Skyler A. Kuhn
02/19/2016

STAT 543: Assignment 2

HW2 is due Thursday Feb 25th at the beginning of the class (hard copy only).
Note: *All work should be shown for full credit. Rounding rules should be followed. For the IMRaD write-up, double space should be used.* ***For the test statistic, critical value, p-value, and the 95% CI provide the answers in two ways: first without using R and then using R (the answers should be approximately the same) and all calculations should be shown.***

1.) Consider the following tests:

i. H_0: p=0.50 vs. H_a: $p > 0.50$, n = 360, *p-hat* =0.56.

ii. H_0: p=0.50 vs. H_a: $p \neq 0.50$, n = 360, *p-hat* = 0.56.

iii. H_0: p=0.37 vs. H_a: $p < 0.37$, n = 1200, *p-hat* = 0.35.

For each of the tests, use the information given to

1a. (10pts) State the assumptions under which the test statistic (Z) can be performed (One of the assumptions can be verified using the information given, verify this assumption).

That the sample used to calculate p-hat is a representative sample of the population of interest. That the sample size is adequately large enough, and that the subjects in the sample from whom the measurements were taken are independent of one another. Given the following information, we can determine if the sample is of a sufficient size using this formula: $np_O > 5$ and $n(1 - p_O) > 5$. Using this formula (calculations below), we can see that frequencies calculated for i, ii, and iii are above the threshold of 5. This means we are able to use the z-test and chi-square test (in R). (See attached notes for Handwritten Calculations):

1b. (10pts) Compute the value of the test statistic (Z) (without using R and then using R) and determine what distribution the test statistics follow. (See calculations below)

R Code:

- #1B. Using R to find the Test Statistic (z):
- #Example: prop.test(x= ... , n= ..., p= pO, correct=FALSE)
- ##i) H0: p=0.50 vs. Ha: p > 0.50 , n = 360, p-hat =0.56
- prop.test(x=201, n=360, p= 0.50, correct= FALSE) # Output: 2.214

- ##ii)H0: p=0.50 vs. Ha: p ≠ 0.50 , n = 360, p-hat = 0.56
- prop.test(x=202, n=360, p= 0.50, correct= FALSE) #Output: 2.319
- ##iii)H0: p=0.37 vs. Ha: p < 0.37 , n = 1200, p-hat = 0.35
- prop.test(x=420, n=1200, p=0.37, correct= FALSE)
- #Output: 1.435 #(should be negative though)

Looking at the calculations (done by hand), you can see that case i and case ii have an answer around z = 2.28. This means that the sample proportion for these two cases (0.56) is slightly more than 2 standard deviations from the above the hypothesized population proportion of 0.50. In case iii, it was shown that z = -1.435. This means the sample proportion 0.35 is 1.4 standard deviations below the hypothesized population proportion of 0.37. The test statistic z has a standard normal distribution (centered around zero).

(See Handwritten Calculations):

1c. (10pts) Compute the (*p*-value) then compare it to α=0.05 (the significance level). State your decisions for each case using the p- value approach to hypothesis testing.

R Code:

- #1C. Using R to find the p-Value:
- #Example: prop.test(x= ..., n= ..., p= pO, alternative= "two.sided,less,greater", correct=FALSE) #alternative depends on Ha
- ##i) H0: p=0.50 vs. Ha: p > 0.50 , n = 360, p-hat =0.56, α=0.05
- prop.test(x=201, n=360, p= 0.50, alternative= "greater", correct= FALSE) # Output: pvalue= 0.01343
- ##ii)H0: p=0.50 vs. Ha: p ≠ 0.50 , n = 360, p-hat = 0.56, α=0.05
- prop.test(x=202, n=360, p= 0.50, alternative= "two.sided", correct= FALSE) # Output: pvalue= 0.02039
- ##iii)H0: p=0.37 vs. Ha: p < 0.37 , n = 1200, p-hat = 0.35, α=0.05
- prop.test(x=420, n=1200, p= 0.37, alternative= "less", correct= FALSE) # Output: pvalue= 0.07564

Looking at the calculations (done by hand), you can see that case i has a p-value of 0.0122, case ii has a p-value of 0.0188, and case iii has a p-value of 0.0735. Given that the p-values for case i and ii are below than the alpha significance level of 0.05, we reject the null hypothesis in favor of the alternative hypothesis. For case iii, a p-value of 0.0735 was observed; this value is more than the significance value of 0.05, so we fail to reject the null hypothesis.

p-value found from the Table:

- case i: p-value = 0.0136, reject Ho
- case ii: p-value = 0.0204, reject Ho
- case iii: p-value = 0.749, fail to reject Ho in favor of Ha

(See Handwritten Calculations for work)

1d. (10pts) i) Find the critical value (*Z**) using the normal distribution table and using *R* software at the significance level alpha = 0.05.

