Assignment 3 - Daniel Yim

Daniel Yim

2023-03-22

#Assignment 3 - 20%

#Question 1 #Consider the gain in weight of 19 female rats between 28 and 84 days after birth. #12 were fed on high protein diet and 7 on a low protein diet. #Using the following data, test the hypothesis that there is no difference #in weight gain between female rats raised on a high protein diet versus those #raised on a low-protein diet. Use a significance level of mieu = 0.05 and assume equal variances. "Hint: var.equal="TRUE" #High protein: 134,146,104,119,124,161,107,83,113,129,97,12 #Low protein: 70,118,101,85,107,132,94 #Unpaired t-test because the two rats were independently tested

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.4.0      v purrr   0.3.5
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
high_pro <-c(134,146,104,119,124,161,107,83,113,129,97,12)
low_pro <-c(70,118,101,85,107,132,94)
t.test(high_pro,low_pro,var.equal = TRUE)
```

```
##
##  Two Sample t-test
##
## data:  high_pro and low_pro
## t = 0.62634, df = 17, p-value = 0.5394
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -23.09271  42.59271
## sample estimates:
## mean of x mean of y
##    110.75    101.00
```

#Decision: Since the p-value is greater than 0.05, we reject the null hypothesis that there is no difference in weight between the two rat groups.

#Question 2 #Load the "MASS package. In the immer dataset of the"MASS" library: #we have a Y1 Yield in 1931, Y2 yield in 1932. #Assuming that the data in immer follows the normal distribution,

```
library(dbplyr)
```

```
##
## Attaching package: 'dbplyr'

## The following objects are masked from 'package:dplyr':
##
##     ident, sql
```

```
library(MASS)
```

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select
```

```
View(immer)
```

#Find the 95% confidence interval estimate of the difference between the mean barley yields between years 1932 and 1932 (Hint: paired t-test).

#Get "p-value" in a variable pvalue and "statistics" in a variable st. (Hint: ttest<-t.test(..,..,...) and then names(ttest))

```
ttest <-t.test(immer$Y1,immer$Y2,paired=TRUE)
ttest
```

```
##
##  Paired t-test
##
## data:  immer$Y1 and immer$Y2
## t = 3.324, df = 29, p-value = 0.002413
## alternative hypothesis: true mean difference is not equal to 0
## 95 percent confidence interval:
##   6.121954 25.704713
## sample estimates:
## mean difference
##        15.91333
```

```
names(ttest)
```

```
##  [1] "statistic"   "parameter"   "p.value"     "conf.int"    "estimate"
##  [6] "null.value"  "stderr"      "alternative" "method"      "data.name"
```

#Question 3 #A professor takes a random sample of students enrolled in her course. #She finds the following: in the sample, there are 25 freshmen, 32 sophomores #and 20 seniors. Test the null hypothesis that freshmen, sophomores, juniors, and seniors #are equally represented among students signed up for this course. #Hint: Chi-square

```r
qchisq(0.05, df=3, lower.tail=FALSE)
```

```
## [1] 7.814728
```

```r
pchisq(4.9158, df=3, lower.tail=FALSE)
```

```
## [1] 0.1780675
```

```r
obs = c(25,32,18,20)
exp = c(1/4,1/4,1/4,1/4)
chisq.test(x=obs,p=exp)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  obs
## X-squared = 4.9158, df = 3, p-value = 0.1781
```

```r
# p-value .1781 > 5%
# test statistic 4.9158 < critical value 7.814728
# do not reject null hypothesis because test statistic is not greater than critical value
```