10 Hypothesis Testing with Two Independent Samples

Previously we have studied:

- the one-sample t-test for population mean μ, using the information provided by a single sample;
- the one-sample z-test for population proportion p, based on one sample;
- the matched pairs t-test based on two observations on each (or identical) subject (which reduces to the one-sample t-test for differenced data).

This week we consider an extension of the above work and study methods to compare *two population means* and *two population proportions*, both based on two independent samples from two populations.

10.1 Two-sample t-test comparing two population means (P.145-149)

10.1.1 Introduction

In every area of human activity, new procedures are invented and existing techniques are revised. Advances occur whenever a new technique is proved to be better than the old. Hence we need to test whether the new method is better than the old one but based on a new experimental design other than matched pairs design.

This section develops a popular statistical test that compares the means of two *independent* populations.

A motivational example: Suppose that there are two types of food available for milking cows. A farmer wishes to test which type of food helps cows to produce more yield of milk.

An experiment: The farmer can select two independent groups of cows who produce similar milk yields. One group is given the food A and the other group is given the food B. After one week, the farmer calculates the means and standard deviations of milk yields for each group and then use his knowledge to decide the type of food which gives better yield.

Further examples:

1. Compare the average age at first marriage of females in two ethnic groups.

2. Compare the average efficiency of two brands of fertilisers.

3. Compare the average marks of statistics students at USYD and UNSW

Note: This type of design do comparison using two independent samples rather than matched pairs. This type of design is necessary in some situations when matched pairs from similar or same subjects are more difficult to form, for example, in the comparison of two ethnic groups where human characteristics are difficult to match.

A statistical test, the *two-sample t-test*, for such comparisons can be developed under the following assumptions:

10.1.2 Test assumptions

1. Two populations are independent and *normally distributed* populations with *equal variances*.

2. Two independent samples are drawn (one from each population).

Notation

Populations:

Two independent populations with means μ_1 and μ_2 and the same variances σ^2.

Population	Variable	Population Mean	Population Variance
1	X_1	μ_1	σ^2
2	X_2	μ_2	σ^2

Samples: Take two independent samples of sizes n_1 and n_2 from each population and calculate the following statistics:

Sample	Variable	Sample Mean	Sample Variance
1	x_1	$\bar{x}_1$	s_1^2
2	x_2	$\bar{x}_2$	s_2^2

10.1.3 The hypothesis

In this (independent) two-sample problem, the null hypothesis of interest is:

$$H_0 : \mu_1 = \mu_2 \text{ or } H_0 : \mu_1 - \mu_2 = 0.$$

The alternative hypothesis, H_1 will be set up according to the specific problem of interest and select one from:

$$H_1 : \mu_1 > \mu_2 \text{ or } H_1 : \mu_1 - \mu_2 > 0 \text{ (one-sided);}$$

$$H_1 : \mu_1 < \mu_2 \ \text{ or } \ H_1 : \mu_1 - \mu_2 < 0 \text{ (one-sided);}$$

$$H_1 : \mu_1 \neq \mu_2 \ \text{ or } \ H_1 : \mu_1 - \mu_2 \neq 0 \text{ (two-sided).}$$

Note: As we have two sample variances s_1^2 and s_1^2, we need to combine them to form a single variance in order to develop a t test. This can be done by combining or pooling s_1^2 and s_1^2 as given below:

10.1.4 Combined or Pooled Variance

It can be shown that the best combination of s_1^2 and s_1^2 to produce the common variance is given by

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

Remarks:

- This combined variance s_p^2 is called the *pooled variance.*
- The pooled variance is simply a weighted average of the two individual sample variances, weighted by their df.

10.1.5 The Test Statistic

It can be proved that under the null hypothesis $H_0 : \mu_1 - \mu_2 = 0$, the test statistic

$$t = \frac{\overline{X}_1 - \overline{X}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

is t-distributed with $n_1 + n_2 - 2$ degrees of freedom.

