MA206, Lesson 9 - Size of Effect

Review: How do you calculate the standardized statistic for one proportion?

$z = \frac{\hat{p}-\pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}}$

Review: What is the R code for a two-tailed test with the z-statistic?

2 * (1 - pnorm(abs(z)))

Define **Significance Level**

A value used as a criterion for deciding how small a p-value needs to be to provide convincing evidence to reject the null hypothesis.

It is the cut line, the accept/reject criteria.

Define **Confidence Interval**

An inference tool used to estimate the values of the parameter, with an associated measure of uncertainty due to the randomness in the sampling method.

All the values we don't reject with a given Significance Level.

How do we interpret a (95%) confidence interval?

"We are (95%) confident that the true parameter lies within the range (a, b)"

How do we calculate a confidence interval for one proportion theoretically?

$\hat{p}\pm$ Margin of Error

$\hat{p} \pm Multiplier \times StandardError(SE)$

$\hat{p} \pm Multiplier \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, if validity conditions are met (10 per category)

What is the R code to find the multiplier for any significance level α?

multiplier <- qnorm(1 - $\frac{\alpha}{2}$)

Common multipliers: 90% = 1.645, 95% = 1.960, 99% = 2.576

What is the **Margin of Error?**

The Margin of Error is the Multiplier (M) multiplied by the Standard Error.

Multiplier $\times$ SE = qnorm(1 - $\frac{\alpha}{2}$) * sqrt(phat * (1 - phat) / n)

1) Students at Hope College were tested to see if they could determine the difference between tap water and bottled water. Of the 63 students tested, 42 correctly identified which water was which. A plausible value simulation was conducted with results represented in the table below.

a) What is the 95% confidence interval based on the results? Include an interpretation of your findings.

Null Hypothesis	0.52	0.53	0.54	0.55	0.56	0.57	0.68
p-value	0.0250	0.0330	0.0530	0.0850	0.0980	0.150	0.1840
Reject or Don't Reject							

Null Hypothesis	0.72	0.73	0.74	0.75	0.76	0.77	0.78
p-value	0.3130	0.2680	0.1920	0.1510	0.0930	0.0770	0.0490
Reject or Don't Reject							

Given these results, the 95% confidence interval is (0.54, 0.77).
We are 95% confidence that the true proportion of students who can correctly identify tap water from bottled water lies between 0.54 and 0.77.

b) What is the 90% Confidence Interval?
(0.57, 0.75)

2) Most people are right handed, and even the right eye is dominant for most people. Developmental biologists have suggested that late-stage human embryos tend to turn their heads to the right. In a study reported in Nature (2003), German bio-psychologist Onur Güntürkün conjectured that this tendency to turn to the right manifests itself in other ways as well, so he studied kissing couples to see which side they leaned their heads to while kissing. He and his researchers observed kissing couples in public places such as airports, train stations, beaches, and parks in Germany. They were careful not to include couples who were holding objects such as luggage that might have affected which direction they turned. For each kissing couple observed, the researchers noted whether the couple leaned their heads to the right or to the left. They observed 124 couples, ages 13 to 70 years.

a) Identify the observational units in this study.
Each kissing couple.

b) Identify the variable recorded in this study. Classify it as categorical or quantitative.
Whether each couple leaned right or left while kissing. This is a categorical variable.

c) Suppose we want to know the true long-run proportion of couples that kiss right in Germany. Would this be a statistic or a parameter? What symbol is used to represent this proportion of the population?
This would be a parameter, represented by π

d) Do we know the exact value of the long run proportion of couples kissing right based on the data? Explain.
No, we don't know the exact value of π. The population parameter is an unknown quantity, but we want to infer it using statistics for a best guess.

e) State the appropriate null and alternative hypotheses, both in words and in terms of the parameter π, for testing the conjecture that kissing couples tend to lean their heads to the right more often.

$H_0 : \pi = 50\%$. The null hypothesis is that the long run proportion of couples who lean their heads to the right is 50%.

$H_a : \pi > 50\%$. The alternate hypothesis is that the long run proportion of couples who lean their heads to the right is greater than 50%.