R code:

- ##i) H0: p=0.50 vs. Ha: p > 0.50 , n = 360, p-hat =0.56
- alpha = 0.05
- ##Right Tailed Test
- z.alpha_right = qnorm(1 - alpha)
- z.alpha_right #Output: Critical Value = 1.644854
- ##ii)H0: p=0.50 vs. Ha: p ≠ 0.50 , n = 360, p-hat = 0.56, α=0.05
- #Two Tailed Test
- z.alpha_two = qnorm(1 - alpha / 2)
- z.alpha_two #Output: Critical Value = 1.959964
- ##iii)H0: p=0.37 vs. Ha: p < 0.37 , n = 1200, p-hat = 0.35, α=0.05
- #Left Tailed Test
- z.alpha_left = qnorm(alpha)
- z.alpha_left #Output: Critical Value = -1.644854

Using the Normal Distribution Table:

Looking at the Normal Distribution Table for a left tailed test with alpha = 0.05, we can see that the critical value is equal to -1.64. For the right tailed test, we get a critical value of 1.64. For the two-tailed test, we get a critical value of 1.96 and -1.96.

ii) State your decisions for each case using the critical value approach to hypothesis testing.

Case i -- H0: p=0.50 vs. Ha: p > 0.50 , n = 360, p-hat =0.56, α=0.05, z = 2.277: Given that the test statistic z = 2.28 and that the critical value is 1.645, we reject the null hypothesis in favor of the alternative hypothesis. This is because the test statistic falls within the rejection region of (1.645, ∞).

Case ii-- H0: p=0.50 vs. Ha: p ≠ 0.50 , n = 360, p-hat = 0.56, α=0.05, z = 2.277 :
Given that the test statistic z = 2.28 and that the critical value and rejection regions for a two tailed test are 1.96 (-∞,1.96) and -1.96 (-1.96,∞), we reject the null hypothesis in favor of the alternative hypothesis. The test statistic falls with this rejection region.

Case iii -- H0: p=0.37 vs. Ha: p < 0.37 , n = 1200, p-hat = 0.35, α=0.05, z= -1.435:
Given that the test statistic z = -1.435 and that the critical value is -1.645, we fail to reject the null hypothesis; it does not fall within the rejection region (-∞,-1.645).

1e. (10pts) Compute the 95% confidence interval for each of the above hypothesis with and without using R) then state your decision for each one based on the 95% confidence interval.

R code (creating a confidence interval function):

- #1E. Using R code to calculate the Confidence Interval:
- ##Case i)
- #Below I am creating a function to calculate the confidence interval
- #Confidence Interval Function:
- confidence.int <- function(phat,po,cv,n){
- pe <- phat
- se <- sqrt(po*(1 - po)/n)
- ci_1 <- pe + (cv * se)
- ci_2 <- pe - (cv * se)
- if (ci_1 < ci_2){
- return (c(ci_1,ci_2))
- }else{
- return (c(ci_2, ci_1))
- } #the if, else statement is just to make sure the value pair it is returning starts with the lowest of the two numbers
- }

Case i, R code:

- #H0: p=0.50 vs. Ha: p > 0.50 , n = 360, p-hat =0.56, alpha = 0.05 (95%), z = 2.277
- confidence.int(phat=0.56, po=0.50, cv= 1.645, n= 360) #calling the confidence interval fucntion
- #Output: CI = (0.5166504, 0.6033496)

Case ii, R code:

- ##Case ii)
- #H0: p=0.50 vs. Ha: p ≠ 0.50 , n = 360, p-hat = 0.56, α=0.05, z = 2.277 :

- confidence.int(phat=0.56, po=0.50, cv= 1.96, n= 360)
- #Output: CI = (0.5083495, 0.6116505)

Case iii, R code:

- ##Case iii)
- #H0: p=0.37 vs. Ha: p < 0.37 , n = 1200, p-hat = 0.35, α=0.05, z= -1.435:
- confidence.int(phat=0.35, po=0.37, cv= -1.645, n= 1200)
- #Output: CI = (0.327073, 0.372927)

