Assignment 6 - Power, Sample Size and Inference for a Population Proportion. Due October 21, 11:59pm 2018

EPIB607 - Inferential Statistics[a]

[a]Fall 2018, McGill University

This version was compiled on October 14, 2018

In this assignment you will practice conducting inference for a one sample proportion as well as conducting power and sample size calculations. Answers should be given in full sentences (DO NOT just provide the number). All figures should have appropriately labeled axes, titles and captions (if necessary). Units for means and CIs should be provided. All graphs and calculations are to be completed in an R Markdown document using the provided template. You are free to choose any function from any package to complete the assignment. Concise answers will be rewarded. Be brief and to the point. Please submit both the compiled HTML report and the source file (.Rmd) to myCourses by October 21, 2018, 11:59pm. Both HTML and .Rmd files should be saved as 'IDnumber_LastName_FirstName_EPIB607_A6'.

Power | Sample Size | Binomial Distribution | One sample proportion | Bootstrap

Template

The `.Rmd` template for Assignment 6 is available here

1. Power and sample size calculations - 1

Suppose we wished to assess, via a formal statistical test, whether (at an *population*, rather than an individual, level) a step-counting device or app is unbiased (H_0) or under-counts (H_1). Suppose we will do so the way Case et al. did, but measuring n persons just once each. We observe the device count when the true count on the treadmill reaches 500. The statistical test will declare the test 'positive' (the departure from 500 is *statistically significant*) if the mean of the n observations is *below* $500 - 1.96 \times s/n^{1/2}$, where s is the SD of the n observations, and the 1.96 (assume the ultimate n will be large enough that the t and z distributions are interchangeable) is chosen to give the test a type I error of 5%. Since each person is only being measured once, we will not be able to distinguish the genuine between-person variance, σ^2_B from the within-person variance, σ^2_W. Thus the sample variance, s^2 will be an estimate of $\sigma^2 = \sigma^2_B + \sigma^2_W$.

1. Using a planned sample size of $n = 25$, and $\sigma = 60$ steps as a pre-study best-guess as to the s that might be observed in them, calculate the critical value $500 - 1.96 \times s/n^{1/2}$.

2. Now imagine that in an infinite sample, the mean would not be the null 500, but $\mu = 470$. Calculate the probability that the mean in the sample of 25 will be less than this critical value. (Use the same s for the alternative that you used for the null.)

3. By trial and error, or better still by deriving a general formula, adjust the n until this probability (i.e., the power) is 80%. Show the 2 probabilities in a diagram, as was done in Figure 4 in section 4.3.2 of the notes.

4. Show that this n satisfies the equality
$$1.96 \times SE_{null}[\bar{y}] + 0.84 \times SE_{alt}[\bar{y}] = 30.$$

2. Attitudes toward school

The Survey of Study Habits and Attitudes (SSHA) is a psychological test that measures the motivation, attitude toward school, and study habits of students. Scores range from 0 to 200. The mean score for U.S. college students is about 115, and the standard deviation is about 30. A teacher who suspects that older students have better attitudes toward school gives the SSHA to 25 students who are at least 30 years of age. Their mean score is $\bar{y} = 132.2$ with a sample standard deviation $s = 28$.

a. The teacher asks you to carry out a formal statistical test for her hypothesis. Perform a test, provide a 95% confidence interval and state your conclusion clearly.
b. What assumptions did you use in part (a). Which of these assumptions is most important to the validity of your conclusion in part (a).

3. Does a full moon affect behavior?

Many people believe that the moon influences the actions of some individuals. A study of dementia patients in nursing homes recorded various types of disruptive behaviors every day for 12 weeks. Days were classified as moon days if they were in a 3-day period centered at the day of the full moon. For each patient, the average number of disruptive behaviors was computed for moon days and for all other days. The hypothesis is that moon days will lead to more disruptive behavior. We look at a data set consisting of observations on 15 dementia patients in nursing homes (available in the `fullmoon.csv` file):

```
fullmoon <- read.csv("fullmoon.csv")
```

```
#     patient moon_days other_days
#  1        1      3.33       0.27
#  2        2      3.67       0.59
#  3        3      2.67       0.32
#  4        4      3.33       0.19
#  5        5      3.33       1.26
#  6        6      3.67       0.11
#  7        7      4.67       0.30
#  8        8      2.67       0.40
#  9        9      6.00       1.59
#  10      10      4.33       0.60
#  11      11      3.33       0.65
#  12      12      0.67       0.69
#  13      13      1.33       1.26
#  14      14      0.33       0.23
#  15      15      2.00       0.38
```

a. Calculate a 95% confidence interval for the mean difference in disruptive behaviors. State the assumptions you used to calculate this interval.
b. Calculate a 95% bootstrap confidence interval for the mean difference in disruptive behaviors and compare to the one obtained in part (a). Comment on the bootstrap sampling distribution and compare it to the assumptions you made in part (a).
c. Test the hypothesis that moon days will lead to more disruptive behavior. State your assumptions and provide a brief conclusion based on your analysis.
d. Find the minimum value of the mean difference in disruptive behaviors ($\bar{y}$) needed to reject the null hypothesis.
e. What is the probability of detecting an increase of 1.0 aggressive behavior per day during moon days? *Hint: calculate the probability of the event calculated in part (d) using a normal distribution with $\mu = 1$ and $\sigma =$ the standard error of the mean*

4. How deep is the ocean?

This question is based on the in-class Exercise on sampling distributions and builds on Question 4 from Assignment 4. For your sample of $n = 20$ of depths of the ocean

a. Calculate a 95% Confidence interval using the t procedure
b. Plot the qnorm, bootstrap, and t procedure confidence intervals on the same plot and comment on the how the t interval compares to the other 2 intervals. You may use the `compare_CI` function provided below to produce the plot.

```
compare_CI <- function(ybar, QNORM, BOOT, TPROCEDURE,
                       col = c("#E41A1C","#377EB8","#4DAF4A")) {

  dt <- data.frame(type = c("qnorm", "bootstrap", "t"),
                   ybar = rep(ybar, 3),
                   low = c(QNORM[1], BOOT[1], TPROCEDURE[1]),
                   up = c(QNORM[2], BOOT[2], TPROCEDURE[2])
  )

  plot(dt$ybar, 1:nrow(dt), pch = 20, ylim = c(0, 5),
       xlim = range(pretty(c(dt$low, dt$up))),
       xlab = "Depth of ocean (m)", ylab = "Confidence Interval Type",
       las = 1, cex.axis = 0.8, cex = 3)

  abline(v = 37, lty = 2, col = "black", lwd = 2)
  segments(x0 = dt$low, x1 = dt$up,
           y0 = 1:nrow(dt), lend = 1,
           col = col, lwd = 4)

  legend("topleft",
         legend = c(eval(substitute( expression(paste(mu," = ",37)))),
                    sprintf("qnorm CI: [%.f, %.f]",QNORM[1], QNORM[2]),
                    sprintf("bootstrap CI: [%.f, %.f]",BOOT[1], BOOT[2]),
                    sprintf("t CI: [%.f, %.f]",TPROCEDURE[1], TPROCEDURE[2])),
         lty = c(1,1,1,1),
         col = c("black",col), lwd = 4)
}

# example of how to use the function:
compare_CI(ybar = 36, QNORM = c(25,40), BOOT = c(31, 38), TPROCEDURE = c(28, 40))
```