

Chapter 11 and 12 Homework

Due 4-27-2018

Data from 10.14

a) Would it be appropriate to use a normal approximation in conducting a statistical test of the research hypothesis that over half of persons suffering from chronic pain are over 50 years of age?

With $\pi(n) > 5$ and $(1-\pi)(n) > 5$, it is appropriate to use a normal approximation.

b) Using the data in the survey, is there substantial evidence ($\alpha = 0.05$) that more than half of persons suffering from chronic pain are over 50 years of age?

$$H_0 : \pi = 0.50$$

$$H_1 : \pi > 0.50$$

Test Statistic:

We have a $\chi^2 = 35.28$ and a p-value of 1.428×10^{-9}

Rejection Region: Reject if p-value is less than 0.05.

Conclusion/Interpretation: Since our p-value is less than 0.05, we have significant evidence to suggest reject the null hypothesis is equal to 0.50. There appears to be evidence that the probability of success is greater than 0.50.

c) Place a 95% confidence interval on the proportion of persons suffering from chronic pain that are over 50 years of age.

We are 95% confident the true percent of people who would be over 50 years of age experiencing chronic pain is between 0.5711 and 0.6389.

a) Place a 95% confidence interval on p1-p2, the difference in the proportions of customers purchasing lawnmowers with and without the warranty.

We are 95% confident the true difference in percents of customers who buy warranties split between customers offered a warranty and those not offered a warranty is between 0.0001 and 0.4799.

b) Test the research hypothesis that offering the warranty will increase the proportion of customers who will purchase a mower. Use $\alpha = 0.05$.

Hypotheses:

$$H_0 : \pi_{Offered} = \pi_{NotOffered}$$

$$H_1 : \pi_{Offered} > \pi_{NotOffered}$$

Test Statistic:

We have a χ^2 of 2.4802.

Rejection Region:

We reject the null hypothesis if our $\chi^2 > \chi_{0.95,1}^2 = 3.841$

Conclusion/Interpretation: Since we do not have a test statistic that is bigger than our rejection region, we will not reject the null hypothesis that the proportions are the same. We do not have enough evidence to say offering customers a warranty will lead to increases warranty purchases.

c) Are the conditions for using a large-sample test to answer the question in part b) satisfied? If not, apply an exact procedure.

Since $(\pi_{NotOffered})(25) < 5$, we have not met the requirements to use a large sample test. We will use Fisher's Exact to look at the results.

Hypotheses:

$$H_0 : \pi_{Offered} = \pi_{NotOffered}$$

$$H_1 : \pi_{Offered} > \pi_{NotOffered}$$

Test Statistic/P-value:

We have a p-value of 0.05683.

Rejection Region:

We will reject if p-value is less than $0.05 = \alpha$.

Conclusion/Interpretation:

With a p-value greater than α , we do not have significant evidence to reject the null hypothesis that the take rates on warranties are the same. There does not appear to be evidence that the take rate increases with an offer for a warranty.

d) Based on your results from parts a) and b), should the dealer offer the warranty?

The dealer should not offer the warranty. The test stat isn't statistically significant at $\alpha = 0.05$.

10.42 (b,d)

b) Is the promotion decision for the fireman related to the age of the fireman? Use $\alpha = 0.05$.

Hypotheses:

H_0 : Promotion as a fireman is independent of age.

H_1 : Promotion as a fireman depends on age.

Test Statistic:

We have a $\chi^2 = 12.796$ and a p-value = 0.0051.

Rejection Region:

We would reject the null hypothesis if our test stat is greater than $\chi^2_{3,0.95} = 7.815$

Conclusion/Interpretation:

There is evidence to reject the null hypothesis that age doesn't impact promotion rate. It appears the promotion of a fireman does depend on their age.

What are some other variables, besides age that needed to be addressed in an age discrimination analysis?

I believe sex/gender, years with the department, and number of previous promotions are adequate variables to split this data on.

10.43

a) Is the age of the fireman related to whether or not the fireman is promoted? Use $\alpha = 0.05$.

Hypotheses:

H_0 Promotion as a fireman is independent of age.

H_1 : Promotion as a fireman depends on age.

Test Statistic: We have a $\chi^2 = 0.15549$ and a p-value = 0.6933.

Rejection Region: We would reject the null hypothesis if our test stat is greater than $\chi^2_{1,0.95} = 3.841$

Conclusion/Interpretation: Since we do not have a test stat that is bigger than our rejection region, we will not reject the null hypothesis. There does not seem to be evidence that promotion as a fireman depends on age.

b) Is your conclusion concerning age discrimination different from your conclusion using the data in Exercise 10.42?

Yes, with a p-value greater than 0.05, it does not appear that the age group is different for promotion potential. However, I am suspicious of lumping everyone into one group that is below 40 years of age.

