Experiment - 9

Applying the t-test for independent and dependent samples

Aim: To understand the Student's t-test for independent and dependent samples using R

Testing of Hypothesis – t-Test

Introduction

If the sample size is less than 30 i.e., n < 30, the sample may be regarded as small sample.

Student's t-test:

The student's t-test is mentioned as a method of testing the theory about the mean of a small sample drawn from a normally distributed population where the standard deviation of the given population is unknown.

The t distribution belonging under a family of curves in which the number of degrees of freedom specifies a particular curve. As the sample size (and the degrees of freedom) increases, the t distribution approaches the bell shape of the standard normal distribution. In common, for tests involving the mean of a sample of size greater than 30, then the normal distribution is applied.

$$t = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}}$$

Paired t- test:

The paired *t*-test is a method used to test whether the mean difference between pairs of measurements is zero or not. We can use the test when your data values are paired measurements. For example, you might have before-and-after measurements for a group of people. Also, the distribution of differences between the paired measurements should be normally distributed.

$$t = \frac{\mu_{\mathrm{d}}}{\frac{s}{\sqrt{n}}}$$

Problem: 1 (Student's t-test)

Two independent samples of sizes 8 and 7 contained the following values:

Sample 1	19	17	15	21	16	18	16	14
Sample 2	15	14	15	19	15	18	16	20

Is the difference between the sample means significant?

Code and Results:

```
# Problem 1

# input the data
sample1=c(19,17,15,21,16,18,16,14)
sample1

## [1] 19 17 15 21 16 18 16 14

sample2=c(15,14,15,19,15,18,16,20)
sample2

## [1] 15 14 15 19 15 18 16 20

# output using t-distribution
t=t.test(sample1,sample2)
t

##
##  Welch Two Sample t-test
##
## data:  sample1 and sample2
## t = 0.44721, df = 13.989, p-value = 0.6616
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -1.898128  2.898128
## sample estimates:
## mean of x mean of y
##      17.0      16.5

# test-statistic
cv=t$statistic
cv

##         t
## 0.4472136

#critical value
tv=qt(0.975,14)
tv

## [1] 2.144787

#conclusion
if(cv <= tv){print("Accept Ho")} else{print("Reject Ho")}

## [1] "Accept Ho"
```

Problem: 2 (Paired t-test)

The following data relate to the marks obtained by 10 students in two test, one held at the beginning of a year and the other at the end of the year after intensive coaching. Do the data indicate that the students have got benefited by coaching?

Test 1	19	17	15	21	16	18	16	14	19	20
Test 2	15	14	15	19	15	18	16	20	22	19

```
# Problem 2

#Paired- t-test
# input the data
test1=c(19,17,15,21,16,18,16,14,19,20)
test1

##  [1] 19 17 15 21 16 18 16 14 19 20

test2=c(15,14,15,19,15,18,16,20,22,19)
test2

##  [1] 15 14 15 19 15 18 16 20 22 19

t=t.test(sample1,sample2,paired=TRUE)
t

##
##  Paired t-test
##
## data:  sample1 and sample2
## t = 0.46771, df = 7, p-value = 0.6542
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -2.02789  3.02789
## sample estimates:
## mean of the differences
##                     0.5

# level of significance
alpha=0.05
# p-value
tv=t$p.value
tv

## [1] 0.6542055

# conclusion
if(tv >alpha){print("Accept Ho")} else{print("Reject Ho")}

## [1] "Accept Ho
```

F-test

When pairs of samples are taken from a normal population, the ratios of the variances of the samples in each pair will always follow the same distribution, the F-distribution.

The F-statistic is simply:

$$F = \frac{\sigma_1^2}{\sigma_2^2}$$

$$\sigma_1^2 = \frac{n_1 s_1^2}{n_1 - 1}, \quad d.f = n_1 - 1$$

$$\sigma_2^2 = \frac{n_2 s_2^2}{n_2 - 1}, \quad d.f = n_2 - 1$$

Problem: 3 (F-test)

Two independent samples of sizes 8 and 7 contained the following values:

Sample 1	19	17	15	21	16	18	16	14
Sample 2	15	14	15	19	15	18	16	20

Is the difference between the sample means significant?

Procedure:

- Import the data set
- Determine the critical value and sample statistic using R functions
- Conclude the problem using R functions

```
Codes and Results:

# Problem 3

# Variance test or F-test
sample1=c(19,17,15,21,16,18,16,14)
sample1

## [1] 19 17 15 21 16 18 16 14

sample2=c(15,14,15,19,15,18,16,20)
sample2

## [1] 15 14 15 19 15 18 16 20

# output using t-distribution
f=var.test(sample1,sample2)
f

##
##  F test to compare two variances
```

```
##
## data:  sample1 and sample2
## F = 1.0588, num df = 7, denom df = 7, p-value = 0.9418
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.2119805 5.2887274
## sample estimates:
## ratio of variances
##           1.058824
```

```
# test-statistic
cv=f$statistic
cv
```

```
##        F
## 1.058824
```

```
#critical value
tv=qf(0.95,7,7)
tv
```

```
## [1] 3.787044
```

```
#conclusion
if(cv <= tv){print("Accept Ho")} else{print("Reject Ho")}
```

```
## [1] "Accept Ho"
```

Conclusion: Student’s t-test and F-test have been explored and executed.