Note: The corresponding df for this two sample problem is 2 less than the total sample size of $n_1 + n_2$. Compare this with the one sample t-test

$$t = \frac{\overline{X} - \mu_0}{s/\sqrt{n}} \sim t_{n-1}$$

is t-distributed with $n-1$ degrees of freedom.

Example: Two independent samples have been taken from two independent normal populations. The observations are:

Sample 1: 8, 5, 7, 6, 9, 7
Sample 2: 2, 6, 4, 7, 6.

Find an estimate of the combined variance (or pooled variance).

Solution:

Sample 1: $n_1 = 6$. $\bar{x}_1 = 7$, $s_1^2 = 2$.
Sample 2: $n_2 = 5$. $\bar{x}_2 = 5$, $s_2^2 = 4$.

Therefore, the combined or pooled variance (estimate) is:

$$s_p^2 = \frac{\qquad\qquad}{\qquad\qquad\qquad\qquad\qquad}$$

Example (cont): State the distribution of $t = \frac{\bar{X}_1 - \bar{X}_2}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$ under H_0.

Solution: Since the $df = 6 + 5 - 2 = 9$

$$t = \frac{\overline{X}_1 - \overline{X}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim \_\_$$

Example: Using the sample information, calculate the value of test statistic.

Solution: $t_0 = \dfrac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} =$ ____________________

Example: Test the null hypothesis against $H_1 : \mu_1 - \mu_2 > 0$.

Solution: P-value = ____________________________________

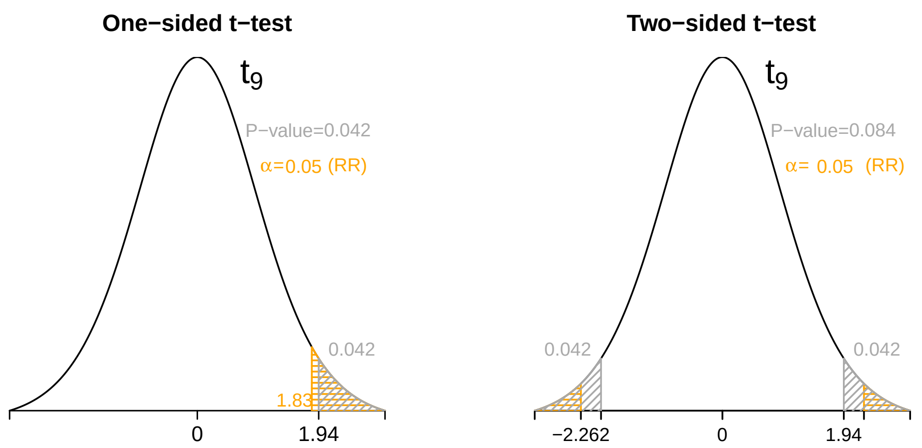
and therefore _________________________.

Example: Test the null hypothesis $H_1 : \mu_1 - \mu_2 = 0$ against the alternative $H_1 : \mu_1 - \mu_2 \neq 0$.

Solution: Sincve the alternative hypothesis is two sided, the corresponding P-value is calculated (as before) as:

____

Therefore ____________________________.



Note: The exact P-value of 0.084 (2-sided) can be obtained from `R` using `2*(1-pt(1.94,9))` and the critical values of 1.833 and 2.262 from t-table.

Read: Example 8.10 (p. 147)

Example: A feeding test is conducted on a herd of 25 dairy cows to compare two diets, A and B. A sample of 13 cows randomly selected from the herd are fed diet A and the remaining cows are fed with diet B. From observations made over a three-week period, the average daily milk production (in L) is recorded for each cow:

	Milk Yield (in L)												
Diet A (x_1)	44	44	56	46	47	38	58	53	49	35	46	30	41
Diet B (x_2)	35	47	55	29	40	39	32	41	42	57	51	39	

Assume these two samples come from independent normally distributed populations with equal variances σ^2.

