Lab 7: One Sample Hypothesis Tests for the Proportion

Yasmine Pascal

03/16/2022

## Instructions

This document is the template you will use to type up your responses to the lab exercises that you can find on Canvas.

To complete the lab type your BRIEF answers and the R code in the spaces provided below, according to these guidelines:

* In general, if an exercise is numbered you will submit the R code that you used to answer the question in the gray area between the three ` marks.
* If you need to submit an answer in addition to the code type this between the line that starts #### and the beginning of the gray area for the code.
* Do not ever type the View() command in a lab report. This works only in RStudio on your computer, so don’t include that command in your lab report or you will not be able to compile the document, a process called *knitting*.
* To produce a document that you can submit on Canvas just click on the Knit button in the upper left quadrant of RStudio.
* We recommend pressing the Knit button after entering each answer: it’s very common to get errors, and they’re easier to fix if you can easily isolate the code that caused the error.
* The Knit button will create two files: one is a .Rmd file (R Markdown) and the other is an .docx file. You will need to submit both these documents for the lab assignment on Canvas.

## Preliminaries

#### Exercise 1

library(tidyverse)  
library(ggformula)  
library(mosaic)  
library(data1135)

## Preliminary Data Analysis

#### Exercise 2

1.39% of policy holders made a claim.

summary(mortality)

## ID Claimed   
## Min. : 1 Min. :0.00000   
## 1st Qu.: 506 1st Qu.:0.00000   
## Median :1011 Median :0.00000   
## Mean :1011 Mean :0.01385   
## 3rd Qu.:1516 3rd Qu.:0.00000   
## Max. :2021 Max. :1.00000

dim(mortality)

## [1] 2021 2

prop(~(Claimed == 1), data= mortality, success= "TRUE")

## prop\_TRUE   
## 0.01385453

#### Exercise 3

The Central Limit Theorem holds because the observations are independent and both np0 and n(1−p0) are greater than 10.

nrow(mortality) \* .011

## [1] 22.231

nrow(mortality) \* (1-.011)

## [1] 1998.769

#### Exercise 4:

The test statistic is 1.23 and the p-value of the test statistic is 0.1093

prop.test(~Claimed, data = mortality, success = "1", correct= FALSE, p= 0.011, alternative= "greater")

##   
## 1-sample proportions test without continuity correction  
##   
## data: mortality$Claimed [with success = 1]  
## X-squared = 1.5137, df = 1, p-value = 0.1093  
## alternative hypothesis: true p is greater than 0.011  
## 95 percent confidence interval:  
## 0.01018147 1.00000000  
## sample estimates:  
## p   
## 0.01385453

#### Exercise 5:

We fail to reject the original hypothesis (HO). We do this because our p- value (0.1093) is greater than a (0.05).

#### Exercise 6:

There is not statistically significant evidence that the proportion of claims has increased from 1.1%

## Using confidence intervals to carry out two-tailed hypothesis tests

#### Exercise 7:

binom.test(~Claimed, data= mortality, success= "1", ci.method = "Wald", conf.level= .95)

##   
## Exact binomial test (Wald CI)  
##   
## data: mortality$Claimed [with success = 1]  
## number of successes = 28, number of trials = 2021, p-value < 2.2e-16  
## alternative hypothesis: true probability of success is not equal to 0.5  
## 95 percent confidence interval:  
## 0.008758506 0.018950549  
## sample estimates:  
## probability of success   
## 0.01385453

#### Exercise 8:

We fail to reject the original hypothesis because the hypothesized value of the population parameter (0.014) falls within the confidence interval [0.009,0.019]. There is not practically significant evidence that the mortality rate has changed.