Case i:
A 95% confidence interval of the population proportion of subjects is (0.517, 0.603). We are 95% confident that the true proportion of subjects is between 51.7% and 60.3%.
Case ii:
A 95% confidence interval of the population proportion of subjects is (0.508, 0.612). We are 95% confident that the true proportion of subjects is between 50.8% and 61.2%
Case iii:
A 95% confidence interval of the population proportion of subjects is (0.327, 0.373). We are 95% confident that the true proportion of subjects is between 32.7% and 37.3%

(Handwritten Calculations, see attached notes for work):

- case i: CI = (0.517, 0.603)
- case ii: CI = (0.509, 0.611)
- case iii: CI = (0.327, 0.373)

2.) A certain research group asserted that the proportion of smokers who began smoking at a young age was 0.90. It was of interest to test whether the true proportion of all smokers who began smoking at a young age is less than 0.9. Another study considered 850 smokers and found that 689 of them began smoking at a young age.

Ho: p = 0.90 vs. Ha: p < 0.90

2a. (10pts) Following the steps of hypothesis testing covered in class, determine the test statistic required to answer the research question (with and without using R).

Ho: (p = 0.90) vs. Ha: (p < 0.90)

R Code:

- #2A. Using R code to calculate the Test Statistic (z):
- #Example: prop.test(x= ... , n= ..., p= pO, correct=FALSE)
- #H0: p=0.90 vs. Ha: p < 0.90 , x= 689, n = 850, p-hat =0.8106
- prop.test(x=689, n=850, p= 0.90, correct= FALSE)
- # Output: z = 8.689(should be negative though), so z = -8.689.

Assuming that the sample is representative, the subjects in the sample from whom measurements were taken are independent of one another, and the sample size is of a sufficient size, we can calculate the test statistic z. In this case, this test statistic is -8.688. This means that the sample proportion for this cases (0.8106) is more than 8 standard deviations below the hypothesized population proportion of 0.9.

(Handwritten Calculations, see attached notes for work):

- Test statistic: z = -8.688

2b. (10pts) Compute the *p*-value first using the Z-Table and then using R.

R code:

- #2B. Using R code to calculate the p-value:
- #H0: p=0.90 vs. Ha: p < 0.90 , n = 850, p-hat = 0.8106, α=0.05, x= 689
- prop.test(x=689, n=850, p= 0.90, alternative= "less", correct= FALSE)
- #Output: pvalue= 2.2e-16 (very small, basically zero)

p-Value Calculation using the Z-Table (assuming z= -8.688):

- p-value = 0.0, reject Ho

2c. (10pts) State your decision based on p-value at 1% level of significance (a = 1% or 0.01). Write a paragraph about the conclusion in words without using statistical terms for the null and the alternative hypothesis.

Our observed p-value was 2.2e-16 (in R, basically zero). This value is less than the significance level $\alpha = 0.01$, so we reject the null hypothesis (Ho: p = 0.90) in favor of the alternative hypothesis (Ha: p < 0.90). Because of this, we would declare that the evidence suggests that the proportion of all smokers who began smoking at a young age is ***likely*** less than 0.9. Thus, those interested in increasing the proportion of smokers who began at a young age (tobacco companies) should increase their marketing efforts.

2d. (5pts) Find the critical value using the Table and R software at 1% level of significance (a = 1% or 0.01) then state your decision based on the critical value at 1% level of significance (a = 1% or 0.01).

R code:

- #2D. Using R code, find the critical value:
- ##Compute critical values at alpha (in this case 0.01 see directly below)
- alpha = 0.01
- #Left Tailed Test
- z.alpha_left = qnorm(alpha)
- #Output: Critical Value = -2.326348

Critical Value using the Table:

- Critical value= -2.33

Given a significance level of $\alpha = 0.01$, a critical value of -2.33, and a test statistic of z = -8.689, we can reject the null hypothesis (Ho: p=0.90) in favor of the alternative hypothesis (Ha: p< 0.90). This is because our test statistic, z = -8.689, falls within the rejection region [-∞, -2.33]. Because of this, we would declare that the evidence suggests that the proportion of all smokers who began smoking at a young age is ***likely*** less than 0.9.

2e. (10pts)Compute the *99% confidence interval* and interpret this confidence interval (with and without using R) then state your decision based on the 99 % confidence interval (a = 1% or 0.01).