(i) Find the mean and the sd for each sample.

(ii) Find an estimate of the 'pooled variance' s_p^2, which estimates the common variance σ^2

(iii) Perform the two-sample t-test to investigate the evidence of a *difference* in true mean milk yields for the two diets.

Solution:

(i) $\bar{x}_1 = 45.15, s_1 = 7.998, n_1 = 13$ for A
$\bar{x}_2 = 42.25, s_2 = 8.740, n_2 = 12$ for B

(ii) The 'pooled' sample variance is

$$
\begin{aligned}
s_p^2 &= \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1+n_2-2)} \\
&= \frac{\qquad}{\rule{6cm}{0.4pt}} \\
&= \rule{1cm}{0.4pt}
\end{aligned}
$$

(iii) The two-sample t-test:

1. Hypotheses: As we want to test whether there is a difference in milk yields, we have a *two-sided* alternatives:

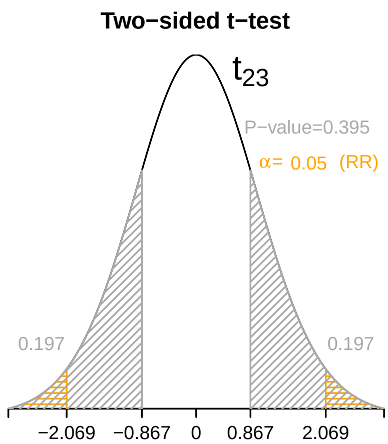
$$H_0 : \mu_1 = \mu_2 \text{ against } H_1 : \mu_1 \neq \mu_2.$$

2. Test Statistic: Under H_0,

$$t_0 = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_$$

3. P-value: ________________

4. Conclusion: ________________________________ ________. There is no significant difference between the two diets.



10.1.6 Confidence interval (CI) for $\mu_1 - \mu_2$

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

$$\bar{x}_1 - \bar{x}_2 \pm t_{n_1+n_2-2,\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

To test $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$ at α significant level:

If $0 \in$ CI, then the data are consistent with H_0.
If $0 \notin$ CI, then there is evidence against H_0.

Example: For the previous example, the 95% CI for $\mu_1 - \mu_2$ is

$$\bar{x}_1 - \bar{x}_2 \mp t_{n_1+n_2-2,\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$= \underline{\hspace{10cm}}$$

Since the CI includes 0, the data are consistent with H_0.

Exercises: Additional problems for practice:

P.152: Q8.1 to Q8.3, Q8.7, Q8.8, Q8.12, Q8.13

P.153: Q8.25, Q8.26.

P.154: Q8.37, Q8.40, 8.45.

Some answers:

Q8.1: $\bar{x} = 20, \quad s = 35, \quad n = 20 \quad \Rightarrow \quad t_{19,0.025} = 2.093$. A 95% CI is

$$\left(20 - 2.093\frac{35}{\sqrt{20}}, 20 + 2.093\frac{35}{\sqrt{20}}\right) = (3.62, 36.38)$$

Since this interval excludes 0, we have evidence against H_0 that the mean change is 0 at the 0.05 significance level if tested against the two-sided alternative.

Q8.7: $s_p^2 = 0.514$; $t_0 = -1.314$; P-value is large; Consistent with H_0.
Q8.12: $s_p^2 = 9.809$; $t_0 = 1.49$; P-value is large; Consistent with H_0.
Q8.13: t=1.49, df=23; Q8.25: t=-3.37, df=9;
Q8.26: t=-1.83, df=29; Q8.37: use two-sample t test;
Q8.40: t=-3.68, df=18; Q8.45: a 95% CI is (2.2,7.2)

10.2 Two-sample z-test for comparing two population proportions (P.157-161)

In some life science problems we need to test whether the two population proportions for a particular attribute are equal.

Motivating example: Suppose that a federal member of the parliament wishes to test whether two suburbs in his electorate have the same unemployment rate.

