

Quantitative Reasoning – LAB 413

Z-tests, T-tests & Interaction Effects

Review Session ▪ Lab 11

Instructor: **Subin Na**

November 19, 2025

Today's Plan

1. Quick Review

- Z-tests and T-tests
- Choosing between them
- Worked exercises
- Interaction effects in regression

2. RStudio Practice

- Interaction effects

01 What is a Z-test?

Key Concepts

A hypothesis test used to compare a **sample mean** to a **population mean** when:

- The population standard deviation (σ) is **known**.
- The sample size is large ($n > 30$).

Formula

$$Z = \frac{\bar{X} - \mu_0}{\sigma / \sqrt{n}}$$

02 What is a T-test?

Key Concepts

A hypothesis test to compare a **sample mean** to a **population mean** when:

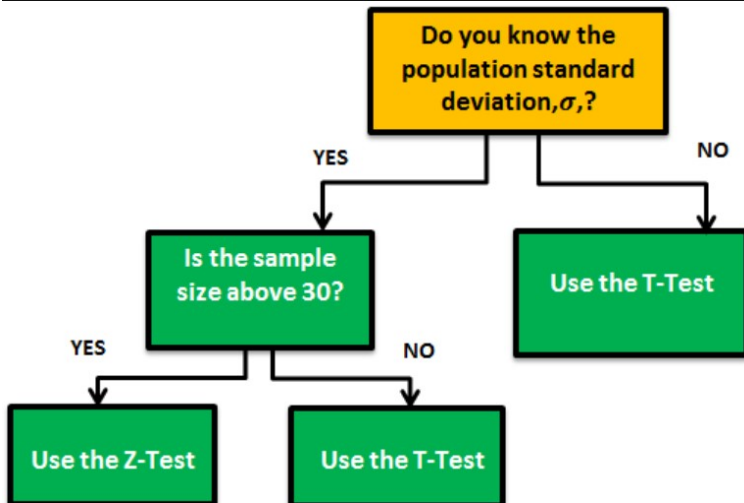
- The population standard deviation is **unknown**.
- You estimate σ using the sample standard deviation (s).

Formula

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

The t -distribution accounts for the additional uncertainty from estimating σ .

Choosing a Test: Z or T?



03 Z-test vs. T-test (Comparison)

Criterion	Z-test	T-test
Sample Size	Typically large ($n > 30$)	Any size, especially $n \leq 30$
Population SD (σ)	Known	Unknown (use sample SD)
Distribution	Standard Normal	t -distribution ($df = n - 1$)
Test Statistic	σ used	s used

Exercise 1

Scenario

You have a sample of **50 people's weights**, and the population standard deviation is **known**. You want to test whether the sample comes from a population with mean = 150 lbs.

Use a Z-test.

Reason: σ known + $n > 30$.

Exercise 2

Scenario

A botanist studies a new soil additive advertised to promote faster plant growth. After applying it across greenhouse units, the botanist measures the final height of **300 plants** grown under the treatment. No prior study provides a reliable estimate of the population standard deviation for this species. The botanist wants to evaluate whether the average plant height differs from the manufacturer's claim of **20 cm**.

Use a T-test.

Reason: σ unknown (the sample SD must be used).

Exercise 3

Scenario

A large hospital system evaluates whether a newly implemented treatment protocol affects average recovery time. Clinical records from past years provide a reliable estimate of the population standard deviation. A research team collects data on **200 patients** who received the updated treatment and records their recovery times, to examine whether the new protocol leads to a *different* average recovery time than the standard approach.

Use a Z-test.

Reason: σ known + large sample.

04 Interaction Effects in Regression

Key Concepts

An **interaction effect** means the effect of one predictor on the outcome *depends on* the value of another predictor.

In R, the ***** operator expands to main effects plus the interaction:

$$x1 * x2 = x1 + x2 + x1:x2$$

- The $x1:x2$ term captures how the slope of one variable *shifts* across levels of the other.

05 Binary $\times$ Continuous Interaction

Key Concepts

When a **binary** variable interacts with a **continuous** one, each group gets its *own slope* for the continuous predictor.

$$Y = \beta_0 + \beta_1 D + \beta_2 X + \beta_3 (D \times X) + \varepsilon$$

- When $D = 0$: slope of X is β_2 .
- When $D = 1$: slope of X is $\beta_2 + \beta_3$.
- β_3 is the **difference in slopes** between the two groups.

RStudio Practice

Regression with Interaction Terms

See `Lab11.Rmd` / `Lab11_1119.R`