R code (creating a function):

- #2E. Using R code to calculate the Confidence Interval:
- #Below I am creating a function to calculate the confidence interval
- #Confidence Interval Function:

- confidence.int <- function(phat,po,cv,n){
- pe <- phat
- se <- sqrt(po*(1 - po)/n)
- ci_1 <- pe + (cv * se)
- ci_2 <- pe - (cv * se)
- if (ci_1 < ci_2){
- return (c(ci_1,ci_2))
- }else{
- return (c(ci_2, ci_1))
- }} # the if, else statement is just to make sure the value pair it is returning starts with the lowest of the two numbers
- #H0: p=0.90 vs. Ha: p < 0.90 , n = 850, p-hat = 0.8106, α=0.05, x= 689, CV= -2.326348
- confidence.int(phat=0.8106, po=0.90, cv= -2.3263, n= 850) #calling the confidence interval function here
- #Output: CI = (0.7866626, 0.8345374) or (0.787, 0.835)

A 99% confidence interval of the population proportion of smokers who began smoking at a young age is (0.787, 0.835). We are 99% confident that the true proportion smokers who began at a young age is between 78.7% and 83.5%.

(Handwritten Calculations, see attached notes for work):

- 99% CI = (0.787, 0.835)

3.) In a certain county, 72% of residents participate in recycling household waste. To promote recycling, a campaign was done. It was of interest to investigate whether the proportion has increased after the campaign where a survey was conducted and 900 households participated and the survey resulted in 674 of residents participated in recycling.

3a. (10 pts) Following the steps of hypothesis testing, determine the test statistic required to answer the research question (follow the step for hypothesis testing covered in class).

Ho: (p = 0.72) vs. Ha: (p > 0.72)

R code:

- #3A. Using R code to calculate the Test Statistic (z):
- #Example: prop.test(x= ... , n= ..., p= pO, correct=FALSE)

- #H0 :p=0.72 vs. Ha: p > 0.72 , n = 900, x= 674, p-hat =0.74889, po=0.72
- prop.test(x=674, n=900, p= 0.72, correct= FALSE) # Output: z = 1.930

Assuming that the sample is representative, the subjects in the sample from whom measurements were taken are independent of one another, and the sample size is of a sufficient size (which it is), we can calculate the test statistic z. In this case, this test statistic is 2.0045. This means that the sample proportion for this cases (0.75) is more than 2 standard deviations above the hypothesized population proportion of 0.72.

(Handwritten Calculations, see attach notes for work):

- Test statistic: z = 2.005

3b. (5pts) Compute the *p*-value (Using the Z-Table) first then using R.

R code:

- #3B. Using R code to calculate the p-value:
- #Ho :p=0.72 vs. Ha: p > 0.72 , n = 900, x= 674, p-hat =0.74889, po=0.72, z = 2.0045
- prop.test(x=674, n=900, p= 0.72, alternative= "greater", correct= FALSE)
- # Output: pvalue= 0.02679, reject the Ho

p-Value from the Table (z = 2.005) and a = 0.05:

- pvalue = (1 - 0.9772) » 0.0228, reject Ho

3c. (5pts) State your decision based on p-value at 5% level of significance. Write a paragraph about the conclusion in words without using statistical terms for the null and the alternative hypothesis.

Our observed p-value was 0.02679 (in R). This value is less than the significance level α = 0.05, so we reject the null hypothesis (Ho: p = 0.72) in favor of the alternative hypothesis (Ha: p > 0.72). Because of this, we would declare that the evidence suggests that the proportion of residents participating in recycling household waste is **likely** more than 72%. Thus, those interested in increasing the popularity of recycling may wish to continue their campaign efforts.

3d. (5pts) Find the critical value using both the Table and R software at 5% level of significance (a = 5% or 0.05) then state your decision based on the critical value at 5% level of significance (a = 5% or 0.05).

R code:

- #3D. Using R code, find the critical value:
- ##Compute critical values at alpha (in this case 0.05, see directly below)
- alpha <- 0.05
- ##Right Tailed Test
- z.alpha_right <- qnorm(1 - alpha)
- z.alpha_right #Output: Critical Value = 1.644854

Critical Value Calculated from the Table:

- Critical Value = 1.64

Given a significance level of α = 0.05, a critical value of 1.645, and a test statistic of z = 2.0045, we can reject the null hypothesis (Ho: p=0.72) in favor of the alternative hypothesis (Ha: p > 0.72). This is because our test statistic, z = 2.0045, falls within the rejection region [1.654,∞]. Because of this, we would declare that the evidence **suggests** that the proportion of those participating in recycling is ***likely*** greater than 0.72.

