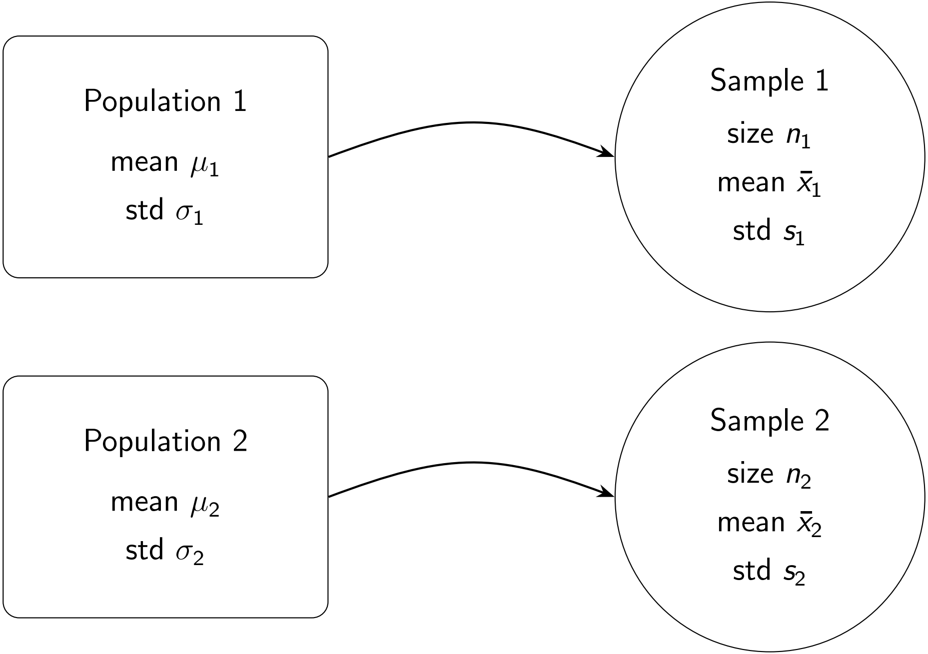
Chapter 9

Hypothesis Tests: Two Samples

Suppose now that there are two populations with means μ_1 and μ_2, and standard deviations σ_1 and σ_2, respectively.

Suppose that we take a random sample of size n_1 from population 1 and a random sample of size n_2 from population 2.

We will denote the sample means by $\bar{x}_1$ and $\bar{x}_2$, and sample standard deviations by s_1 and s_2, respectively.


Population 1
mean μ_1
std σ_1
Sample 1
size n_1
mean $\bar{x}_1$
std s_1
Population 2
mean μ_2
std σ_2
Sample 2
size n_2
mean $\bar{x}_2$
std s_2

Independent Sampling

In this chapter, we will assume that these two samples are independent.

If the samples are NOT independent, our formulas will not work. For example:

- If the same subjects are measured twice: for example, measuring the mood of the subjects before and after watching a comedy.
- If the observations are paired or correlated, such as fathers' heights and their sons' heights.

What can we say about the distribution of $\overline{X}_1 - \overline{X}_2$, the difference between the two sample means?

Since $\overline{X}_1$ and $\overline{X}_2$ are two independent random variables, and since

$$\sigma_{\bar{x}_1} = \frac{\sigma_1}{\sqrt{n_1}}, \quad \sigma_{\bar{x}_2} = \frac{\sigma_2}{\sqrt{n_2}}$$

we have

$$E(\overline{X}_1 - \overline{X}_2) = E(\overline{X}_1) - E(\overline{X}_2) = \mu_1 - \mu_2$$

$$\mathrm{Var}(\overline{X}_1 - \overline{X}_2) = \mathrm{Var}(\overline{X}_1) + \mathrm{Var}(\overline{X}_2) = \sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2$$

$$\Longrightarrow \sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\sigma_{\bar{x}_1}^2 + \sigma_{\bar{x}_2}^2}$$

$$= \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Hypothesis Tests: Two Samples

Aim: Comparing two population means, i.e. making inferences about $\mu_1 - \mu_2$.