Chapter 10 HW

Kyle Ligon

April 27, 2018

10.14 a) Would it be appropriate to use a normal approximation in conducting a statistical test of the research hypothesis that over half of persons suffering from chronic pain are over 50 years of age?

```
(424/800)*(800) > 5
```

```
## [1] TRUE
```

```
(376/800)*(800) > 5
```

```
## [1] TRUE
```

Yes, it is appropriate.

b) Using the data in the survey, is there substantial evidence ($\alpha = 0.05$) that more than half of persons suffering from chronic pain are over 50 years of age?

```
old <- prop.test(x = 484, n = 800, p = 0.5, alternative = "greater", correct = FALSE)
old
```

```
##
## 1-sample proportions test without continuity correction
##
## data: 484 out of 800, null probability 0.5
## X-squared = 35.28, df = 1, p-value = 1.428e-09
## alternative hypothesis: true p is greater than 0.5
## 95 percent confidence interval:
## 0.576263 1.000000
## sample estimates:
##      p
## 0.605
```

c) Place a 95% confidence interval on the proportion of person suffering from chronic pain that are over 50 years of age.

```
0.605 - 1.96*(sqrt(0.605*0.395/800))
```

```
## [1] 0.5711244
```

```
0.605 + 1.96*(sqrt(0.605*0.395/800))
```

```
## [1] 0.6388756
```

10.18 a) Place a 95% confidence interval on $p_1 - p_2$, the difference in the proportions of customers purchasing lawnmowers with and without the warranty.

```
mower <- prop.test(x = c(10, 4), n = c(25, 25), alternative = "two.sided",
                  correct = FALSE)
mower$conf.int
```

```
## [1] 0.0001465295 0.4798534705
## attr(,"conf.level")
## [1] 0.95
```

- b) Test the research hypothesis that offering the warranty will increase the proportion of customers who will purchase a mower. Use $\alpha = 0.05$.

```
mower.test <- prop.test(x = c(10,4), n = c(25,25), alternative = "greater",
                        correct = TRUE)
mower.test

##
## 2-sample test for equality of proportions with continuity
## correction
##
## data:  c(10, 4) out of c(25, 25)
## X-squared = 2.4802, df = 1, p-value = 0.05765
## alternative hypothesis: greater
## 95 percent confidence interval:
## -0.001291378  1.000000000
## sample estimates:
## prop 1 prop 2
##  0.40  0.16
```

- c) Are the conditions for using a large-sample test to answer the question in part b) satisfied? If not, apply an exact procedure.

```
0.4*25 > 5

## [1] TRUE
0.6*25 > 5

## [1] TRUE
0.16*25 > 5

## [1] FALSE
0.84*25 > 5

## [1] TRUE
x <- matrix(c(10, 4, 15, 21), byrow = FALSE, nrow = 2, ncol = 2)
fisher.test(x, alternative = "greater")

##
## Fisher's Exact Test for Count Data
##
## data:  x
## p-value = 0.05683
## alternative hypothesis: true odds ratio is greater than 1
## 95 percent confidence interval:
##  0.9619875      Inf
## sample estimates:
## odds ratio
##  3.4113
```

- d) Based on your results from parts a) and b), should the dealer offer the warranty? The dealer should not offer the warranty. The test stat isn't statistically significant at $\alpha = 0.05$.

10.42 (b,d) b) Is the promotion decision for the fireman related to the age of the fireman? Use $\alpha = 0.05$.

```
promote <- matrix(c(9, 41, 29, 39, 34, 46, 12, 38), nrow = 2, ncol = 4)
promote
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    9   29   34   12
## [2,]   41   39   46   38
```

```
promote_chi <- chisq.test(x = promote)
promote_chi
```

```
##
## Pearson's Chi-squared test
##
## data:  promote
## X-squared = 12.796, df = 3, p-value = 0.0051
```

- d) What are some other variables, besides age that needed to be addressed in an age discrimination analysis?
I believe sex/gender, years with the department, and number of previous promotions are adequate variables to split this data on.

10.43 a) Is the age of the fireman related to whether or not the fireman is promoted? Use $\alpha = 0.05$.

```
new_promote <- matrix(c(38, 80, 46, 84), nrow = 2, ncol = 2)
```

```
chi_promote <- chisq.test(x = new_promote, correct = )
chi_promote
```

```
##
## Pearson's Chi-squared test with Yates' continuity correction
##
## data:  new_promote
## X-squared = 0.15549, df = 1, p-value = 0.6933
```

- b) Is your conclusion concerning age discrimination different from your conclusion using the data in Exercise 10.42?
Yes, with a p-value greater than 0.05, it does not appear that the age group is different for promotion potential. However, I am suspect of lumping everyone into one group that is below 40 years of age.