To test this, the member can take two independent samples (one from each suburb) and calculate the proportion of the unemployment. However, these two proportions can not show whether any difference between them is sufficiently large to support his claim. Therefore, we need to develop a proper statistical test.

10.2.1 Assumption

1. Two independent samples
2. *Both* sample sizes are large: both $n_1 \geq 30$ and $n_2 \geq 30$.

10.2.2 Hypotheses

Null hypothesis of interest: As we would like to compare two proportions p_1 and p_2 for each of the populations,

$$H_0 : p_1 = p_2 \text{ or equivalently } H_0 : p_1 - p_2 = 0$$

Alternative hypothesis: Depending on the specific problem, it can be:

$$\begin{aligned} H_1 :\quad & p_1 > p_2 \text{ or equivalently } H_1 : p_1 - p_2 > 0 \text{ (one-sided)},\\ H_1 :\quad & p_1 < p_2 \text{ or equivalently } H_1 : p_1 - p_2 < 0 \text{ (one-sided)},\\ H_1 :\quad & p_1 \neq p_2 \text{ or equivalently } H_1 : p_1 - p_2 \neq 0 \text{ (two-sided)}. \end{aligned}$$

To develop a suitable test statistic, we need a single estimate for the proportion based on two independent samples under H_0 of equal proportions. This combined or pooled estimate is obtained using the formula given below:

10.2.3 Combined or pooled proportion

Suppose that x_1 and x_2 are the number of "successes" in each independent sample, and n_1 and n_2 their respective sample sizes. Under the null hypothesis that two population proportions are equal, we estimate this common proportion using:

$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2},$$

Note: It is clear that

$$\hat{p} = \frac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2},$$

where $\hat{p}_1$ and $\hat{p}_2$ are the estimates of the two proportions based on two independent samples.

Remark: $\hat{p}$ is just a weighted average of the two sample proportions $\hat{p}_1$ and $\hat{p}_2$, weighted by their sample sizes.

10.2.4 The test Statistic

The formula for the test statistic is

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \sim \mathrm{N}(0,1)$$

since under the null hypothesis,

$$\mathrm{Var}(\hat{p}_1 - \hat{p}_2) = \mathrm{Var}(\hat{p}_1) + \mathrm{Var}(\hat{p}_2) = \frac{\hat{p}(1-\hat{p})}{n_1} + \frac{\hat{p}(1-\hat{p})}{n_2} = \hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)$$

and the distribution of z is N(0,1) by Central Limit Theorem when both n_1 and n_2 are large.

Note that the denominator uses the pooled estimate $\hat{p}$ under H_0.

10.2.5 The P-value

The P-value for the probability of observed and more extreme values is calculated as before and it depends on the direction of the alternative hypothesis.

Example 1: A university wants to determine whether a recently hired instructor is working out. Twenty-three out of 30 of Instructor A's students passed a certain test on the first try. In comparison, 57 out of 72 of more experienced Instructor B's students passed the test on the first try. Is Instructor A's success rate *worse* than Instructor B's?

Solution: Let

- p_1 = success rate of Instructor A's students
- p_2 = success rate of Instructor B's students

be the true proportions of interest.

Note: In this two-sample problems it is important to identify which parameter corresponds to which population. Here we have $X_1 = 23$, $n_1 = 30$, $X_2 = 57$ and $n_2 = 72$.

