

Chapter 2 - t -tests

1. Introduction

In this chapter, you will learn how to use the t -test to:

- compare the mean of a *single* population to a fixed value
- compare means of two *independent* populations with one another
- compare means of two *dependent* populations with one another

We will also touch upon the F -test for equal variance, and various tests of normality. Make sure you understand each of the following concepts:

- Population
- Sample
- Hypotheses:
- Null hypothesis
- Alternative hypothesis
- Two-sided alternative hypothesis
- Left one-sided alternative hypothesis
- Right one-sided alternative hypothesis
- Test statistic
- Significance level
- Critical value
- p -value
- Rejection of the null hypothesis
- Non-rejection of the null hypothesis
- Confidence interval

Note that we *always* make our conclusion in terms of the null hypothesis, i.e.

- Reject the null hypothesis, or
- Do not reject the null hypothesis.

Do not refer to “acceptance” of any hypothesis – a hypothesis can never be proven true; the only conclusion that can be made is that the hypothesis is either plausible or implausible based on the statistical evidence.

Background reading material

- Biometry 212: note set 16
- Biometry 242: note sets 1 and 3
- Clewer and Scarisbrick: chapters 5 and 6

2. One sample t -test (C&S 5.2, p. 39)

Import into *R* the `chap2data1.csv` file. The data represent 15-day comb masses of male chicks that each received a certain sex hormone. Use *R* to test the following hypotheses:

- That the true population mean comb mass is equal to 120g vs. the (two-sided) alternative that the true population mean is not equal to 120g.
- That the 15-day comb masses of chicks given sex hormone A are on average equal to 90g vs. the (one-sided right) alternative that they are greater than 90g.
- That the average 15-day comb mass for chicks on this hormone is equal to 100g vs. the (one-sided left) alternative that they are less than 100g.

One sample t -test:

```
t.test(x, y = NULL,
       alternative = c("two.sided", "less", "greater"),
       mu = 0, paired = FALSE, var.equal = FALSE,
       conf.level = 0.95)
```

The t -test is based upon two assumptions. Therefore, before you can believe the results of a t -test, you have to make sure that the assumptions are tenable for the particular data set. The one-sample t -test assumes:

- Normally-distributed data
- Independent observations

Proper *randomisation* in the sampling procedure should ensure independence of the observations.

To test the normality assumption we may use the Shapiro-Wilk test. In addition, histograms and quantile-quantile (Q-Q) plots can be drawn as rough indications of the level of normality, or the lack of it.

Shapiro-Wilk test

```
shapiro.test(x)
```

Where x is the single variable which you want to test.

QQ plot

```
qqPlot(chap2data1[, "Comb.mass"], main="QQ plot", ylab="Comb  
mass", xlab="Quantiles")
```

Investigate the assumption of normality for the *Comb mass* data.

Remember, tests and graphs become more accurate the more data we have, all else being equal.

If the data are not normal, the data can be transformed, or a *non-parametric* test can be performed instead of a *t*-test. This will be discussed in detail in a later chapter.

3. Two independent samples *t*-test (C&S 6.2, p. 51)

Import into *R* the file `chap2data2.csv`.

The data represent 15-day comb masses of eleven male chicks that each received sex hormone A, and 15-day comb masses of eleven other male chicks that each received sex hormone B. The chicks were randomly assigned to receive one of the two sex hormones.

When the true *means* of two populations are to be compared using a *t*-test, it is first necessary to determine whether we can assume equality of the population *variances* or not. Depending on whether this assumption of *homoscedasticity* can be made, we will use either a *pooled t-test* or a *t-test for unequal variances* to compare the true means. To test for homoscedasticity of two populations we make use of the *F*-test (C&S 6.5, p. 58).

Test the following hypotheses about the population mean comb mass using *R*:

- That the mean comb mass for hormone A chicks is the same as that for hormone B chicks.
- That the mean comb mass for hormone A chicks is 20g greater than that for hormone B chicks.
- That the mean comb mass for hormone B chicks is 12g less than that for hormone A chicks vs. the alternative that the mean mass difference is *greater* than 12g.

F-test for homoscedasticity

```
with(chap2data2, tapply(Comb.mass, Hormone, var, na.rm=TRUE))
var.test(Comb.mass~Hormone, alternative="two.sided", data=chap2data2)
```

t-test for two independent samples

```
t.test(x, y,
       alternative = c("two.sided", "less", "greater"),
       mu = 0, paired = FALSE, var.equal = FALSE,
       conf.level = 0.95)
```

The assumptions underlying this *t*-test are the same as those of the one-sample *t*-test: normally-distributed data and independent observations. In fact, the data from both samples *separately* must be normally distributed. These assumptions can be tested in the same way as for the one-sample case. Do so using the imported two-hormone data.

