**STAT 40001/MA 59800 Statistical Computing Fall 2017**

**Lab-11**

1. The data set *normtemp* in *UsingR* package contains measurements of 130 healthy randomly selected individuals. The variable temperature contains normal body temperature. Does the data support that the average body temperature is 98.6 0 F. Perform the hypothesis test.

> library(UsingR)

> data("normtemp")  
> t.test(normtemp$temperature,mu=98.6,conf.level=0.9)

One Sample t-test

data: normtemp$temperature

t = -5.4548, df = 129, p-value = 2.411e-07

alternative hypothesis: true mean is not equal to 98.6

90 percent confidence interval:

98.14269 98.35577

sample estimates:

mean of x

98.24923

p-value<0.05, reject the H0 hypothesis

1. A certain brand of apple juice is supposed to have 64 ounces of juice. In order to check whether the whether the filling machine was calibrated correctly, a sample data was collected.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 64.05 | 64.05 | 64.03 | 63.97 | 63.95 | 64.02 |
| 64.01 | 63.99 | 64.00 | 64.01 | 64.06 | 63.94 |
| 63.98 | 64.05 | 63.95 | 64.01 | 64.08 | 64.01 |
| 63.95 | 63.97 | 64.10 | 63.98 |  |  |

Based on these data check whether the machine is calibrated correctly.

> x = scan()

1: 64.05 64.05 64.03 63.97 63.95 64.02

7: 64.01 63.99 64.00 64.01 64.06 63.94

13: 63.98 64.05 63.95 64.01 64.08 64.01

19: 63.95 63.97 64.10 63.98

23:

Read 22 items

> t.test(x,mu=64)

One Sample t-test

data: x

t = 0.76426, df = 21, p-value = 0.4532

alternative hypothesis: true mean is not equal to 64

95 percent confidence interval:

63.98748 64.02706

sample estimates:

mean of x

64.00727

1. The waiting time (mins) of 100 bank customers before service is being rendered are provided below

0.8, 0.8, 1.3, 1.5, 1.8, 1.9, 1.9, 2.1, 2.6, 2.7, 2.9, 3.1, 3.2, 3.3, 3.5, 3.6, 4.0, 4.1, 4.2, 4.2, 4.3, 4.3, 4.4, 4.4, 4.6, 4.7, 4.7, 4.8, 4.9, 4.9, 5, 5.3, 5.5, 5.7, 5.7, 6.1, 6.2, 6.2, 6.2, 6.3, 6.7, 6.9, 7.1, 7.1, 7.1, 7.1, 7.4, 7.6, 7.7, 8, 8.2, 8.6, 8.6, 8.6, 8.8, 8.8, 8.9, 8.9, 9.5, 9.6, 9.7, 9.8, 10.7, 10.9, 11, 11, 11.1, 11.2, 11.2, 11.5, 11.9, 12.4, 12.5, 12.9, 13, 13.1, 13.3, 13.6, 13.7, 13.9, 14.1, 15.4, 15.4, 17.3, 17.3, 18.1, 18.2, 18.4, 18.9, 19, 19.9, 20.6, 21.3, 21.4, 21.9, 23.0, 27, 31.6, 33.1, 38.5

1. Construct a 95% confidence interval for waiting time for the bank customers.

> t.test(data)$conf.int

[1] 8.441023 11.312977

attr(,"conf.level")

[1] 0.95

1. Construct a 99% confidence interval for waiting time for the bank customers.

> t.test(data,conf.level =0.99)$conf.int

[1] 7.976271 11.777729

attr(,"conf.level")

[1] 0.99

1. Do you have enough evidence to conclude that it takes on average more than 8 minutes before you are served?

> t.test(data,mu=8,alt="greater")

One Sample t-test

data: data

t = 2.5936, df = 99, p-value = 0.005468

alternative hypothesis: true mean is greater than 8

95 percent confidence interval:

8.675376 Inf

sample estimates:

mean of x

9.877

Reject the H0 hypothesis

1. In order to investigate the possible relationship between marijuana smoking and a deficit in performance on a task measuring short term memory--the digit span task from the Wechsler Adult Intelligence Scale two groups of ten subjects were tested. One group, the "nonsmokers," claimed not to smoke marijuana. A second group, the "smokers," claimed to smoke marijuana regularly. Below are the scores.

nonsmokers : 18,22,21,17,20,17,23,20,22,21

smokers : 16,20,14,21,20,18,13,15,17,21

Do we have enough evidence that the nonsmoker has higher score than smokers?

