Lab4

# Project Goals

In this lab assignment, you will be asked to:

* Construct a 99% confidence interval
* Use bootstrapping to estimate a 99% confidence interval
* Compare confidence intervals of different bootstrap sample sizes
* Compare confidence intervals for different levels of confidence.

Open a new RMarkdown file entitled lab4\_lastname. Copy the code exactly as it is written below. Answer each question fully. Save as a Word file and upload the Word file to Canvas under the assignment submission. Make sure that you include all code and resulting output from the code.

# Construct Confidence Intervals

1. The drug Eliquis (apixaban) is used to help prevent blood clots in certain patients. In clinical trials, among 5924 patients treated with Eliquis, 153 developed the adverse reaction of nausea (based on data from Bristol-Myers Squibb Co.). Construct a 99% confidence interval for the true proportion of people who take Eliquis and develop nausea in R. State this confidence interval and interpret it in the context of the problem. You will construct this in R two ways: One is through the process of writing our own function and, two, using a function within R. Compare those two confidence intervals. Are they equivalent? Bonus point if you can tell me why they are not.

# a. We can write our own functions in R  
# Credit: https://www.r-bloggers.com/calculating-confidence-intervals-for-proportions/  
# For the function, you must specify the variables that your function will use  
  
confidence <- function(n, p, z){  
 # Create an empty list for our confidence interval  
 out <- list()  
 # Formulas for calculating confidence interval for single proportion, lb is the lower bound and ub is the upper bound of the confidence interval.  
 out$lb <- p - z\*sqrt((p\*(1-p))/n)  
 out$ub <- p + z\*sqrt((p\*(1-p))/n)  
 out  
}  
  
# qnorm takes the lower bound of a standard normal distribution to find the critical z-value  
z <- qnorm(.005,lower.tail=FALSE)  
  
# We then substitute the values from the problem to get our confidence interval.  
confidence(n = 5924, p = 153/5924, z)

#b. We can use the function prop.test()

prop.test(x=153, n=5924, p=0.5, alternative = "two.sided", conf.level = .99 )

# Bootstrapping

1. Use the bootstrap method to construct a 99% confidence interval estimate of the proportion of patients who experience nausea. Use 1000 bootstrap samples. How does the result compare to the confidence intervals found in part A? Your answer simply needs to address its similarity or difference. Do they seem “close”?

# Original Sample - "A" for number of adverse reaction, "B" for no adverse reaction  
smp <- c(rep("A", 153), rep("B", 5924-153))  
  
# Initialize the storage of our sample statistic, single proportion   
bootstrap.stats <- c()  
  
# Create a "for loop" to run 1000 bootstrap samples   
for (i in 1:1000) {  
 # Sampling with replacement of the sample we created above  
 bootstrap <- sample(smp, replace = TRUE)   
 # Calculating the bootstrap statistics of each bootstrap sample  
 bootstrap.stats <- c(bootstrap.stats, sum(bootstrap == "A")/5924)  
}  
  
#Calculating the 99% confidence interval using our bootstrap samples   
quantile(bootstrap.stats, probs = c(0.005, 0.995))

# Comparing Bootstrap Samples

1. Repeat the process of constructing a 99% confidence interval estimate using 10000 bootstrap samples. Compare and discuss your confidence intervals found using 1000 bootstrap samples and 10000 bootstrap samples.

# The following code is the same as above, but we will be using 10000 samples  
smp <- c(rep("A", 153), rep("B", 5924-153))  
bootstrap.stats <- c()  
  
# Change number of bootstrap samples from 1000 to 100000  
for (i in 1:10000) {  
 bootstrap <- sample(smp, replace = TRUE)   
 bootstrap.stats <- c(bootstrap.stats, sum(bootstrap == "A")/5924)  
}  
  
#Calculating the 99% confidence interval using our bootstrap samples   
quantile(bootstrap.stats, probs = c(0.005, 0.995))

1. Redo problem 3, but this time with a 95% level of confidence. Compare to the intervals in problem 3. Are they wider or narrower? Explain in complete sentences.