3e. (10pts) Compute the *95% confidence interval* (with and without using R), and interpret this confidence interval then state your decision based on the 95 % confidence interval (α = 5% or 0.05).

R code:

- #3E. Using R code to calculate the Confidence Interval:
- #Below I am creating a function to calculate the confidence interval
- #Confidence Interval Function:
- confidence.int <- function(phat,po,cv,n){
- pe <- phat
- se <- sqrt(po*(1 - po)/n)
- ci_1 <- pe + (cv * se)
- ci_2 <- pe - (cv * se)
- if (ci_1 < ci_2){
- return (c(ci_1,ci_2))
- }else{
- return (c(ci_2, ci_1))

- }} #the if, else statement is just to make sure the value pair it is returning starts with the lowest of the two numnbers
- #H0: p=0.72 vs. Ha: p > 0.72 , n = 900, p-hat = 0.7489, α=0.05, x= 674, CV= 1.645
- confidence.int(phat=0.7489, po=0.72, cv= 1.645, n= 900) #calling the confidence interval
- #Output: CI = (0.7242799,0.7735201)

A 95% confidence interval of the population proportion of residents participating in recycling household waste is (0.724, 0.773). We are 95% confident that the true proportion of residents participating in recycling household waste is between 72.4% and 77.3%.

(Handwritten Calculations, see notes for work):

- CI = (0.724, 0.773)

For this problem only (problem 4), the questions should be answered in an IMRaD write-up style (following the steps covered in class for hypothesis testing). For the write-up, use the sample example given at the end of Chapter 2 in the book page 32, and use the Process given after this problem to simplify the write-up procedure and to help you check that your write-up includes all the required pieces.

4. (70pts) In a certain country, it was asserted that 12% of the population does not have health insurance. A certain group in one of the states believes that this rate is lower for that state and was thus interested in testing whether the true proportion of the state's population that do not have health insurance is less than 0.12. A random sample was considered (n= 1550) and it was found that 165 did not have health insurance.

4a. Following the steps of hypothesis testing (covered in class and summarized in the process given below), determine the test statistic required to answer the research question at the 5% level of significance (a = 5% or 0.05) (remember to state the hypothesis and check for the assumptions).

Ho: p = 0.12 vs. Ha: p < 0.12

R code:

- #4A. Using R code to determine the Test Statistic (z):
- #Example: prop.test(x= ... , n= ..., p= pO, correct=FALSE)
- #H0 :p=0.12 vs. Ha: p < 12 , n = 1550, x= 165, p-hat =0.1065, po=0.12
- prop.test(x=165, n=1550, p= 0.12, correct= FALSE)
- # Output: z = -1.641

Assuming that the sample is representative, the subjects in the sample from whom measurements were taken are independent of one another, and the sample size is of a sufficient size (which it is, see notes), we can calculate the test statistic z. In this case, this test statistic is -1.641 (R calculation). This means that the sample proportion for this cases (0.1065) is more than 1.6 standard deviations below the hypothesized population proportion of 0.12.

(Handwritten Calculations, see attached notes for work):

- Test Statistic: z = -1.6356

4b. Compute the *p*-value (Using the Z-Table and using R) and state your decision based on p-value.

R code:

- #4B. Using R code to calculate the p-value:
- #H0 :p=0.12 vs. Ha: p < 0.12 , n = 1550, x= 165, p-hat =0.1065, po=0.12, α = 0.05
- prop.test(x=165, n=1550, p= 0.12, alternative= “less”, correct= FALSE)
- # Output: pvalue= 0.05035
- #fail to reject the Ho

p-Value from the Table:

- pvalue = 0.0505, fail to reject

Our observed p-value was 0.0505 (from the table). This value is greater than the significance level α = 0.05, so we fail to reject the null hypothesis (Ho: p = 0.12) in favor of the alternative hypothesis (Ha: p < 0.12). Because of this, we would declare that the evidence suggests that the proportion of the population in this specific state that does not have health insurance is **likely** 12%. Thus, those interested in increasing the enrollment of individuals without health insurance may wish to continue their efforts.

4c. (10pts) Find the critical value using the Table and R software at 5% level of significance (a = 5% or 0.05) then state your decision based on the critical value at 5% level of significance (a = 5% or 0.05).

