Chapter 9 Inferences Based on Two Samples

Ch 10 $\xrightarrow{3080}$

9.1 z Tests and CIs for a Difference Between Two Population Means

In this section we will explore tests and CIs for the difference $\mu_1 - \mu_2$ of two population means.

Ex: Two medications have different recovery times.

$H_0: \mu_1 - \mu_2 = 0$ (two medications have same mean)

$H_a: \mu_1 - \mu_2 > 0$ ($\mu_1 > \mu_2$)

Assumptions: 1. $X_1, X_2, \ldots, X_m$ is a random sample from a distribution with mean μ_1 and variance σ_1^2.

2. $Y_1, Y_2, \ldots, Y_n$ is a random sample from a distribution with mean μ_2 and variance σ_2^2.

3. X and Y samples are independent.

Our point estimator for $\mu_1 - \mu_2$ is $\overline{X} - \overline{Y}$.

$$E[\overline{X} - \overline{Y}] = E[\overline{X}] - E[\overline{Y}] = \mu_1 - \mu_2$$

$$Var(\overline{X} - \overline{Y}) = Var(\overline{X}) + (-1)^2 Var(\overline{Y}) = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}$$

$$SD(\overline{X} - \overline{Y}) = \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

If both populations are normally distributed and σ_1 and σ_2 are known,

then $$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}} \sim N(0,1)$$

This can be used to do hypothesis tests and CIs. Since we usually don't know σ_1 and σ_2, we won't go into those in detail, but the book has more information.

Large Sample Tests: If we have large samples ($m > 40$ and $n > 40$), then $$Z = \frac{\overline{X} - \overline{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \approx N(0,1).$$

Test Structure:

delta

Hypotheses: $H_0: \mu_1 - \mu_2 = \Delta_0$

$H_a: \mu_1 - \mu_2 > \Delta_0$, $H_a: \mu_1 - \mu_2 < \Delta_0$, or $H_a: \mu_1 - \mu_2 \neq \Delta_0$.

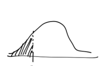
Test Statistic:

$$z = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}, \quad Z \approx N(0,1).$$

P-value: $H_a: \mu_1 - \mu_2 > \Delta_0$, P-value $= 1 - \text{pnorm}(z)$

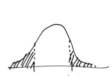
z

$H_a: \mu_1 - \mu_2 < \Delta_0$, P-value $= \text{pnorm}(z)$



z

$H_a: \mu_1 - \mu_2 \neq \Delta_0$, P-value $= 2 \cdot \text{pnorm}(-|z|)$



$-|z|$ $|z|$

Conclusion: If P-value $\leq \alpha$, reject H_0.

If P-value $> \alpha$, fail to reject H_0.

Example 1: We want to test whether Brand A batteries last longer on average than Brand B batteries. We obtained random samples of the lifetimes in hours of the two brands of batteries under certain conditions.

Brand	Sample Size	Sample Mean	Sample SD
A	50	121	12
B	70	117	13

Does this data provide strong evidence that Brand A batteries last longer than Brand B batteries? Test with $\alpha = 0.05$.

Let μ_1 be the average lifetime for Brand A and μ_2 for Brand B.

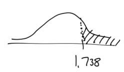
Hypotheses: $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 > 0$.

Test Statistic: $$z = \frac{\bar{X} - \bar{Y} - 0}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} = \frac{121 - 117}{\sqrt{\frac{12^2}{50} + \frac{13^2}{70}}} \approx 1.738$$

P-value: P-value = pnorm(1.738, lower.tail = FALSE)

≈ 0.041.



Conclusion: Since P-value $\leq \alpha$, there is enough evidence to reject H_0, and conclude that Brand A batteries do last longer on average.

Using a Comparison to Identify Causality:

We may want to compare the effects of two different treatments, or treatment vs control (no treatment).

If the individuals or objects used in the comparison are not assigned at random to the two groups, the study is called observational.

Ex: One group had previous infection, and the other group didn't.

It is difficult to draw conclusions about causality based on observational studies:

Ex: People's behavior or economic differences complicate what exactly caused the differences.

A lurking variable could be causing both effects.

A randomized controlled experiment results when investigators randomly assign subjects to the two treatments. This gives more confidence that statistical significance in the experiment comes from the difference in the treatments.

Ex: Drug trials, randomly assigned either drug or placebo.

Large Sample CI for $\mu_1 - \mu_2$

When $m > 40$ and $n > 40$,

$$Z = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \approx N(0,1).$$

$$P\left(-z_{\frac{\alpha}{2}} < \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} < z_{\frac{\alpha}{2}}\right) \approx 1 - \alpha$$

Solving algebraically for $\mu_1 - \mu_2$ in the center, we get the CI.

