Lab 8: Inference for means

Yasmine Pascal

03/25/2022

## 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)

## Exploratory data analysis

#### Exercise 2

Each row represents a person who viewed the ad, either on Facebook or Instagram

dim(AdsManager)

## [1] 2667 6

#### Exercise 3

Out of the 2667 who viewed the ad, 2462 people did not click on the ad and 2520 people did not make a purchase.

summary(AdsManager)

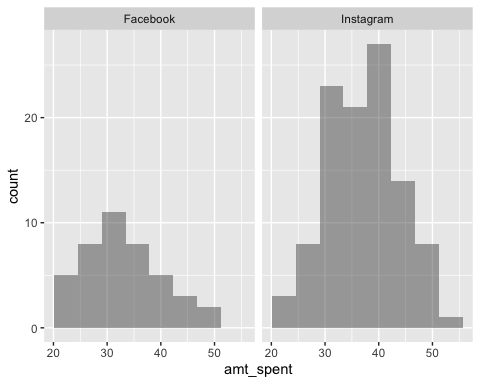
## platform age\_group device\_type clicked   
## Length:2667 18 - 24:882 Length:2667 Length:2667   
## Class :character 25 - 34:974 Class :character Class :character   
## Mode :character 35 - 49:626 Mode :character Mode :character   
## 50+ :185   
##   
##   
##   
## pageviews amt\_spent   
## Min. :0.000 Min. :20.57   
## 1st Qu.:2.000 1st Qu.:30.80   
## Median :3.000 Median :36.44   
## Mean :2.629 Mean :36.07   
## 3rd Qu.:3.000 3rd Qu.:41.32   
## Max. :6.000 Max. :51.77   
## NA's :2462 NA's :2520

#### Exercise 4

This histogram shows that Instagram users weren’t only clicking on the ads and making more purchases than Facebook users, Instagram users also made more expensive purchases.

gf\_histogram(~amt\_spent| platform, data= AdsManager, bins = 8)

## Warning: Removed 2520 rows containing non-finite values (stat\_bin).



#### Exercise 5

On average, Instagram users spent more ($37.31) than Facebook users ($32.96).

favstats(~amt\_spent | platform, data= AdsManager)

## platform min Q1 median Q3 max mean sd n missing  
## 1 Facebook 20.57 27.245 32.445 37.12 48.41 32.96024 7.330657 42 868  
## 2 Instagram 20.87 31.960 37.670 41.87 51.77 37.31200 6.718973 105 1652

## Inference

#### Exercise 6

We can assume that the observations in each group is independent, because it is unlikely that one person’s purchase affects or influences another person’s purchase. We can also assume that the groups are independent of each other because it is unlikely that a users purchase from Facebook affects another user’s purchase on Instagram. Finally the sample size of each group is larger than 30. Facebook has 42 users who made purchases, while Instagram has 105 users.

#### Exercise 7

We reject the null hypothesis, because the p-value (2.2e-16) is smaller than alpha (.05). There is statistically significant evidence that purchasers who saw the ad on Instagram spent more per visit than purchasers who saw the ad on Facebook.

t.test(amt\_spent ~ platform, data= AdsManager, mu = 63, alternative= "less")

##   
## Welch Two Sample t-test  
##   
## data: amt\_spent by platform  
## t = -51.514, df = 70.065, p-value < 2.2e-16  
## alternative hypothesis: true difference in means between group Facebook and group Instagram is less than 63  
## 95 percent confidence interval:  
## -Inf -2.172374  
## sample estimates:  
## mean in group Facebook mean in group Instagram   
## 32.96024 37.31200

#### Exercise 8

Research question: Is there a difference in the number of pages users clicked on beyond the landing page, between mobile devices and other devices?

#### Exercise 9

Define any notation (like ) you use in H0 and HA below: mu\_m = μmobile mu\_d= mean of desktop and tablet =/ not equal

H0:mu\_m - mu\_d = 0

HA:mu\_m - mu\_m =/ 0

#### Exercise 10

We fail to reject the null hypothesis because the p-value (0.8819) is greater than alpha (.05). There is not statistically significant evidence that there is a difference in the number of pages users clicked on beyond the landing page between mobile devices and other devices.

t.test(pageviews ~(device\_type == "mobile"), data= AdsManager, mu= 0, alternative = "two.sided")

##   
## Welch Two Sample t-test  
##   
## data: pageviews by device\_type == "mobile"  
## t = 0.149, df = 77.868, p-value = 0.8819  
## alternative hypothesis: true difference in means between group FALSE and group TRUE is not equal to 0  
## 95 percent confidence interval:  
## -0.4070855 0.4729478  
## sample estimates:  
## mean in group FALSE mean in group TRUE   
## 2.653846 2.620915