Module 06 HW

Daniel Jackson

October 4th, 2023

Word Document

**Exercises 5.2, 5.4, 5.10, 5.22**

**Exercise 5.2 Identify Parameter, part II**

1. Proportion – Categorical response.
2. Mean – Numerical response.
3. Proportion – Categorical response.
4. Proportion – Categorical response.
5. Mean – Numerical response.

**Exercise 5.4 Unexpected expense**

1. 765 adults in the United States.
2. Parameter being estimated is that 322 adults could not cover a $400 unexpected expense without borrowing money or going into debt.
3. Point estimate for the parameter: 322/765 = 0.421. p^ = 0.421
4. We use sampling error to measure uncertainty of point estimate.
5. Check conditions to see if we can use central limit theorem to confirm normal like distribution.

Since the sample is a random sample, we can conclude each observation is independent.

n X p^= 765 X 0.421 = 322.07 which is greater than or equal to 10

n(1-p^) = 765 X (1-0.421) = 442.94 which is greater than or equal to 10

Since both conditions are true, we can assume point estimation distribution follows normal distribution.

Find sampling error:

SE = sqrt(p^(1-p^/n) = sqrt(0.2438/765) = 0.018

1. Using R-studio:

pnorm(0.421, mean = 0.5, sd = sqrt(0.25/765))

This returns:

6.210504e-06. This is an insanely small probability so she should be surprised with her findings!

1. p = 0.4

SE = sqrt(0.4(1-0.4)/765) = 0.0177

The resulting value does not change much. There is a difference of 0.0003 between the two sampling errors.

**Exercise 5.10 Twitter users and news, part II**

1. Since α = 0.01, that means we are looking for 99% confidence interval. We know SE = 0.024

We will use point estimate +/- z\* X SE

We know z\* = 2.58 for 99% confidence intervals

p^ = 0.5

0.5 +/- 2.58 X 0.024

(0.439, 0.561)

We are 99% confident that between 43.9% and 56.1% U.S. adult Twitter users get some news from Twitter. So, the statement is false: Since a portion of the 99% confidence interval is below, 50%, the data does not provide statistically significant evidence that more than half of U.S. adult Twitter users get some news from Twitter.

1. False, knowing that the SE is 2.4% gives us no insight to 97.6% of all U.S. adult Twitter users being included in the study.
2. False, more data will always reduce the sampling error of the point estimate. In the SE equation, the sample size, n, is on the dominator within the square root. Therefore, the higher the sample size, the lower the sampling error.
3. False, a 99% confidence interval will be wider than a 90% confidence interval. The more confident that are looking to be, the wider the interval will be. To test this, let’s do part b.) again with a 90% confidence interval. This results in a confidence interval between 46.1% and 53.9%. This range delta is 7.8% compared to delta range of 12.2% in part b.).

**Exercise 5.22 Getting enough sleep**

Null hypothesis H0 :p=0.5

Alternative hypothesis HA :p̸=05

p-nought = 0.5

p^ = 0.723

Look at sampling distribution to see if it p-nought follows normal distribution.

Since sample population was randomly sampled, we know each observation is independent of one another.

n X p-nought = 400 X 0.5 = 200 which is greater than or equal to 10

p-nought X (1- p-nought) = 400 X (1-0.5) = 200 which is greater than equal to 10

Both conditions are met so we know p-nought somewhat follows normal distribution.

Mean is 0.5.

SE = sqrt(0.25/400) = 0.025

Compute z-score of p^ = (0.723-0.5)/0.025 = 8.92

Using Rstudio, we know pnorm(8.92) = 1. Since we are looking for tail area for p^, we subtract: 1-pnorm(8.92) to get 0. Since p-value = 0 is less than 0.01, we reject null hypothesis that p = 0.5.