

## Chapter 18

### Tests of Hypotheses for one Population Mean

In this chapter, we will discuss the following topics:

- How to find quantiles of the  $t$ -distribution using the R-function **qt**.
- How to perform a significance test for testing the mean of a population with unknown standard deviation using the R-function **t.test**.
- How to perform a matched pairs  $t$  procedure using the R-function **t.test**.

We have two arguments in the function **qt(p,df)**, where  $p$  is the probability and  $df$  is the degrees of freedom.

- The function **qt(p,df)** returns the quantile for which there is a probability of  $p$  of getting a value less than or equal to it. Thus, the quantile is the value  $t$  such that  $P(T \leq t) = p$  for a given  $p$ , where  $T$  follows a  $t$ -distribution.
- The function **qt(p,df,lower.tail=FALSE)** returns the quantile for which there is a probability of  $p$  of getting a value greater than or equal to it. Thus, the quantile is the value  $t$  such that  $P(T \geq t) = p$  for a given  $p$ . The argument **lower.tail=TRUE** is the default in R.

The function **t.test** has several optional arguments.

- The function

```
t.test(x,mu)
```

performs a two-sided **t.test** on the sample  $x$  with the mean of the null hypothesis specified by  $mu$ . Notice that the default in R is  $mu=0$  and with alternative hypothesis  $\mu \neq \mu_0$ .

- We can also specify that the test is one sided by adding the argument **alternative=c("greater")** for  $\mu > \mu_0$  as in

```
t.test(x,alternative=c("greater"),mu).
```

- For testing against the alternative hypothesis  $\mu < \mu_0$ , we would add the argument **alternative=c("less")** as in

```
t.test(x,alternative=c("less"),mu).
```

- In addition, the desired confidence level can be specified by including the argument **conf.level=0.99** to get a 99% confidence interval as in

```
t.test(x,mu,conf.level=0.99).
```

- To perform a matched pairs  $t$  test on two simple random samples  $x$  and  $y$ , we would add the argument **paired=TRUE** as in

```
t.test(x,y,paired=TRUE).
```

### The $t$ -Distribution

Assume a simple random sample,  $X_1, X_2, \dots, X_n$ , of size  $n$  is drawn from a normal population with unknown mean  $\mu$  and unknown standard deviation  $\sigma$ . Since we do not know  $\sigma$ , we estimate it by the sample standard deviation,  $S$ . We then replace the standard deviation,  $\frac{\sigma}{\sqrt{n}}$ , of  $\bar{X}$  by its standard error  $\frac{S}{\sqrt{n}}$ . The resulting one sample statistics,

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}},$$

has the  $t$  distribution with  $n - 1$  degrees of freedom. The tails of the  $t$  distribution are heavier than those of the Normal distribution. In addition, we notice that if the sample size is large, the distribution of  $t$  is approximately standard normal.

**Problem.** Find the critical value for a one sided  $t$  test with 5 degrees of freedom and significance level  $\alpha = 0.05$ .

**Solution.** We want to find  $t^*$  such that  $P(T \leq t^*) = 0.95$ . We obtain:

```
> qt(0.95,5)
[1] 2.015048
```

Hence  $t^* = 2.015$ .

**Problem.** Suppose that we have a simple random sample of size 20 and we calculate the one-sample  $t$ -statistics. Find the critical value  $t^*$  such that

- $t$  has probability 0.01 to the right of  $t^*$ .
- $t$  has probability 0.90 to the left of  $t^*$ .

**Solution to part (a).** The degrees of freedom is 19. We want to find  $t^*$  such that  $P(T \geq t^*) = 0.01$ . We obtain:

```
> qt(0.01,19,lower.tail=FALSE)
[1] 2.539483
```

Hence,  $t^* = 2.539$ .

**Solution to part (b).** We want to find  $t^*$  such that  $P(T \leq t^*) = 0.9$ . We obtain:

```
> qt(0.9,19)
[1] 1.327728
```

Hence,  $t^* = 1.328$ .

### The One-Sample $t$ test

Assume a simple random sample,  $X_1, X_2, \dots, X_n$ , of size  $n$  is drawn from a normal population with unknown mean  $\mu$  and unknown standard deviation. However, the normal condition can be somewhat relaxed if the distribution is symmetric and with one peak unless we have a very small sample. A level  $C$  confidence interval for  $\mu$  is

$$\bar{x} \pm t^* \frac{s}{\sqrt{n}},$$

where the critical value  $t^*$  is defined such that the area under the  $t$ -distribution density curve is  $C$  between  $-t^*$  and  $t^*$ .

