MA256 Lesson 4 - Significance - Strength of Evidence (1.3, 1.4, 1.5)

Define and describe the standardized statistic for a proportion:

z =

When can we use a theory-based approach to calculate a p-value? (for a one-proportion z-test)

Using the theoretical approach, what is the expected standard deviation of a null distribution? Why?

SD =

Using the theoretical approach, how do we calculate our standardized statistic?

z =

Guidelines for using p-value/standardized test statistic for strength of evidence.

	p-value	standardized test-stat
Weak Evidence against the null		
Moderate Evidence against the null		
Strong Evidence against the null		
Very Strong Evidence against the null		

What factors impact the strength of evidence?

Using a theoretical z-score, how do we calculate the p-value? (Hint: Take a look at the Course Guide - Block 2/Lesson 3-4)

1) Twenty-eight firsties miss recall formation because they partied hard over the weekend, but blamed their lateness on a flat tire. The TAC team brings them into their office and asks them one question which will determine if they get hours or not. Which tire went flat? This question works if we assume that each tire is equally likely to be chosen, but it has been proposed that people tend to answer "right front" more often. The results of the cadets' responses are shown below.

Left Front	Left Rear	Right Front	Right Rear
6	4	14	4

a) What is the research question?

b) Identify the observational units in this study.

c) Describe the parameter of interest (in words).

d) State the appropriate null and alternate hypotheses to be tested, both in words and symbols.

```
# pi <- XXXXX
```

e) What is our observed statistic? What is our sample size?

```
# n <- XXXXX
# phat <- XXXXX
# c(n, phat)
```

f) Does our sample meet the validity conditions to use a theory-based test?

g) Assume that validity conditions are met. What is the theory-based standardized statistic and p-value?

```
# z <- XXXXX
# z
# XXXXX  # use a function of pnorm() here
```

h) Summarize the conclusion that you draw from this study and your analysis. Explain your reasoning.

i) Simulate the results of the cadets answering the TAC team. Use 1000 replications. List the simulated p-value and interpret the strength of evidence (as compared to your answer above).

```
# set.seed(XXXXX)
# M <- XXXXX
# pi <- XXXXX
# ncadets <- XXXXX
# n.succ <- XXXXX
# phat <- XXXXX
# RES <- data.frame(res = rep(NA, M))  # create a data frame to hold the results of the simulation
#
```

```
# for(i in 1:M){
#   myobs <- rbinom(1, XXXXX, XXXXX)
#   RES$res[i] <- myobs
# }
# RES %>% ggplot(aes(x = XXXXX)) + geom_histogram() +
#   geom_vline(xintercept = n.succ, color = "red")
#
# sum(XXXXXs >= XXXXX) / XXXXX
```

Suppose this study were repeated with only 14 cadets and 7 of them answered "front right." Use this reduced sample scenario to answer parts *j* through *l*.

j) What would you expect to happen to the strength of evidence against the null hypothesis in this case?

k) Does our reduced sample meet the validity conditions to use a theory-based test?

l) Using our reduced sample, calculate the new p-value and standardized statistic. Specify if you simulated or used theoretical methods.

```
# PUT SOMETHING USEFUL HERE
```

2) An article published in *College Mathematics Journal* (Eyler, Shalla, Doumaux, and McDevitt, 2009) found that players tend to not prefer scissors when playing Rock-Paper-Scissors. You want to test if people really choose scissors less, and conduct a test. You played 120 games and your friend chose scissors 31 times.

a) List the null and alternate hypothesis in words and symbols.

b) Calculate the mean and standard deviation associated with your null distribution. Calculate the observed proportion that chose scissors.

```
# pi <- XXXXX
# n <- XXXXX
# sd.RPS <- sqrt(XXXXX)
# phat <- XXXXX
# c(pi, sd.RPS, phat)
```

c) What is the standardized statistic (z) for your test? p-value? Comment on the strength of evidence.

```
# z <- XXXXX
# pval <- XXXXX
# c(z, pval)
```

d) If you repeated the test another 240 times and your friend chose scissors the same proportion of times ($\hat{p} = 0.258333$), would you expect your strength of evidence to increase, decrease, or stay the same?

e) If we repeated the experiment with a different friend and our sample size stayed the same (120), but the number of times he chose scissors was 38, would the strength of evidence increase, decrease, or stay the same?

f) What if we used our original experimental data ($\frac{31}{120}$ scissors), but instead we wanted to do a two-sided test instead of a one-sided test. Would our strength of evidence increase, decrease, or stay the same? Use R to verify.

```
# PUT SOMETHING USEFUL HERE
```