Homework 12

Peyton Hall

04/11/2024

```r
library(stats)
library(readxl)
```

1. A clinic developed a new scheduling method to reduce the amount of unscheduled nurse visits. The proportion of unscheduled nurse visits was 40% before this intervention. After this intervention, in a sample of 100 clinic visits, 25 of them were unscheduled nurse visits. Use the 0.05 significance level to test the claim that the proportion of unscheduled nurse visits is significantly less than 40%. Show your R code. Write down the H0 and Ha and the decision to address the original claim in the context in R Markdown using comments. $H_0 : P = .40$ vs $H_A : P < .40$

```r
#(x,n,p)
prop.test(25,100,p=.4,alternative = "less")
```

```
##
##  1-sample proportions test with continuity correction
##
## data:  25 out of 100, null probability 0.4
## X-squared = 8.7604, df = 1, p-value = 0.001539
## alternative hypothesis: true p is less than 0.4
## 95 percent confidence interval:
##  0.00000 0.33249
## sample estimates:
##    p
## 0.25
```

The p value is 0.001539. The P-Value is low because, compared to the significance level, it is less than .05. Therefore, we reject H_0 and go with the H_A. There is sufficient evidence to support the claim that the proportion of unscheduled nurse visits is significantly less than 40%.

2. A pharmaceutical company developed a new drug (A) to treat migraine. The company would like to compare the proportion of people who responded positively to drug A is significantly higher than the proportion of people who responded positively to the standard drug (drug B). On one sample of 200 people who took A, 130 responded positively, and on another sample of 100 people who took B, 50 people responded positively. Use the 0.05 significance level to test the claim that the proportion of people who responded positively to A is significantly higher than the proportion who responded positively to B. Show your R code. Write down the H0 and Ha and the decision to address the original claim in the context in R Markdown using comments. $H_0 : P_A = P_B$ vs $H_A : P_A > P_B$ let p_A = Drug A let p_B = Drug B

```r
AB <- c(130, 50)
N_AB <- c(200, 100)
prop.test(AB, N_AB, alternative = "greater")
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  AB out of N_AB
## X-squared = 5.6406, df = 1, p-value = 0.008774
## alternative hypothesis: greater
## 95 percent confidence interval:
##  0.0432961 1.0000000
## sample estimates:
## prop 1 prop 2
##   0.65   0.50
```

The probability value is 0.008774. That is, lower than the 0.05 significance level. Therefore, we reject the H_0 and go with the H_A. There is sufficient evidence to support the claim that the proportion of people who responded positively to A is significantly higher than the proportion who responded positively to B

3. A car manufacturer aims to improve the quality of the products by reducing defects and hence increasing customer satisfaction. He monitors the efficiency of two assembly lines in the shop. In line A, there are 18 defects reported out of 200 samples. In line B, there are 25 defects out of 600 cars. Use the 0.05 significance level to test the claim that the proportion of defects in line A is significantly different from the proportion in line B. Write down the H0 and Ha, and the decision to address the original claim in the context in R Markdown using comments. $H_0 : P_A = P_B$ vs $H_A : P_A \neq P_B$

```r
car_AB <- c(18, 25)
car_N_AB <- c(200, 600)
prop.test(car_AB, car_N_AB, alternative = "two.sided") # syntax for not equal
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  car_AB out of car_N_AB
## X-squared = 5.9722, df = 1, p-value = 0.01453
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  0.002236347 0.094430320
## sample estimates:
##     prop 1     prop 2
## 0.09000000 0.04166667
```

The p-value is 0.01453. The significance level is 0.05. The p-value is less than the significance level. Therefore, we would reject the H_0 and go with the H_A. There is sufficient evidence to support the claim that the proportion of defects in line A is significantly different from the proportion in line B.

4. The data "COVID19 Vaccine Preference" on D2L showed a sample of 50 individuals' preference for the type of COVID19 vaccine. Use the table() function to get the counts of each preference first. Use the 0.05 significance level to test the claim that the proportion of preferring Moderna vaccine is significantly different from the proportion of preferring Pfizer. (use a comment to show your null and alternative hypothesis and the decision to the claim) $H_0 : P_M = P_Z$ vs $H_A : P_M \neq P_Z$ let P_M = proportion of Moderna preference let P_Z = proportion of Pfizer preference

```r
COVID19_Vaccine_Preference <- read_excel("~/Desktop/Data211/Week 13/COVID19 Vaccine Preference.xlsx")
COVID19_Vaccine_Preference
```

```
## # A tibble: 50 x 2
##       ID Preference
##    <dbl> <chr>
##  1     1 Moderna
##  2     2 Moderna
##  3     3 Pfizer
##  4     4 Moderna
##  5     5 Moderna
##  6     6 Pfizer
##  7     7 Moderna
##  8     8 Moderna
##  9     9 Moderna
## 10    10 Moderna
## # i 40 more rows
```

```r
table(COVID19_Vaccine_Preference$Preference)
```

```
##
## Moderna  Pfizer
##      26      24
```

```r
length(COVID19_Vaccine_Preference$Preference)
```

```
## [1] 50
```

```r
Preference <- c(26,24) # total of 50
V <- c(50,50)

prop.test(Preference,V,alternative = "two.sided") #(x,n,p)
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  Preference out of V
## X-squared = 0.04, df = 1, p-value = 0.8415
## alternative hypothesis: two.sided
## 95 percent confidence interval:
##  -0.1758395  0.2558395
## sample estimates:
## prop 1 prop 2
##   0.52   0.48
```

The p-value is 0.8415. The significance level is 0.05. The p-value is greater than the significance level. Therefore, we fail to reject the H_0. There is insufficient evidence to support the claim that the proportion of preferring Moderna vaccine is significantly different from the proportion of preferring Pfizer.

5. The data "onlineLearning" on D2L recorded the students' ratings of synchronous online classes and asynchronous online classes. Use the 0.05 significance level to test the claim that the proportion of a rating of 5 for synchronous classes is significantly higher than the proportion of a rating of 5 for asynchronous classes. Use table() to get the corresponding counts. $H_0 : P_S = P_A$ vs $H_A : P_S > P_A$

```r
onlineLearning <- read_excel("~/Desktop/Data211/Week 13/onlineLearning.xlsx")
onlineLearning
```

```
## # A tibble: 79 x 3
##       ID  Sync Async
##    <dbl> <dbl> <dbl>
##  1     1     5     5
##  2     2     5     4
##  3     3     5     5
##  4     4     5     2
##  5     5     4     4
##  6     6     5     2
##  7     7     4     2
##  8     8     5     5
##  9     9     4     4
## 10    10     5     5
## # i 69 more rows
```

```r
n<-as.numeric(length(onlineLearning$ID))
n<-c(n,n)
x<-c(as.numeric(table(onlineLearning$Async)[5]), as.numeric (table(onlineLearning$Sync)[3]))
prop.test (x, n, alternative = "less")
```

```
##
##  2-sample test for equality of proportions with continuity correction
##
## data:  x out of n
## X-squared = 4.6747, df = 1, p-value = 0.0153
## alternative hypothesis: less
## 95 percent confidence interval:
##  -1.00000000 -0.04153469
## sample estimates:
##    prop 1    prop 2
## 0.2658228 0.4430380
```

The p-value is 0.0153 The significance level is 0.05. The p-value is significantly lower than the significance level. Therefore, we would reject the H_0 and go with the H_A. There is sufficient evidence to support the claim that the proportion of a rating of 5 for synchronous classes is significantly higher than the proportion of a rating of 5 for asynchronous classes.