We would reject the null hypothesis if the value of  $t$  falls within the critical region; that is, if  $t > t^*$  for testing  $H_a : \mu > \mu_0$ , if  $t < -t^*$  for testing  $H_a : \mu < \mu_0$ , where  $t^*$  in both cases is defined such that  $P(T > t^*) = \alpha$ ; and finally if  $|t| > t^*$  for testing  $H_a : \mu \neq \mu_0$ , where in the latter case  $t^*$  is defined such that  $P(|T| > t^*) = \alpha$ . Recall from chapter 15, that the probability that  $t$  falls within the critical region when the null hypothesis is true is equal to the chosen significance level.

**Problem.** In 1879, Michelson [2] determined the speed of light. He had  $n = 100$  observations (he subtracted 299000 from the values and the units are km/sec). Stigler [3] reported the true speed of light to be 710.5. Michelson's data can be found in R by typing:

```
> morley
```

The mean in Michelson's data is  $\bar{x} = 852.4$  and the standard deviation is 79.01.

(A) Determine if Michelson's mean of the data is significantly different than 710.5 at the 5% significance level.

(B) Find a 99% confidence interval for the mean of the speed of light from Michelson's data.

**Solution to part (a).** We will include a stem plot to check that the data is approximately Normal. We wish to test:

$$H_0 : \mu = 710.5 \text{ against } H_a : \mu \neq 710.5.$$

Since we need to estimate the standard deviation from the data, we perform a two-sided  $t$  test in R:

```
table=morley
x=table$Speed
> x
 [1] 850 740 900 1070 930 850 950 980 980 880 1000 980 930 650 760
[16] 810 1000 1000 960 960 960 940 960 940 880 800 850 880 900 840
[31] 830 790 810 880 880 830 800 790 760 800 880 880 880 860 720
[46] 720 620 860 970 950 880 910 850 870 840 840 850 840 840 840
[61] 890 810 810 820 800 770 760 740 750 760 910 920 890 860 880
[76] 720 840 850 850 780 890 840 780 810 760 810 790 810 820 850
[91] 870 870 810 740 810 940 950 800 810 870
```

```
> stem(x)
```

The decimal point is 2 digit(s) to the right of the |

```
6 | 2
6 | 5
7 | 222444
7 | 566666788999
8 | 0000011111111112233444444444
8 | 55555555666677778888888888999
9 | 0011233444
9 | 55566667888
10 | 000
10 | 7
```

```
> t.test(x,mu=710.5)
```

One Sample t-test

```
data: x
t = 17.9596, df = 99, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 710.5
95 percent confidence interval:
 836.7226 868.0774
sample estimates:
mean of x
 852.4
```

The data is approximately Normal and no outliers do appear. Thus, we can use the  $t$  test. The  $p$ -value  $< 2.2 \times 10^{16}$  is smaller than 0.05 so we reject the null hypothesis. Thus, there is strong evidence that Michelson's determination of the speed of light is significantly different than Stigler's value of speed of light.

**Solution to part (b).** We obtain:

```
> t.test(x,mu=710.5,conf.level=0.99)
```

One Sample t-test

```
data: x
t = 17.9596, df = 99, p-value < 2.2e-16
alternative hypothesis: true mean is not equal to 710.5
99 percent confidence interval:
 831.6486 873.1514
sample estimates:
mean of x
 852.4
```

Hence, a 99% confidence interval for the mean is (831.6486, 873.1514). The true mean 710.5 is not contained in this interval so we again can see that the mean of Michelson's measurements of the speed of light is significantly different from 710.5 at the 1% significance level.

**Explanation.** The explanation of the code is as follows:

- The default in R is, **alternative=c("two-sides")**, for the alternative hypothesis.
- We set  $\mu_0=710.5$  to indicate that  $\mu_0 = 710.5$  in the null hypothesis.

- The line

```
data: x
```

indicates which data that are being tested.

- The line

```
t = 17.9596, df = 99, p-value < 2.2e-16
```

gives the value of the  $t$  statistics, the degrees of freedom, and the p-value.

- The line

```
alternative hypothesis: true mean is not equal to 710.5
```

indicates what the null hypothesis is and states the alternative hypothesis.

- The line

```
95 percent confidence interval:  
836.7226 868.0774
```

states the 95 percent confidence interval for the population mean  $\mu$ .

- The line

```
sample estimates:  
mean of x  
852.4
```

gives the sample mean,  $\bar{x}$ , of the data.

- Adding the argument **conf.level=0.99** to the function **t.test** returns a 99% confidence interval.