H0 : nonsmoker <= smokers  
Ha : nonsmoker > smokers  
> t.test(NS,S,alt="greater",var.equal = TRUE)

Two Sample t-test

data: NS and S

t = 2.2573, df = 18, p-value = 0.01833

alternative hypothesis: true difference in means is greater than 0

95 percent confidence interval:

0.6026879 Inf

sample estimates:

mean of x mean of y

20.1 17.5  
p-value < 0.05, Reject H0

1. Two identical footballs, one air-filled and one helium-filled, were used outdoors on a windless day at The Ohio State University's athletic complex. Each football was kicked 39 times and the two footballs were alternated with each kick. The experimenter recorded the distance traveled by each ball. The data are available in the link below

<http://lib.stat.cmu.edu/DASL/Datafiles/Heliumfootball.html>

Perform the hypothesis to test that there is a significant difference in the mean distance travel based on whether the ball is filled with air or with helium.

data = read.table("C:\\Users\\Administrator\\Desktop\\statistical computing\\lab11\\lab11#5.txt",header = T)  
> t.test(data$Air,data$Helium,paired = T)

Paired t-test

data: data$Air and data$Helium

t = -0.41976, df = 38, p-value = 0.677

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-2.687423 1.764346

sample estimates:

mean of the differences

-0.4615385  
p-value > 0.05, there are no enough evidence to demonstrate that the differences between the distances of the two balls is not equal to 0.

1. Furness and Bryant (1996) compared the metabolic rates of male and female breeding northern fulmars (data described in Logan (2010) and Quinn (2002)).

|  |  |
| --- | --- |
| Sex | Metabolic rate |
| Female | 728 |
| Female | 1087 |
| Female | 1091 |
| Female | 1361 |
| Female | 1491 |
| Female | 1956 |
| Male | 526 |
| Male | 606 |
| Male | 843 |
| Male | 1196 |
| Male | 1946 |
| Male | 2136 |
| Male | 2309 |
| Male | 2950 |

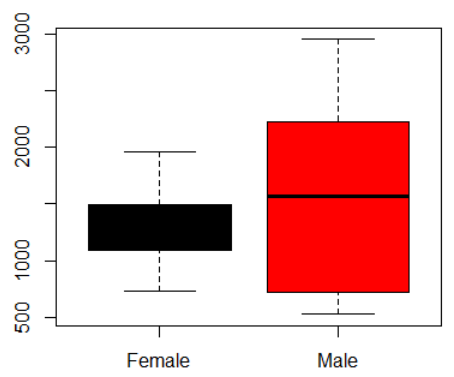
1. Display the metabolic rate of Female and Male group using side-by-side boxplot

> Sex = c("Female","Female","Female","Female","Female","Female","Male","Male","Male","Male","Male","Male","Male","Male")

> MetabolicRate = c(728,1087,1091,1361,1491,1956,526,606,843,1196,1946,2136,2309,2951)

> data = data.frame(Sex,MetabolicRate)

> boxplot(MetabolicRate~factor(Sex),col=c(1,2))



1. Test the hypothesis whether there is a difference in Metabolic rate based on gender

>t.test(data[data[,'Sex']=='Male',"MetabolicRate"],data[data[,'Sex']=='Female',"MetabolicRate"],var.equal = T)

Two Sample t-test

data: data[data[, "Sex"] == "Male", "MetabolicRate"] and data[data[, "Sex"] == "Female", "MetabolicRate"]

t = 0.70122, df = 12, p-value = 0.4965

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-586.7591 1143.6757

sample estimates:

mean of x mean of y

1564.125 1285.667

p-value > 0.05, there is no evidence to demonstrate that the difference in means is not equal to 0.