Checking normality assumption (two or more groups simultaneously):

```
shapiro.test(chap2data2$Comb_mass[chap2data2$Hormone == "A"])
shapiro.test(chap2data2$Comb_mass[chap2data2$Hormone == "B"])
```

Test the normality of the hormone treatment groups separately for the comb mass data. As before, if the data are not normally distributed, a *non-parametric* test must be performed instead of a *t*-test.

If the data consist of paired observations, consider a *t*-test for *dependent* samples, explained in the next section.

4. Two dependent samples *t*-test (C&S 5.8, p. 46)

Often, the data from two samples are not unrelated but *paired*. For example, ten people are about to embark on a diet: they are weighed *before* and again *after*. The *same* people are weighed twice.

This is done either by

- 1 telling the software to perform a *paired t-test*, or
- 2 manually calculating a column of differences in the paired values and then performing a one-sample *t*-test on this new column.

Import the *before* and *after* masses from the file `chap2data3.csv`. Note the format of the data. Now test the following hypotheses:

- That the diet has no effect (or is ineffective).

- That on average the diet leads to weight-loss of 10kg.
- That on average the diet leads to weight-loss of 12kg vs. the alternative that the average weight-loss is less than 12kg.

Paired samples *t*-test:

```
t.test(x, y,
       alternative = c("two.sided", "less", "greater"),
       mu = 0, paired = TRUE, var.equal = FALSE, conf.level = 0.95)
```

This test is based on the assumption that the *paired differences* are normally distributed and independent of one another. The assumption of normality can be tested in the same way as in the one sample case, but based on a column of differences in the paired values. Use *R* to check whether the paired differences of the weight loss data came from a normally distributed population.

Once again, if the assumption of normality is not valid, the differences can be transformed, or a *non-parametric* test can be performed instead of a *t*-test.

5. Power analysis

Consider the following definitions:

Type I error – rejecting the null hypothesis when it is true (false positive)

Type II error – not rejecting the null hypothesis when it is false (false negative)

Power – the probability to not commit a type II error.

A power analysis investigates and uses the relationship between power and sample size to determine the optimal sample size for your experiment. This relationship is investigated under various circumstances of natural variation, level of significance and effect sizes (smallest detectable differences).

2.5.1 Learning by example

Consider again the data in the file `chap2data3.csv`. Let's investigate the relationship between power and sample size for the `comb.mass` variable given that the estimate of the natural variation is the pooled variance, the level of significance is equal to 5% and that we want to detect a difference between the two hormones as small as 35g. Calculate the power for sample sizes of 5, 7, 9, 11, 13 and 15. Visualize the results with a scatter plot. Investigate this relationship for changes in the smallest detectable difference (δ) values. Use 20g, 35g and 50g in the analysis. What is the effect of changing the sample size and the smallest detectable difference on power?

```

#estimate the pooled variance
hA <- chap2data2[chap2data2$Hormone=="A",][,1]
hB <- chap2data2[chap2data2$Hormone=="B",][,1]
SSA <- sum(hA^2)-sum(hA)^2/length(hA)
SSB <- sum(hB^2)-sum(hB)^2/length(hB)
sp2 <- (SSA+SSB)/(length(hA)+length(hB)-2)
#OR
sp2 <- anova(lm(Comb.mass~Hormone,data=chap2data2))[2,3]
#Example, for n=11
power.t.test(power=NULL,n=11,sd=sp2^.5,delta=35)
#Extract only the value for power
power.t.test(power=NULL,n=11,sd=sp2^.5,delta=35)$power
#Tabulate all results
#set delta
delta <- 35
pwrTable <- data.frame(sample_size=c(5,7,9,11,13,15),

power=c(power.t.test(power=NULL,n=5,sd=sp2^.5,delta= delta)$power,

power.t.test(power=NULL,n=7,sd=sp2^.5,delta= delta)$power,

power.t.test(power=NULL,n=9,sd=sp2^.5,delta= delta)$power,

power.t.test(power=NULL,n=11,sd=sp2^.5,delta= delta)$power,

power.t.test(power=NULL,n=13,sd=sp2^.5,delta= delta)$power,

power.t.test(power=NULL,n=15,sd=sp2^.5,delta= delta)$power))
#plot results
plot(pwrTable)

```

6. Exercises

Exercise 1

With the data from C&S Example 5.1 (p. 41) (`chap2exer1.csv`), test if the mean linseed yield is equal to 2.0. Use a 5% significance level and clearly state your conclusion.

Exercise 2

Do C&S Exercise 5.1 (p. 45) (`chap2exer2.csv`). Test if the population mean is different from 8. Carry out the appropriate test, starting (and testing) any assumptions you make.

Exercise 3

With the data from C&S Table 5.2 (p. 48) (`chap2exer3.csv`)

Perform a paired t -test to see if the mean yield from the two wheat varieties differ. Use a 5% significance level. Calculate the sample differences between the two varieties (per farm), and perform a test to see if the mean difference is equal to zero.

Exercise 4

Use R to do C&S Exercise 6.2 (p. 61) (`chap2exer4.csv`). That is, test if Varieties A and B are significantly different. Data are the yields (in t/ha).