1. Hypotheses: As we want to test whether Instructor A's success rate *worse*, we have a *lower side* alternatives:

$$H_0 : p_1 = p_2 \text{ vs. } H_1 : p_1 < p_2.$$

2. Test statistic:

Preliminary calculations:
Sample proportions are:

$$\hat{p}_1 = \frac{X_1}{n_1} = \underline{\hspace{3cm}}, \qquad \hat{p}_2 = \frac{X_2}{n_2} = \underline{\hspace{3cm}}$$

Combined or the pooled proportion is:

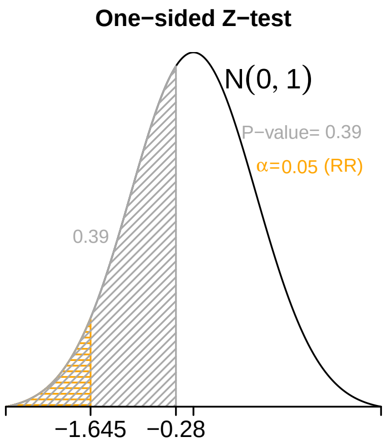
$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \underline{\hspace{6cm}}$$

Hence the test statistic is:

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \underline{\hspace{8cm}}$$

3. P-value: ____

4. Conclusion: ______________________________________
___. That is, there is insufficient evidence that Instructor A is inferior to Instructor B in terms of their student success rate.



Example 2: On October 23, 2009, an outbreak of mumps was reported in Borough Park, Brooklyn. Fifty-seven children were diagnosed with this childhood disease. Surprisingly, 43 of the children had the recommended two doses of MMR vaccine which is supposed to protect against the disease. In the past, from a sample of 100 children with mumps in New York State, 83% of them had the recommended two doses of the vaccine. Test the hypothesis that the MMR vaccination rate for the two groups is *different* at $\alpha = 0.05$.

Solution: Let

- p_1 = proportion of vaccinated children with mumps in Boro Park
- p_2 = proportion of vaccinated children with mumps in NYS

1. Hypotheses: As we want to test whether the rates are *different*, we have a *two-sided* alternatives:

$$H_0 : p_1 = p_2 \text{ vs. } H_1 : p_1 \neq p_2.$$

2. Test statistic:
Preliminary calculations:

$$\hat{p}_1 = \frac{X_1}{n_1} = \underline{\hspace{3cm}}, \qquad \hat{p}_2 = \frac{X_2}{n_2} = \underline{\hspace{3cm}},$$

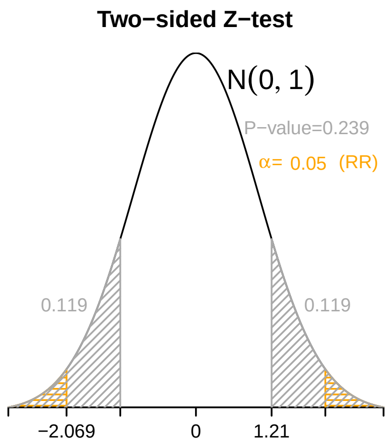
$$\hat{p} = \frac{x_1 + x_2}{n_1 + n_2} = \underline{\hspace{6cm}}$$

Hence the test statistic is:

$$z_0 = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} = \underline{\hspace{9cm}}$$

3. ***P*-value:** ______________________

4. Conclusion: ______________________
___. That is, there is insufficient evidence to conclude that vaccination rate is different among children with mumps in Boro Park and in NYS.



10.2.6 Confidence interval (CI) for $p_1 - p_2$

A $(1-\alpha)100\%$ CI for $p_1 - p_2$ is given by:

$$\hat{p}_1 - \hat{p}_2 \mp z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)},$$

where $z_{1-\alpha/2}$ is the corresponding value from the N(0,1) table with a lower area $1-\alpha/2$.

Example: Find a 95% CI for $p_1 - p_2$ for example 1.

Solution: The 95% CI for $p_1 - p_2$ is:

$$\hat{p}_1 - \hat{p}_2 \mp z_{1-\alpha/2}\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}$$

$$= \underline{\hspace{10cm}}$$

$$= \underline{\hspace{8cm}}$$

Since the 95% CI contain 0, there is no significant difference between the two success rates. This result agrees with the result from hypotheses testing.

Exercise: Find a 95% CI for $p_1 - p_2$ for example 2.

Answer: (-0.2051, 0.0539)