Chapter 3

Hypothesis tests for two means

3.1 Introduction

Last week you were introduced to the concept of hypothesis testing in statistics, and we considered hypothesis tests for the mean if we have a single sample drawn from a single population. If we have *two* independent random samples from *two* populations, we can compare the two sample means in a test for two means (*c.f.* comparing one sample mean to a *proposed value* in the one–sample case). We use the same framework for hypothesis testing as for the one–sample tests:

1. State the **null hypothesis**, H_0;
2. State the **alternative hypothesis**, H_1;
3. Calculate a **test statistic**;
4. Find the ***p*–value**, and
5. Use table 2.1 to form your **conclusions**.

However, the calculations required for the test statistic in step 3 are slightly different.

3.2 Testing two means

Recall that, in the test for one mean, there were two cases: population variance (σ^2) known and population variance unknown. Similarly, when comparing two means, we can consider a test where both population variances are known and both are unknown.

3.2.1 Both population variances (σ_1^2 and σ_2^2) known

1. **State the null hypothesis**
   This time, the null hypothesis is
   $$H_0 \;:\; \mu_1 = \mu_2,$$
   i.e. the two population means are equal.
2. **State the alternative hypothesis**
   We usually test against the (two–tailed) alternative:
   $$H_1 \;:\; \mu_1 \neq \mu_2,$$
   i.e. the population means *are not* equal. However, if we have reason to believe that one population mean is larger (or smaller!) than the other, we might want to use the (one–tailed) alternatives:
   $$\begin{aligned} H_1 &: \mu_1 > \mu_2, \qquad \text{or} \\ H_1 &: \mu_1 < \mu_2. \end{aligned}$$
3. **Calculate the test statistic**
   The test statistic for a two–sample test (when both population variances are known) is
   $$z = \frac{|\bar{x}_1 - \bar{x}_2|}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}},$$
   where $\bar{x}_1$, $\bar{x}_2$, n_1 and n_2 are the means and sample sizes of samples 1 and 2 (respectively), and σ_1^2 and σ_2^2 are the corresponding population variances.
4. **Find the p–value of the test**
   As in the tests for one mean, we use statistical tables to obtain our p–value, or rather a *range* for our p–value. Since, in this case, both population variances are known, we refer to standard normal tables (i.e. table 2.2) as before, the critical values depending on whether we are testing against a one– or two–tailed alternative hypothesis.
5. **Reach a conclusion**
   Use table 2.1 to form a conclusion (i.e. retain or reject the null hypothesis), remembering to also word your conclusions in plain English and in the context of the question posed.

Example
Before a training session for call centre employees a sample of 50 calls to the call centre had an average duration of 5 minutes, whereas after the training session a sample of 45 calls had an average duration of 4.5 minutes. The population variance is known to have been 1.5 minutes before the course and 2 minutes afterwards. Has the course been effective?

3.2.2 σ_1^2 and σ_2^2 unknown

In the more likely situation where the population variances are unknown, we use the test statistic

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}},$$

where s is a "pooled standard deviation", and is found as

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

Also, as with the test for one mean when the population variance was unknown, we need to use t–tables (table 2.3) to obtain our p–value for the test. The degrees of freedom is now found as $\nu = n_1 + n_2 - 2$. Apart from these changes, the hypothesis test follows the same format as that for which both population variances are known.

Example
A company is interested in knowing if two branches have the same level of average transactions. The company sample a small number of transactions and calculates the following statistics:

Shop 1	$\bar{x}_1 = 130$	$s_1^2 = 700$	$n_1 = 12$
Shop 2	$\bar{x}_2 = 120$	$s_2^2 = 800$	$n_2 = 15$

Test whether or not the two branches have (on average) the same level of transactions.

Steps 1 and 2 *(hypotheses)*
Our null and alternative hypotheses are:

$$\begin{aligned} H_0 &: \mu_1 = \mu_2 \qquad \text{versus} \\ H_1 &: \mu_1 \neq \mu_2. \end{aligned}$$

Step 3 *(calculating the test statistic)*
Since both population variances are unknown (only the *sample* values are given), the test statistic is

$$t = \frac{|\bar{x}_1 - \bar{x}_2|}{s \times \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}};$$

thus, we first need to obtain the pooled variance s. This is given as

$$
\begin{aligned}
s &= \sqrt{\frac{(n_1-1)s_1^2+(n_2-1)s_2^2}{n_1+n_2-2}} \\
&= \sqrt{\frac{(12-1)\times 700+(15-1)\times 800}{12+15-2}} \\
&= \sqrt{\frac{11\times 700+14\times 800}{25}} \\
&= \sqrt{\frac{7700+11200}{25}} \\
&= \sqrt{\frac{18900}{25}} \\
&= \sqrt{756} \\
&= 27.495.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
t &= \frac{|130-120|}{27.495\times\sqrt{\frac{1}{12}+\frac{1}{15}}} \\
&= \frac{10}{27.495\times\sqrt{0.15}} \\
&= \frac{10}{10.649} \\
&= 0.939.
\end{aligned}
$$

Step 4 *(finding the p–value)*
Since both population variances are unknown, we use t–tables to obtain our critical value. The degrees of freedom, $\nu = n_1+n_2-2$, i.e. $\nu = 12+15-2 = 25$. Under a two–tailed test, and using table 2.3, we get the following critical values:

Significance level	10%	5%	1%
Critical value	1.708	2.060	2.787

Our test statistic $t = 0.939$ lies to the left of the first critical value, and so our p–value is bigger than 10%.