Two-sample z-test for the difference of means (for large samples)

Two-sample z-test for the difference of means (for large samples)

Assumptions:

- The two samples are random and independent,
- $n_1 \geq 30$, and
- $n_2 \geq 30$

Two-sample z-test for the difference of means (for large samples)

- null hypothesis: $H_0 : \mu_1 - \mu_2 = 0$
- alternative hypothesis:
  - One-tailed, upper-tailed: $H_1 : \mu_1 - \mu_2 > 0$
  - One-tailed, lower-tailed: $H_1 : \mu_1 - \mu_2 < 0$
  - Two-tailed: $H_1 : \mu_1 - \mu_2 \neq 0$
- Test statistic:

$$z_c = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \quad \text{if } \sigma_1, \sigma_2 \text{ are known}$$

$$z_c = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad \text{if } \sigma_1, \sigma_2 \text{ are unknown}$$

- Decision:

$$\text{One-tailed case: if } |z_c| > z_\alpha, \text{ we reject } H_0$$

$$\text{Two-tailed case: if } |z_c| > z_{\alpha/2}, \text{ we reject } H_0$$

Example

In a school, some students use an external paid resource called "mini-courses."

A random sample of 121 students who do not use mini-courses has a mean GPA of 2.9 and a standard deviation of 1.1.

A random sample of 32 students who use mini-courses has a mean GPA of 2.3 and a standard deviation of 0.8.

At significance level $\alpha = 0.01$, test whether students who do not use mini-courses have a higher GPA on average.

Example (cont'd)

Let μ_1 be the mean GPA of students who do not use mini-courses, and let μ_2 be the mean GPA of students who use mini-courses.

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 > 0$$

$$z_c = \frac{2.9 - 2.3}{\sqrt{\frac{1.1^2}{121} + \frac{0.8^2}{32}}} = \frac{0.6}{\sqrt{0.01 + 0.02}} \approx 3.46.$$

$$z_\alpha = z_{0.01} = 2.326$$

Since $|z_c| = 3.46 > 2.326 = z_\alpha$, we reject H_0.
There is enough statistical evidence at significance level 0.01 to conclude that students who do not use mini-courses have a higher GPA on average than students who use mini-courses.

Example

A company has two factories producing the same type of battery.

A random sample of 100 batteries from Factory A has a mean lifetime of 42.3 hours and a standard deviation of 5.4 hours.

A random sample of 90 batteries from Factory B has a mean lifetime of 40.8 hours and a standard deviation of 4.9 hours.

At significance level $\alpha = 0.05$, test whether the average lifetimes of batteries produced by the two factories are different.

Example (cont'd)

Let μ_1 be the mean lifetime of batteries produced by Factory A, and let μ_2 be the mean lifetime of batteries produced by Factory B.

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0$$

$$z_c = \frac{42.3 - 40.8}{\sqrt{\frac{5.4^2}{100} + \frac{4.9^2}{90}}} = \frac{1.5}{\sqrt{0.2916 + 0.2668}} \approx 2.01.$$

$$z_{\alpha/2} = z_{0.025} = 1.96$$

Since $|z_c| = 2.01 > 1.96 = z_{\alpha/2}$, we reject H_0. There is enough statistical evidence at significance level 0.05 to conclude that the average lifetimes of batteries produced by the two factories are different.

Two-sample z-test for the difference of means (for small samples, σ_1 and σ_2 known)

Two-sample z-test for the difference of means (for small samples, σ_1 and σ_2 known)

Assumptions:

- The two samples are random and independent,
- $n_1 < 30$ or $n_2 < 30$,
- the populations are normally distributed,
- σ_1 and σ_2 are known

Two-sample z-test for the difference of means (for small samples, σ_1 and σ_2 known)

- null hypothesis: $H_0 : \mu_1 - \mu_2 = 0$
- alternative hypothesis:
  - One-tailed, upper-tailed: $H_1 : \mu_1 - \mu_2 > 0$
  - One-tailed, lower-tailed: $H_1 : \mu_1 - \mu_2 < 0$
  - Two-tailed: $H_1 : \mu_1 - \mu_2 \neq 0$
- Test statistic:

$$z_c = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

- Decision:

$$\text{One-tailed case: if } |z_c| > z_\alpha, \text{ we reject } H_0$$
$$\text{Two-tailed case: if } |z_c| > z_{\alpha/2}, \text{ we reject } H_0$$

Example

A delivery company wants to compare the average delivery times of two routes.

Delivery times are assumed to be normally distributed. From past records, the company knows that the population standard deviations are $\sigma_1 = 5$ minutes for Route A and $\sigma_2 = 6$ minutes for Route B.

