Quantitative Reasoning – LAB 413

Confidence Intervals & Hypothesis Testing

Week 9 Recap ▪ Lab 7

Instructor: **Subin Na**

October 22, 2025

Today's Plan

1. **Quick Recap**
   - Overview
   - Confidence Interval: Concept
   - Confidence Interval: In Regression
   - Hypothesis Testing: Logic
   - One- vs. Two-Sided Tests
   - Interpreting Significance
   - Statistical vs. Practical Significance
   - Testing Errors
2. **RStudio Practice**

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

Key Concepts

- A **point estimate** gives our best single guess of the population value (imperfect because of random sampling variability).
- A **confidence interval (CI)** provides a *range of plausible values* for that population parameter.

$$\text{CI} = \text{point estimate} \pm (\text{critical value}) \times (\text{standard error})$$

03 Confidence Intervals (2): In Regression

Key Concepts

The CI shows how much **uncertainty** surrounds our estimated effect.

Example. A one-month increase in age predicts a 0.25-inch increase in height (95% CI: 0.22–0.28).

- If we repeated this study many times, **95% of the intervals** we compute would contain the true effect.

04 Hypothesis Testing: Logic (1)

Key Concepts

Hypothesis testing is a **structured way to evaluate a claim** about a population parameter.

Null hypothesis H_0	Alternative hypothesis H_1
"no effect" or "no difference"	the effect or difference we want to detect

05 Hypothesis Testing: Logic (2)

Steps

1. State H_0 and H_1.
2. Collect data and calculate the test statistic (z).
3. Compute the p-value $\rightarrow$ probability of observing such an extreme statistic if H_0 were true.
4. Compare the p-value with α (e.g., 0.05) $\rightarrow$ decide whether to reject H_0.

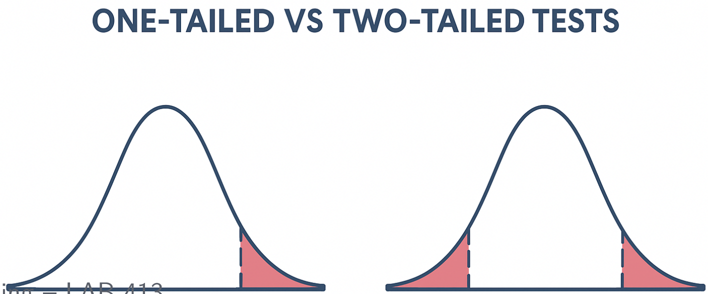
$$H_0,\ H_1 \longrightarrow \text{test statistic } z \longrightarrow p\text{-value} \longrightarrow \text{compare with } \alpha \longrightarrow \text{decision}$$

One- vs. Two-Sided Tests

Key Concepts

The *p*-value depends on the type of alternative hypothesis.

- **One-sided:** tests for an effect in one direction (increase *or* decrease only).
- **Two-sided:** tests for any difference; doubles the one-sided tail area.



Worked Example: Coffee Consumption

Scenario

A national survey reports adults drink an average of **2.8 cups** of coffee per day ($\sigma = 1.1$). You collect data from $n = 50$ Penn students and find they drink **3.2 cups** on average.

At $\alpha = 0.05$, is there evidence that Penn students drink *more* coffee than the national average?

Reference: Z-Score Table

Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1.0	0.84134	0.84375	0.84614	0.84849	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2.0	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3.0	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99896	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976
3.5	0.99977	0.99978	0.99978	0.99979	0.99980	0.99981	0.99981	0.99982	0.99983	0.99983
3.6	0.99984	0.99985	0.99985	0.99986	0.99986	0.99987	0.99987	0.99988	0.99988	0.99989
3.7	0.99989	0.99990	0.99990	0.99990	0.99991	0.99991	0.99992	0.99992	0.99992	0.99992
3.8	0.99993	0.99993	0.99993	0.99994	0.99994	0.99994	0.99994	0.99995	0.99995	0.99995
3.9	0.99995	0.99995	0.99996	0.99996	0.99996	0.99996	0.99996	0.99996	0.99997	0.99997

Source: statisticsbyjim.com/hypothesis-testing/z-table/

Worked Example: Solution

Given

$$\begin{aligned} \mu_0 &= 2.8 \text{ cups/day} \\ \bar{x} &= 3.2 \text{ cups/day} \\ \sigma &= 1.1 \\ n &= 50 \\ \alpha &= 0.05 \end{aligned}$$

Test

$$\begin{aligned} H_0 &: \mu = 2.8 \quad \text{(No difference)} \\ H_1 &: \mu > 2.8 \quad \text{(drink more)} \end{aligned}$$

$$z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{3.2 - 2.8}{1.1/\sqrt{50}} \approx 2.57$$

$p = 0.0051 < 0.05 \Rightarrow$ **Reject** H_0
Strong evidence Penn students drink more coffee than the national average.

07 Interpreting Significance

Key Concepts

- A "significant" result means our data are *unlikely* if H_0 were true — **not** that the effect is large or important.
- Report exact p-values **and** effect sizes, not just "significant" vs. "not significant."

Students drank an average of **3.2** cups/day, compared to the national mean of **2.8**
(mean difference $= 0.4$ cups, $p = 0.005$, 95% CI [0.13, 0.67]).

08 Statistical vs. Practical Significance

- **Statistical significance:** result unlikely due to chance ($p < \alpha$).
- **Practical significance:** result is large or meaningful in context.
- **Causal significance:** result can be interpreted causally (requires experimental / randomized design).

Type	**Example**	**Interpretation**
Statistical	$p = 0.03$	Evidence against H_0
Practical	Effect $= 0.01$ vs. 2.0	Magnitude matters
Causal	RCT effect	Can infer causality

09 Testing Errors

Key Concepts

- **Type I Error (α):** rejecting a *true* null hypothesis (false positive).
- **Type II Error (β):** failing to reject a *false* null hypothesis (false negative).
- **Power ($1 - \beta$):** probability of correctly rejecting a false null.

High power $\Rightarrow$ a good chance of detecting an effect when it truly exists.

RStudio Practice

Confidence Intervals & One-Sample t-test

See `Lab7.Rmd` / `Lab7_1022.R`