Step 5 *(conclusion)*
Using table 2.1, we see that, since our p–value is larger than 10%, we have no evidence to reject the null hypothesis. Thus, we retain H_0 and conclude that there is no significant difference between the average level of transactions at the two shops.

3.3 Exercises

1. A chain of record shops believes that its Northumberland Street store (Shop 1) is more successful than its Metro Centre branch (Shop 2). The management take a random sample of daily takings and obtains the following summary statistics:

Shop 1	$\bar{x}_1 = £15000$	$s_1^2 = 400$	$n_1 = 10$
Shop 2	$\bar{x}_2 = £14250$	$s_2^2 = 600$	$n_2 = 20$

   Is the management's belief correct?

2. (a) An airline company, EasyAir, advertises 5 hour flights from Glasgow to Cairo. Another company flying the same route, RyanJet, suspects EasyAir of false advertising, and samples 20 of EasyAir's flights between Glasgow and Cairo. The mean flight time from this sample is 5 hours and 10 minutes with a standard deviation of 20 minutes. Are RyanJet's suspicions supported?

   (b) RyanJet claim that their Glasgow to Cairo flights are, on average, faster than EasyAir's flights. A sample of 23 RyanJet flights between Glasgow and Cairo has a mean flight time of 5 hours and 4 minutes with a standard deviation of 22 minutes. Are RyanJet's flights on this route shorter than EasyAir's?

3. The mean age of students attending a college is 19.5 years. The ages of a sample of 7 students who have been offered college accommodation are given below. Test the hypothesis that younger students are favoured.

   17.9 18.2 19.1 20.3 17.8 17.4 17.8

4. During the streaking phase of the 1970's, a psychological test designed to determine extroversion was applied to a group of 19 admitted male streakers and a control group of 19 male non–streakers. The results were

streakers	non–streakers
$\bar{x}_1 = 15.26$	$\bar{x}_2 = 13.90$
$s_1^2 = 2.62$	$s_2^2 = 4.11$

   Are streakers more extrovert than non–streakers?

5. In the comparison of two kinds of paint, a consumer testing service finds that eight one–gallon cans of "Wilko's Best", and ten one–gallon cans of "Dulor", cover the following areas (in square feet):

"Wilko's Best"	542	546	550	548	540	537	548	553		
"Dulor"	521	532	498	512	551	540	500	513	520	523

   The population standard deviation coverage for "Wilko's Best" and "Dulor" is 31 square feet and 26 square feet respectively. Perform a test to find out whether or not there is any difference in mean coverage between the two brands.

One–tailed test	10%	5%	2.5%	1%	0.5%
Two–tailed test	20%	10%	5%	2%	1%
Critical value	1.282	1.645	1.96	2.326	2.576

Table 3.1: Tabulated values of z for which $\Pr(Z > z) = p$, where Z has a standard normal distribution

	One–tailed test	10%	5%	2.5%	1%	0.5%
	Two–tailed test	20%	10%	5%	2%	1%
	1	3.078	6.314	12.706	31.821	63.657
	2	1.886	2.920	4.303	6.965	9.925
	3	1.638	2.353	3.182	4.541	5.841
	4	1.533	2.132	2.776	3.747	4.604
	5	1.476	2.015	2.571	3.365	4.032
	6	1.440	1.943	2.447	3.143	3.707
	7	1.415	1.895	2.365	2.998	3.449
ν	8	1.397	1.860	2.306	2.896	3.355
	9	1.383	1.833	2.262	2.821	3.250
	10	1.372	1.812	2.228	2.764	3.169
	11	1.363	1.796	2.201	2.718	3.106
	12	1.356	1.782	2.179	2.681	3.055
	13	1.350	1.771	2.160	2.650	3.012
	14	1.345	1.761	2.145	2.624	2.977
	15	1.341	1.753	2.131	2.602	2.947
	16	1.337	1.746	2.120	2.583	2.921
	17	1.333	1.740	2.110	2.567	2.898
	18	1.330	1.734	2.101	2.552	2.878
	19	1.328	1.729	2.093	2.539	2.861
	20	1.325	1.725	2.086	2.528	2.845
	21	1.323	1.721	2.080	2.518	2.831
	22	1.321	1.717	2.074	2.508	2.819
	23	1.319	1.714	2.069	2.500	2.807
	24	1.318	1.711	2.064	2.492	2.797
	25	1.316	1.708	2.060	2.485	2.787
	26	1.315	1.706	2.056	2.479	2.779
	27	1.314	1.703	2.052	2.473	2.771
	28	1.313	1.701	2.048	2.467	2.763
	29	1.311	1.699	2.045	2.462	2.756
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
	∞	1.282	1.645	1.960	2.326	2.576

Table 3.2: Tabulated values of t for which $\Pr(|T| > t) = p$, where T has a t–distribution with ν degrees of freedom