### Matched pairs $t$ test

Let  $X = (X_1, \dots, X_n)$  and  $Y = (Y_1, \dots, Y_n)$  be two random samples of size  $n$  measured on the same subjects such that  $X$  and  $Y$  are **dependent**.

For example  $X$  could be the old treatment given to a group of patients and  $Y$  could be the new treatment give to the same group of patients. We then want to measure the effect of  $Y$  compared to  $X$ . Let  $W = X - Y$  be the difference between  $X$  and  $Y$ . Let  $\mu_W$  and  $s_w$  denote the mean and standard deviation of  $W$ , respectively. The null hypothesis is then  $H_0 : \mu_W = 0$  and the alternative hypothesis is either one-sided, ( $\mu_W < 0$  or  $\mu_W > 0$ ), or two sided,  $\mu_W \neq 0$ .

The t-statistics is

$$t = \frac{\bar{X} - \bar{Y}}{\frac{S_{x-y}}{\sqrt{n}}} = \frac{\bar{W}}{\frac{S_w}{\sqrt{n}}}$$

which has a t-distribution with  $n - 1$  degrees of freedom.

**Problem.** In 1908, William Gosset [1] under the pseudonym **Student** published the article "The Probable error of a mean". He seeded plots with two different types of seed; regular and kiln-dried, and each type of seed was planted in 11 adjacent plots, so called split plots. His article is maybe the most famous one in statistical literature and is being read even today. Here are his data given in the 1908 paper:

```
X= 1903 1935 1910 2496 2108 1961 2060 1444 1612 1316 1511
Y= 2009 1915 2011 2463 2180 1925 2122 1482 1542 1443 1535
```

```
X=REG=Corn yield (lbs/acre) from regular seed.\\
Y=KILN = Corn yield from kiln-dried seed.\\
```

Is the mean corn yield from kiln-dried seed greater than the mean corn yield from regular seed at the 5% significance level?

**Solution.** Define  $W = X - Y$ . We wish to test whether the population mean of regular seed is smaller than the population mean of kiln-dried seed. This corresponds to testing whether the difference between their means is zero. According to how we defined  $W$ , the alternative hypothesis is that the difference in their means is smaller than zero. Thus, we are testing:

$$H_0 : \mu_W = 0 \text{ against } H_a : \mu_W < 0.$$

We use a matched pairs  $t$  test since we have 11 *pairs* of split plots. We include a stem plot:

```
> x=c(1903, 1935, 1910, 2496, 2108, 1961, 2060, 1444, 1612, 1316, 1511)
> y= c(2009, 1915, 2011, 2463, 2180, 1925, 2122, 1482, 1542, 1443, 1535)
> w=x-y
> stem(w)
```

The decimal point is 2 digit(s) to the right of the |

```
-1 | 310
-0 | 76
-0 | 42
```

```

0 | 234
0 | 7

> t.test(x,y,alternative=c("less"),mu=0,paired=TRUE)

      Paired t-test

data:  x and y
t = -1.6905, df = 10, p-value = 0.06091
alternative hypothesis: true difference in means is less than 0
95 percent confidence interval:
 -Inf 2.433766
sample estimates:
mean of the differences
      -33.72727

```

There are no signs of major departures from normality in the data of differences and no outliers do appear. The p-value is  $0.06091 > 0.05$  so we accept the null hypothesis at the 5% level. Therefore one should not treat the seeds by kiln drying them. However, this does not prove that the corn yields are equal, only that the data of differences does not deviate enough from zero for us to reject the null hypothesis. Maybe a larger sample size would have returned a different result.

**Explanation.** The code can be explained as follows:

- The argument **mu=0** and **alternative=c("less")** specifies that we are testing the null hypothesis that the mean is zero against the alternative hypothesis that the mean is smaller than zero.
- We added the argument **paired=TRUE** to perform a matched pairs *t* test. If we omit this entry, R will perform a regular two-sample t.test. (The two-sample t.test will be explained in the next chapter.)

## References

- [1] W.S. Gosset, "*The Probable Error of a Mean*". *Biometrika*, 6, pp 1-25, 1908. The data and the story are also listed on DASL at <http://lib.stat.cmu.edu/DASL/>
- [2] A. A. Michelson, *Experimental determination of the velocity of light made at the United States Naval Academy, Annapolis*. *Astronomic Papers* 1 1358. U.S. Nautical Almanac Office. (See Table 24.), 1882
- [3] S. M. Stigler, *Do robust estimators work with real data?* *Annals of Statistics* 5, 1055-1098. (See Table 6.), 1977. The data and the story are also listed on DASL at <http://lib.stat.cmu.edu/DASL/>

00