Güntürkün's findings are compiled in the "Kissing.csv" dataset available on Teams. Load the data into R to analyze it, using the TidyVerse Tutorial and Course Guide to help with coding as needed.

f) Calculate the sample proportion of the observed couples who leaned their heads to the right while kissing. Also indicate the symbol used to denote this value.

$\hat{p} = \frac{80}{124} = 0.645$

g) Do we meet the validity conditions to conduct a theoretical test? Justify your answer.

Yes, we meet the validity conditions because 80 right and 44 left are both larger than 10.

h) If we wanted to do strength of evidence testing using theoretical methods, which test would we use?

As we are assessing categorical variables, we would use a one proportion z-test.

i) Using the applet to simulate, assess the strength of evidence that the sample data provide for Güntürkün's conjecture that kissing couples tend to lean their heads to the right more often than they would by random chance. Report the approximate p-value and summarize your conclusion about this strength of evidence.

Should be roughly 0 - this gives very strong evidence that the proportion is not equal to 0.5

j) Now use theoretical methods to test whether the data provide evidence that the probability that a couple leans their heads to the right while kissing (π) is different from 0.60 ($H_a \neq$. (Note that this question changes both the null hypothesis and the alternate hypothesis) Report the standardized statistic, p-value, and comment on the strength of evidence.

z = 1.02653

p-value = 0.3046419

This is weak evidence against the null hypothesis, we would say that 0.6 is plausible for the true proportion of couples who kiss while leaning right.

k) Using theoretical methods, calculate the 95% confidence interval. Interpret your results.

Our confidence interval is $\hat{p}\pm$ M * SE $= 0.645 \pm qnorm(1 - \frac{0.05}{2}) * \sqrt{\frac{0.645*(1-0.645)}{124}} = (0.5609, 0.7294)$

We are 95% confident that the true parameter lies between 0.5609 and 0.7294.

l) Does your confidence interval include 0.50? Does it include 0.60? Explain how your answers relate to the strength of evidence tests conducted in 9 and 10 above.

Our interval does not include 0.5 but it does include 0.6. This makes sense since, with a signifance level of 0.05, we already concluded in 9 that 0.50 is not feasible, but 0.60 is.

m) Now suppose we were to use a significance level of 0.01 instead of 0.05. How would you expect the interval of plausible values to change: wider, narrower, or no change? Explain your reasoning.

We expect it to get wider, we are including more values in our confidence interval

n) Calculate the corresponding 99% confidence interval. Did it behave as expected?

Our confidence interval is $\hat{p} \pm \text{M * SE} = 0.645 \pm qnorm(1 - \frac{0.01}{2}) * \sqrt{\frac{0.645*(1-0.645)}{124}} = (0.534, 0.756)$
This is a wider interval than our 95% confidence interval above, as expected.

o) Based on your 99% confidence interval, what can be said about the p-value for testing a null hypothesis of 0.78?

As it falls outside of our 99% confidence interval, we know that the p-value must be less than 0.01 for a two-sided test and a null hypothesis of $\pi = 0.78$.

p) Can your results be generalized? Explain your reasoning.

We cannot generalize our results as the couples selected were not truly random. We do not know what parts of Germany were surveyed to know what groups were or were not included. Additionally, outgoing couples may have a different tendency than those who do not kiss in public places.

3) In order to determine who would accept a Facebook friend request from someone they didn't know, a student researcher made up a phony Facebook profile that represented a female student at her college. She then sent out 118 friend requests and 61 of these accepted the request.

a) Do we meet validity conditions to use theoretical methods?

Yes we do, we have 61 successes and 57 failures.

b) Find a 90% confidence interval for the proportion of all students that would accept a Facebook request from a female at their college that they did not know.

$\hat{p} \pm Multiplier \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

(0.441, 0.593)

4) In order to determine who would accept a Facebook friend request from someone they didn't know, a student researcher made up a phony Facebook profile that represented a male student at her college. She then sent out 101 friend requests and 18 of these accepted the request.

a) Do we meet validity conditions to use theoretical methods?

Yes we do, we have 18 successes and 83 failures.

b) Find a 99% confidence interval for the proportion of all students that would accept a Facebook request from a male at their college that they did not know.

$\hat{p} \pm Multiplier \times \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

(0.080, 0.276)