed} - \text{null guess}}{\text{SE}}$$

$$\text{Large } |z| \Rightarrow \text{ evidence against } H_0.$$

p-Values & Significance

p-value

The probability of observing data as extreme as ours *if H_0 were true.*

p-value range	Interpretation	Symbol
$p \geq 0.10$	Not significant	
$0.05 \leq p < 0.10$	Marginally significant	$*$
$p < 0.05$	Significant	$**$
$p < 0.01$	Highly significant	$***$

Decision rule (α)

- If $p < \alpha$: reject H_0.
- If $p \geq \alpha$: fail to reject H_0.

Never "accept H_0" — only **reject** or **fail to reject**.

RStudio Practice

Regression Confidence Intervals with the `Prestige` data

See `Lab6.Rmd` / `Lab6_1015.R`