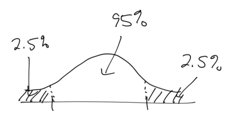
If $m > 40$, $n > 40$ then a $100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$ is

$$\bar{x} - \bar{y} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}} .$$

Example 2: Two versions of a standardized test are being studied. Random samples of students take the two versions:

Test Version	sample size	sample mean	sample SD
A	100	44	8
B	90	40	7

Find a 95% CI for $\mu_1 - \mu_2$, the difference in average scores on the two versions. Is it plausible that $\mu_1 = \mu_2$?



$$z_{\frac{\alpha}{2}} = \text{qnorm}(0.975) \approx 1.96$$

$$\bar{x} - \bar{y} \pm z_{\frac{\alpha}{2}} \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

$$44 - 40 \pm 1.96 \sqrt{\frac{8^2}{100} + \frac{7^2}{90}}$$

$\boxed{(1.867,\ 6.133)}$ is our 95% CI for $\mu_1 - \mu_2$.

If $\mu_1 = \mu_2$, then $\mu_1 - \mu_2 = 0$ which is not in the CI. So it's not plausible that $\mu_1 = \mu_2$.

9.2 The Two-Sample t-Test and CI:

In this section we develop a test and CI for $\mu_1 - \mu_2$ when the populations are both approximately normal. (Small sample sizes are okay.)

When the population distributions are both normal,

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \quad \text{has approximately a } t \text{ distribution with}$$

$$\nu \approx \frac{\left(\frac{S_1^2}{m} + \frac{S_2^2}{n}\right)^2}{\frac{\left(\frac{S_1^2}{m}\right)^2}{m-1} + \frac{\left(\frac{S_2^2}{n}\right)^2}{n-1}} \quad \text{degrees of freedom (round down to nearest integer).}$$

A $100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$ is

$$\bar{x} - \bar{y} \pm t_{\frac{\alpha}{2}, \nu} \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

The t-Test for $\mu_1 - \mu_2$ has this structure:

Hypotheses: $H_0: \mu_1 - \mu_2 = \Delta_0$ vs $H_a: \mu_1 - \mu_2 \begin{smallmatrix} > \\ < \\ \neq \end{smallmatrix} \Delta_0$.

Test Statistic: $t = \dfrac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$

P-value: Areas under the t_ν curve.

Conclusion: reject H_0 when P-value $\leq \alpha$, else fail to reject H_0.

Example 1: Suppose μ_1 and μ_2 are the mean stopping distances at 50 mph for cars of a certain type equipped with two different braking systems. We obtain the following data:

Brake Type	Sample Size	Sample Mean	Sample SD
A	8	120	5.1
B	7	115	5.2

Assume both populations are normal. Find a 95% CI for $\mu_1 - \mu_2$.

Perform a test with $\alpha = 0.05$ of whether or not $\mu_1 = \mu_2$.

A $100(1-\alpha)\%$ CI for $\mu_1 - \mu_2$ is

$$\bar{x} - \bar{y} \pm t_{\frac{\alpha}{2}, \nu} \sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}, \quad \nu = \frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{\left(\frac{s_1^2}{m}\right)^2}{m-1} + \frac{\left(\frac{s_2^2}{n}\right)^2}{n-1}} = \frac{\left(\frac{5.1^2}{8} + \frac{5.2^2}{7}\right)^2}{\frac{\left(\frac{5.1^2}{8}\right)^2}{8-1} + \frac{\left(\frac{5.2^2}{7}\right)^2}{7-1}} \approx 12.66 \searrow 12$$

$$t_{\frac{\alpha}{2}, \nu} = qt(0.975, df = 12) \approx 2.179$$

$$120 - 115 \pm 2.179 \cdot \sqrt{\frac{5.1^2}{8} + \frac{5.2^2}{7}}$$

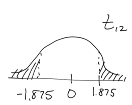
$(-0.812, 10.812)$ is our 95% CI for $\mu_1 - \mu_2$.

Hypotheses: $H_0: \mu_1 - \mu_2 = 0$

$H_a: \mu_1 - \mu_2 \neq 0$

Test Statistic: $t = \dfrac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}} = \dfrac{120 - 115 - 0}{\sqrt{\frac{5.1^2}{8} + \frac{5.2^2}{7}}} \approx 1.875$

P-value: $2 \cdot pt(-1.875, df = 12) \approx 0.0853$



Conclusion: We fail to reject H_0 because p-value $> \alpha$. There is insufficient evidence to conclude the means are different.