A random sample of 10 deliveries from Route A has a mean delivery time of 34.2 minutes.

A random sample of 12 deliveries from Route B has a mean delivery time of 39.1 minutes.

At significance level $\alpha = 0.05$, test whether Route A has a lower average delivery time than Route B.

Example (cont'd)

Let μ_1 be the mean delivery time of Route A, and let μ_2 be the mean delivery time of Route B.

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 < 0$$

$$z_c = \frac{34.2 - 39.1}{\sqrt{\frac{5^2}{10} + \frac{6^2}{12}}} = \frac{-4.9}{\sqrt{2.5 + 3}} \approx -2.09.$$

$$z_\alpha = z_{0.05} = 1.645$$

Since $|z_c| = 2.09 > 1.645 = z_\alpha$, we reject H_0.

There is enough statistical evidence at significance level 0.05 to conclude that Route A has a lower average delivery time than Route B.

Two-sample t-test for the difference of means (for small samples, common σ unknown)

Two-sample t-test for the difference of means (for small samples, common σ unknown)

Assumptions:

- The two samples are random and independent,
- $n_1 < 30$ or $n_2 < 30$,
- the populations are normally distributed,
- $\sigma_1 = \sigma_2 = \sigma$, and
- σ is unknown

If $\sigma_1 = \sigma_2 = \sigma$, then the standard error becomes

$$\sigma_{\bar{x}_1 - \bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} = \sqrt{\frac{\sigma^2}{n_1} + \frac{\sigma^2}{n_2}}$$

$$= \sigma\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Since σ is unknown, we estimate it by the *pooled standard deviation*:

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}.$$

Under these assumptions, we use the t-distribution with

$$\text{degrees of freedom} = (n_1 - 1) + (n_2 - 1) = n_1 + n_2 - 2$$

The test statistic is

$$t_c = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

where

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

Two-sample t-test for the difference of means (for small samples, common σ unknown)

- null hypothesis: $H_0 : \mu_1 - \mu_2 = 0$
- alternative hypothesis:
  - One-tailed, upper-tailed: $H_1 : \mu_1 - \mu_2 > 0$
  - One-tailed, lower-tailed: $H_1 : \mu_1 - \mu_2 < 0$
  - Two-tailed: $H_1 : \mu_1 - \mu_2 \neq 0$
- Test statistic:

$$t_c = \frac{\bar{x}_1 - \bar{x}_2}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \text{where} \quad s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

- Decision:

$$\text{One-tailed case: if } |t_c| > t_\alpha, \text{ we reject } H_0$$
$$\text{Two-tailed case: if } |t_c| > t_{\alpha/2}, \text{ we reject } H_0$$

where the t-curve is based on $n_1 + n_2 - 2$ degrees of freedom.

Example

One sample from a normally distributed population has values

$$3,\ 4,\ 6,\ 7,\ 7,\ 13,\ 16.$$

Another sample from another normally distributed population has values

$$1,\ 2,\ 5,\ 10,\ 17.$$

Assume that the two populations have the same standard deviation.

At significance level $\alpha = 0.05$, test whether the first population has a higher average than the second population.

Example (cont'd)

Let μ_1 be the mean of the first population, and let μ_2 be the mean of the second population.

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 > 0$$

$$\bar{x}_1 = \frac{56}{7} = 8, \qquad \bar{x}_2 = \frac{35}{5} = 7.$$

$$n_1 = 7, \qquad n_2 = 5, \qquad \text{degrees of freedom} = 7 + 5 - 2 = 10.$$

Example (cont'd)

$$s_1 = \sqrt{\frac{(-5)^2 + (-4)^2 + (-2)^2 + (-1)^2 + (-1)^2 + 5^2 + 8^2}{6}} = \sqrt{\frac{136}{6}}.$$

$$s_2 = \sqrt{\frac{(-6)^2 + (-5)^2 + (-2)^2 + 3^2 + 10^2}{4}} = \sqrt{\frac{174}{4}}.$$

$$s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{6 \cdot \frac{136}{6} + 4 \cdot \frac{174}{4}}{10}} = \sqrt{\frac{136 + 174}{10}} = \sqrt{31}.$$

Example (cont'd)

$$t_c = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{8 - 7}{\sqrt{31}\sqrt{\frac{1}{7} + \frac{1}{5}}} = \frac{1}{\sqrt{31}\sqrt{\frac{12}{35}}} \approx 0.307.$$

$$t_\alpha = t_{0.05} = 1.812 \qquad \text{with 10 degrees of freedom.}$$

Since

$$|t_c| = 0.307 < 1.812 = t_\alpha,$$

we fail to reject H_0.

