Lab 11

Alexander Hernandez

09/29/2022

```
library(UsingR)
```

```
## Loading required package: MASS
## Loading required package: HistData
## Loading required package: Hmisc
## Loading required package: lattice
## Loading required package: survival
## Loading required package: Formula
## Loading required package: ggplot2
##
## Attaching package: 'Hmisc'
## The following objects are masked from 'package:base':
##
##     format.pval, units
##
## Attaching package: 'UsingR'
## The following object is masked from 'package:survival':
##
##     cancer
```

1) Does the data in normtemp support that average body temp is 98.6? Hypothesis test

```
# Null:         u  = 98.6
# Alternative:  u != 98.6
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
```

```
# The p-value is 2.411e^-07,
# we have enough evidence to reject the null hypothesis.
```

2) Waiting time of 100 bank customers

```
waiting_data = c(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)
```

a) construct a 95% confidence interval

```
t.test(waiting_data, conf.level=0.95)$conf.int
```

```
## [1]  8.441023 11.312977
## attr(,"conf.level")
## [1] 0.95
```

b) construct a 99% confidence interval

```
t.test(waiting_data, conf.level=0.99)$conf.int
```

```
## [1]  7.976271 11.777729
## attr(,"conf.level")
## [1] 0.99
```

c) Is there enough evidence to conclude that it takes on average more than 8 minutes to be served

```
# Null:         u  = 8
# Alternative:  u >= 8
t.test(waiting_data, mu=8,
       alternative="greater")
```

```
##
##  One Sample t-test
##
## data:  waiting_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
```

```
# With a p-value of 0.0055,
# there is enough evidence to reject the null hypothesis.
```

3) Smokers and Non-Smokers. Is there evidence that the nonsmokers has a higher score than smokers?

```
# Null:         u(ns) - u(s) = 0
# Alternative:  u(ns) - u(s) > 0
smokers = c(16,20,14,21,20,18,13,15,17,21 )
nonsmokers = c(18,22,21,17,20,17,23,20,22,21)

t.test(nonsmokers, smokers, alt="greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  nonsmokers and smokers
## t = 2.2573, df = 16.376, p-value = 0.01899
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  0.5919017       Inf
## sample estimates:
## mean of x mean of y
##      20.1      17.5
```

```
# With a p-value of 0.019,
# we have enough evidence to reject the null hypothesis.
```

4) Pulse rate Experiment

```
pulse_data = read.table("http://www.statsci.org/data/oz/ms212.txt", header=TRUE)
```

a) Test the hypothesis of whether there is a difference in pulse rate if the students were sitting

Null: u(norm) - u(sitting) = 0

```
# Alternative:  u(norm) - u(sitting) != 0

# First version:
t.test(pulse_data$Pulse2~pulse_data$Ran, alt="greater")
```

```
##
```

```
##  Welch Two Sample t-test
##
## data:  pulse_data$Pulse2 by pulse_data$Ran
## t = 13.156, df = 57.309, p-value < 2.2e-16
## alternative hypothesis: true difference in means between group 1 and group 2 is greater than 0
## 95 percent confidence interval:
##  45.38385      Inf
## sample estimates:
## mean in group 1 mean in group 2
##       126.84783        74.85714
```

```
# Second version:
R = subset(pulse_data, pulse_data$Ran=="1")
S = subset(pulse_data, pulse_data$Ran=="2")
t.test(R$Pulse2, S$Pulse2, alt="greater")
```

```
##
##  Welch Two Sample t-test
##
## data:  R$Pulse2 and S$Pulse2
## t = 13.156, df = 57.309, p-value < 2.2e-16
## alternative hypothesis: true difference in means is greater than 0
## 95 percent confidence interval:
##  45.38385      Inf
## sample estimates:
## mean of x mean of y
## 126.84783  74.85714
```

```
# With a p-value of 2.2e^-16, we have enough evidence to reject the null hypothesis.
```

b) Test the hypothesis whether the average pulse rate for running students increased by 10 after they ran

```
# Null:         u(norm) - u(ran) = 10
# Alternative:  u(norm) - u(ran) > 10
length(R$Age)
```

```
## [1] 46
```

```
t.test(R$Pulse2, R$Pulse1,
       alt="greater",
       mu=10,
       paired=T)
```

```
##
##  Paired t-test
##
## data:  R$Pulse2 and R$Pulse1
## t = 13.311, df = 45, p-value < 2.2e-16
## alternative hypothesis: true mean difference is greater than 10
## 95 percent confidence interval:
##  46.16911      Inf
## sample estimates:
## mean difference
##         51.3913
```

