MAST7866 — Foundations of Data Science

Computing Session 9 – Hypothesis Testing 2

Within this worksheet we are going to be learning about how to perform hypothesis tests to compare two different populations.

Task 1 - Introduction

We often want to make comparisons between different groups of individuals – for example do boys or girls perform better at school exams?

Within this worksheet we will see how we can perform tests to determine if one population mean or proportion is significantly different from another population. Firstly, we need to determine if our two samples are independent or dependent.

Two samples are **independent** if they are drawn from two different populations and the elements of one sample have no relationship to the elements of the second sample. If the elements of the two samples are somehow related, then the samples are said to be **dependent**.

Salaries of male and female executives – Suppose we want to estimate the difference between the mean salaries if all male and all female executives. To do so, we draw samples, one from the population of male executives and another from the population of female executives. These two samples are **independent** because they are drawn from different populations, and the samples have no effect on one another.

Weights before and after a weight loss program – Suppose we want to estimate the difference between the mean weights of all participants before and after a weight loss program. To accomplish this, suppose we take a sample of 40 participants and measure their weights before and after the completion of this program. Note that these two samples include the same 40 participants. This is an example of two **dependent** samples.

Task 2 – Testing for differences between means (independent samples)

Suppose we have two independent samples from two different populations that are referred to as population 1 and population 2. Let μ_1 denote the mean of population 1 and μ_2 the mean of population 2. We will be testing whether:

H_0: $\mu_1 = \mu_2$

H_1: $\mu_1 \neq \mu_2$

Note that we can instead have one-sided tests which will have alternative hypotheses $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$. It will generally be the case that the population variance for each of the two populations is unknown. We can adapt the test depending on whether we believe the unknown population variance is thought to be the same or not within each of the two populations.

Example 1

A study of the effect of caffeine on muscle metabolism used 18 male volunteers who each underwent arm exercise tests. Nine of the men were randomly selected to take a capsule containing pure caffeine one hour before the tests whilst the other men received a placebo capsule. During each exercise the subject's respiratory exchange ratio (RER) was measured. The results were as follows:

Placebo	Caffeine
105	96
119	99
100	94
97	89
96	96
101	93
94	88
95	105
98	88

Does caffeine affect RER?

Let us start by inputting these data into R:

```
> placebo <- c(105, 119, 100, 97, 96, 101, 94, 95, 98)
> caffeine <- c(96, 99, 94, 89, 96, 93, 88, 105, 88)
```

Now let us perform a two sample t-test:

```
> t.test(placebo,caffeine)
```

You will see the following output:

```
     Welch Two Sample t-test

data:  placebo and caffeine
t = 1.9948, df = 14.624, p-value = 0.06505
alternative hypothesis: true difference in means is not equal
to 0
95 percent confidence interval:
 -0.4490961 13.1157627
sample estimates:
mean of x mean of y
100.55556  94.22222
```

We get a p-value of 0.06505 which implies that we have insufficient evidence to reject the null hypothesis of the placebo and caffeine capsules having equal effect.

Note that as a default the `t.test` function assumes that the population variances are not equal. If they can be assumed to be equal you can use the command `var.equal=TRUE`. The default is that you are performing a two-sided test. You can use the commands `alternative="less"` or `alternative="greater"` to specify a one-tailed test.

Challenge 1

We are now going to return to the Motor Car Trends data in R (`mtcars`). We are interested in determining if there is a different in the miles per gallon between automatic cars and manual cars.

```
> t.test(mtcars$mpg~mtcars$am)
```

Note that we are using a different command inside the `t.test` function because variable am is binary and subdivides the two populations. What do you conclude?

Task 3 – Testing for differences between means (dependent samples)

When we have dependent samples we have paired data, and thus we need to perform a paired t-test.

Example 2

Let us return to the anorexia data we have previously encountered. We want to assess whether there has been a significant increase in the weight of the patients before and after the study.

```
> library(MASS)
> data(anorexia)
> attach(anorexia)
> t.test(Postwt,Prewt,paired=TRUE, alternative="greater")

     Paired t-test

data:  Postwt and Prewt
t = 2.9376, df = 71, p-value = 0.002229
alternative hypothesis: true difference in means is greater than
0
95 percent confidence interval:
 1.195825      Inf
sample estimates:
mean of the differences
               2.763889
```

Hence we can conclude that there is strong evidence to reject the null hypothesis of no change in weight and conclude that there has been a significant increase in weight after treatment.

Challenge 2

Ten athletes ran a 400m race at sea level and at a later meeting ran another 400m race at high altitude. Their times in seconds were as follows:

Athlete	1	2	3	4	5	6	7	8	9	10
Sea Level	48.3	47.9	50.2	51.7	46.5	44.9	45.2	47.7	48.4	49.1
High altitude	48.7	49.2	50.1	51.9	48.2	45.8	48.0	47.3	50.2	51.5

Test whether the athletes are performing equally well at sea level and high altitude.

Task 4 – Testing for differences between proportions

Example 3

Suppose we have two drugs which are used to treat patients with a certain type of cancer. In order to compare their effectiveness, a clinical trial was planned. 75 patients were given drug A whilst 60 patients were given drug B. The number of patients who survived for one year beyond diagnosis in each group was as follows: Drug A: 49, Drug B: 34. How do we test whether both drugs are equally effective?

To do this we return to the `prop.test` function we used earlier for one sample problems.

```
> prop.test(x = c(49, 34), n = c(75, 60))


     2-sample test for equality of proportions with
     continuity correction

data:  c(49, 34) out of c(75, 60)
X-squared = 0.72294, df = 1, p-value = 0.3952
alternative hypothesis: two.sided
95 percent confidence interval:
 -0.0936275  0.2669608
sample estimates:
   prop 1    prop 2
0.6533333 0.5666667
```

The p-value is 0.3952 and from this we can conclude that there is insufficient evidence to reject the null hypothesis that the drugs are equally effective.

Challenge 3

A sample of 500 male registered voters showed that 57% of them voted in the last presidential election. Another sample of 400 female registered voters showed that 55% of them voted in the same election. Can you conclude that the proportion of male voters is different to the proportion of female voters?