R code:

- #4C. Using R to compute Critical Values:
- alpha = 0.05
- #Left Tailed Test
- z.alpha_left <- qnorm(alpha)
- z.alpha_left #Output: Critical Value = -1.644854

Critical Value from the Table:

- Critical Value = -1.64

Given a significance level of α = 0.05, a critical value of -1.645, and a test statistic of z = -1.6356, we fail to reject the null hypothesis (Ho: p=0.12) in favor of the alternative hypothesis (Ha: p < 0.12). This is because our test statistic, z = -1.6356, falls outside of the rejection region [-∞, -1.645]. Because of this, we would declare that the evidence **suggests** that the proportion of the population (for this specific state) without health insurance is ***likely*** equal to 0.12.

4d. (15pts) Compute the *95% confidence interval* and interpret this confidence interval then state your decision based on the 95 % confidence interval (a = 5% or 0.05).

R code (see code above for how I created the function):

- #4D: Using R to Calculate Confidence Intervals
- #Calling the function I created earlier (look above to see code)
- confidence.int(phat=0.1065, po=0.12, cv= -1.645, n= 1550)
- #Output: CI = (0.09292212, 0.12007788)

A 95% confidence interval of the population proportion of people in this state without health insurance is (0.093, 0.120). We are 95% confident that the true proportion of people in this state who do not have health insurance is between 9.3% and 12.0%.

(Handwritten Calculations, see attached notes for work):

- CI = (0.09292, 0.1201)

Follow the process below for Problem 4:

1. State research question in form of testable hypothesis.

2. Determine whether assumptions are met.

   (a) Representative sample. ✓

   (b) Independent measurements. ✓

   (c) Sample size: calculate expected frequencies ($np0$ and $n(1 - p0)$) ✓
   (See hand calculations, the sample is large enough) ✓

3. Summarize data.

   (a) If $np0 > 10$ and $n(1 - p0) > 10$, then report: frequency, sample size, sample proportion and CI. ✓

   (b) If $np0 < 10$ or $n(1 - p0) < 10$, then report: frequency and sample size. ✓

4. Calculate test statistic. ✓

5. Compare test statistic to critical value or **<u>calculate p-value</u>** (Use the p-value approach only) ✓

6. Make decision (reject H0 or fail to reject H0). ✓

7. Summarize with IMRaD write-up ✓

Introduction: Health Insurance protects general population from the unknown. It protects the patient from incurring the inflated costs of medical procedures. Paying for one emergency room visit out-of-pocket my cost upwards of two-thousand dollars. It important that the general population has health insurance for this very reason. If a patient does not pay for a procedure, the hospital or doctor (private practice) will have to write off the bill. This is one of the reasons why healthcare costs are so expensive. Actuaries working the either the hospital or for Health Insurance Providers are then interested in the proportion of individuals in a population that have health insurance; they can use these numbers to determine the price of appointed care. That being said, it was of interest to determine whether the proportion of a state's population with no health insurance was less than 0.12.

Methods: The frequency of people reporting no health insurance enrollment out of 1550 is reported, and the proportion of the population reporting no health insurance is summarized with a sample proportion and a 95% confidence interval. We test the null hypothesis of a 0.12 health insurance un-enrollment rate (Ho: $p = 0.12$) against a one-sided alternative hypothesis that the un-enrollment rate is less than 0.12 (Ha: $p < 0.12$) by using a chi-square-test with a the significance level $a = 0.05$. We will reject the null hypothesis in favor of the alternative hypothesis if the p-value is less than alpha ($a = 0.05$); otherwise, we will fail to reject the null hypothesis. The R statistical software used in coordination with RStudio, Version 0.99.486, was used for all analyses.

Results: Assuming that the sample was representative and the subjects were independent, and the sample was large enough to conduct the statistical analysis (it is).

Out of a sample of 1550 total people, 165 reported no current enrollment with a health insurance provider (p-hat = 0.107, 95%CI: 0.093, 0.120). Using this data, a chi-squared test yielded a p-value of 0.0504, which is more than the stated significance level alpha. Thus, we fail to reject the null hypothesis.

Discussion: The sample data suggest that the proportion of individuals in a state's population who are not currently enrolled with a Health Insurance Provider is likely equal to 0.12. Thus, those working in the field of financial risk assessment--whether it be in a hospital or with a Health Insurance Provider such as Anthem BCBS--may wish to keep the cost care high.