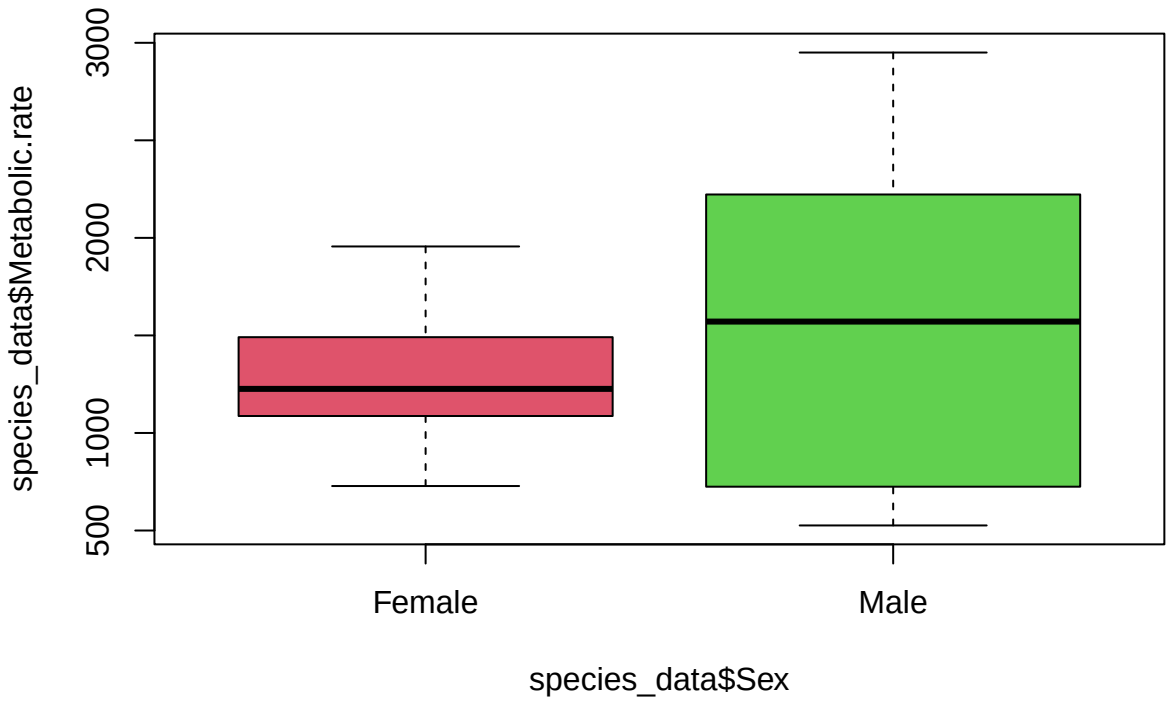
```
# With a p-value of 2.2e^-16,
# we have evidence to reject the null hypothesis.
```

5) Furness and Bryant Metabolic rates of male and female breeding fulmars

a) Display the metabolic rate of female and male group using side-by-side boxplot

```
species_data = read.csv('C:\\repos\\STAT 50001\\Lab 11\\species.csv', sep='\t')

boxplot(species_data$Metabolic.rate ~ species_data$Sex, col=c(2,3))
```



b) Test the hypothesis whether there is a difference in metabolic rate based on gender

```
# Null:         u(f) - u(m)  = 0
# Alternative:  u(f) - u(m) != 0
t.test(species_data$Metabolic.rate ~ species_data$Sex)
```

```
##
##  Welch Two Sample t-test
##
## data:  species_data$Metabolic.rate by species_data$Sex
## t = -0.77341, df = 10.466, p-value = 0.4564
```

```
## alternative hypothesis: true difference in means between group Female and group Male is not equal to
## 95 percent confidence interval:
##  -1075.375   518.708
## sample estimates:
## mean in group Female   mean in group Male
##             1285.667             1564.000
```

```
# with a p-value of 0.456,
# we do not have enough evidence to reject the null.
```