There is not enough statistical evidence at significance level 0.05 to conclude that the first population has a higher average than the second population.

Example

A company compares the average production times of two machines.

The production times are assumed to be normally distributed, and the two populations are assumed to have the same standard deviation.

A random sample of 12 products from Machine A has a mean production time of 52.4 seconds and a standard deviation of 4.1 seconds.

A random sample of 10 products from Machine B has a mean production time of 47.8 seconds and a standard deviation of 3.8 seconds.

At significance level $\alpha = 0.05$, test whether Machine A has a higher average production time than Machine B.

Example (cont'd)

Let μ_1 be the mean production time of Machine A, and let μ_2 be the mean production time of Machine B.

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 > 0$$

$$\text{degrees of freedom} = n_1 + n_2 - 2 = 12 + 10 - 2 = 20.$$

$$s_p = \sqrt{\frac{(12-1)(4.1)^2 + (10-1)(3.8)^2}{12+10-2}} \approx 3.97.$$

Example (cont'd)

$$t_c = \frac{52.4 - 47.8}{3.97\sqrt{\frac{1}{12} + \frac{1}{10}}} \approx 2.71.$$

$$t_\alpha = t_{0.05} = 1.725 \quad \text{with 20 degrees of freedom.}$$

Since $|t_c| = 2.71 > 1.725 = t_\alpha$, we reject H_0.

There is enough statistical evidence at significance level 0.05 to conclude that Machine A has a higher average production time than Machine B.

Example

A bank wants to compare the average waiting times at two branches.

Waiting times are assumed to be normally distributed, and the two populations are assumed to have the same standard deviation.

A random sample of 12 customers from Branch A has a mean waiting time of 8.0 minutes and a standard deviation of 2.6 minutes.

A random sample of 10 customers from Branch B has a mean waiting time of 6.9 minutes and a standard deviation of 2.3 minutes.

At significance level $\alpha = 0.05$, test whether the average waiting times at the two branches are different.

Example (cont'd)

Let μ_1 be the mean waiting time at Branch A, and let μ_2 be the mean waiting time at Branch B.

$$H_0 : \mu_1 - \mu_2 = 0$$
$$H_1 : \mu_1 - \mu_2 \neq 0$$

$$\text{degrees of freedom} = n_1 + n_2 - 2 = 12 + 10 - 2 = 20.$$

Example (cont'd)

$$s_p = \sqrt{\frac{(12-1)(2.6)^2 + (10-1)(2.3)^2}{12+10-2}} = \sqrt{\frac{11 \cdot 6.76 + 9 \cdot 5.29}{20}} \approx 2.47.$$

$$t_c = \frac{8.0 - 6.9}{2.47\sqrt{\frac{1}{12} + \frac{1}{10}}} \approx 1.04.$$

$$t_{\alpha/2} = t_{0.025} = 2.086 \qquad \text{with 20 degrees of freedom.}$$

Since

$$|t_c| = 1.04 < 2.086 = t_{\alpha/2},$$

we fail to reject H_0.

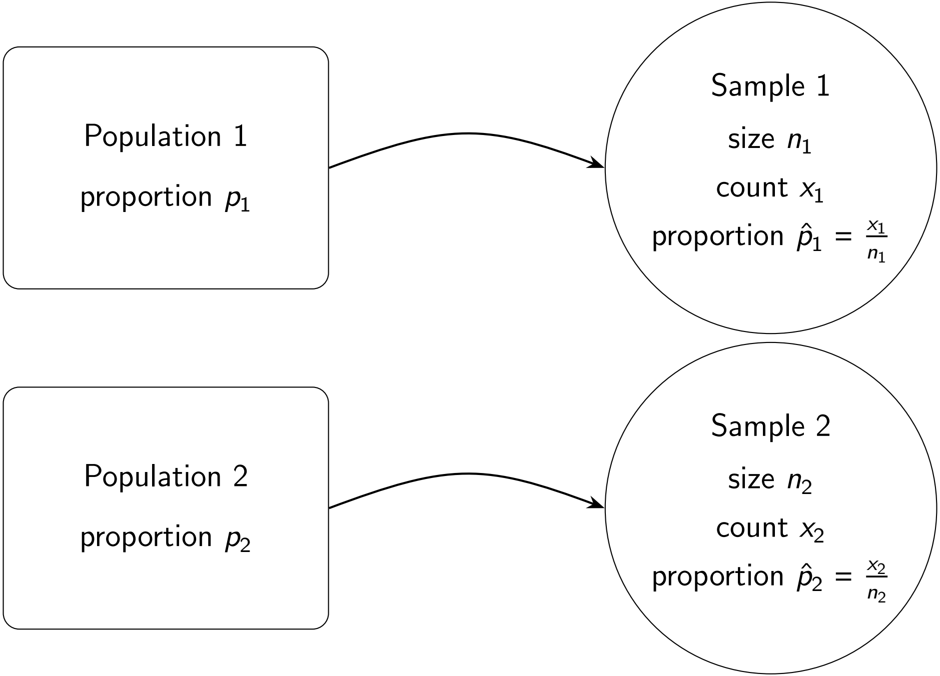
There is not enough statistical evidence at significance level 0.05 to conclude that the average waiting times at the two branches are different.

Side Note

For the case $\sigma_1 \neq \sigma_2$, see page 473 of the textbook.

(This case is not included in the scope of this course.)

Two-sample z-test for the difference of proportions


Population 1
proportion p_1
Sample 1
size n_1
count x_1
proportion $\hat{p}_1 = \frac{x_1}{n_1}$
Population 2
proportion p_2
Sample 2
size n_2
count x_2
proportion $\hat{p}_2 = \frac{x_2}{n_2}$

Two-sample z-test for the difference of proportions

Assumptions:

- The two samples are random and independent,
- $x_1 = n_1 \hat{p}_1 \geq 15$,
- $n_1 - x_1 = n_1(1 - \hat{p}_1) \geq 15$,
- $x_2 = n_2 \hat{p}_2 \geq 15$, and
- $n_2 - x_2 = n_2(1 - \hat{p}_2) \geq 15$

Here, x_1 and x_2 are the counts of individuals or items with the characteristic we are interested in.

Two-sample z-test for the difference of proportions

- null hypothesis: $H_0 : p_1 - p_2 = 0$
- alternative hypothesis:
  - One-tailed, upper-tailed: $H_1 : p_1 - p_2 > 0$
  - One-tailed, lower-tailed: $H_1 : p_1 - p_2 < 0$
  - Two-tailed: $H_1 : p_1 - p_2 \neq 0$
- Test statistic:

$$z_c = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}_0(1-\hat{p}_0)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad \text{where} \quad \hat{p}_0 = \frac{x_1 + x_2}{n_1 + n_2}$$

- Decision:

$$\text{One-tailed case: if } |z_c| > z_\alpha, \text{ we reject } H_0$$
$$\text{Two-tailed case: if } |z_c| > z_{\alpha/2}, \text{ we reject } H_0$$

Example

Among 113 boys tested for a virus, 34 have antibodies.

Among 139 girls tested for the same virus, 54 have antibodies.

At significance level $\alpha = 0.05$, test whether the proportion of girls with antibodies is higher than the proportion of boys with antibodies.

Example (cont'd)

Let p_1 be the proportion of girls with antibodies, and let p_2 be the proportion of boys with antibodies.

$$H_0 : p_1 - p_2 = 0$$
$$H_1 : p_1 - p_2 > 0$$

$$\hat{p}_1 = \frac{54}{139} \approx 0.3885, \qquad \hat{p}_2 = \frac{34}{113} \approx 0.3009.$$

$$\hat{p}_0 = \frac{54 + 34}{139 + 113} = \frac{88}{252} \approx 0.3492.$$

Example (cont'd)

$$z_c = \frac{0.3885 - 0.3009}{\sqrt{0.3492(1 - 0.3492)\left(\frac{1}{139} + \frac{1}{113}\right)}} \approx 1.45.$$

$$z_\alpha = z_{0.05} = 1.645.$$

Since

$$|z_c| = 1.45 < 1.645 = z_\alpha,$$

we fail to reject H_0.

There is not enough statistical evidence at significance level 0.05 to conclude that the proportion of girls with antibodies is higher than the proportion of boys with antibodies.