# Name:

# Classwork 18: **Yeager, Hirschi, & Josephs’ Respect Model**

1. Let’s review the Central Limit Theorem.

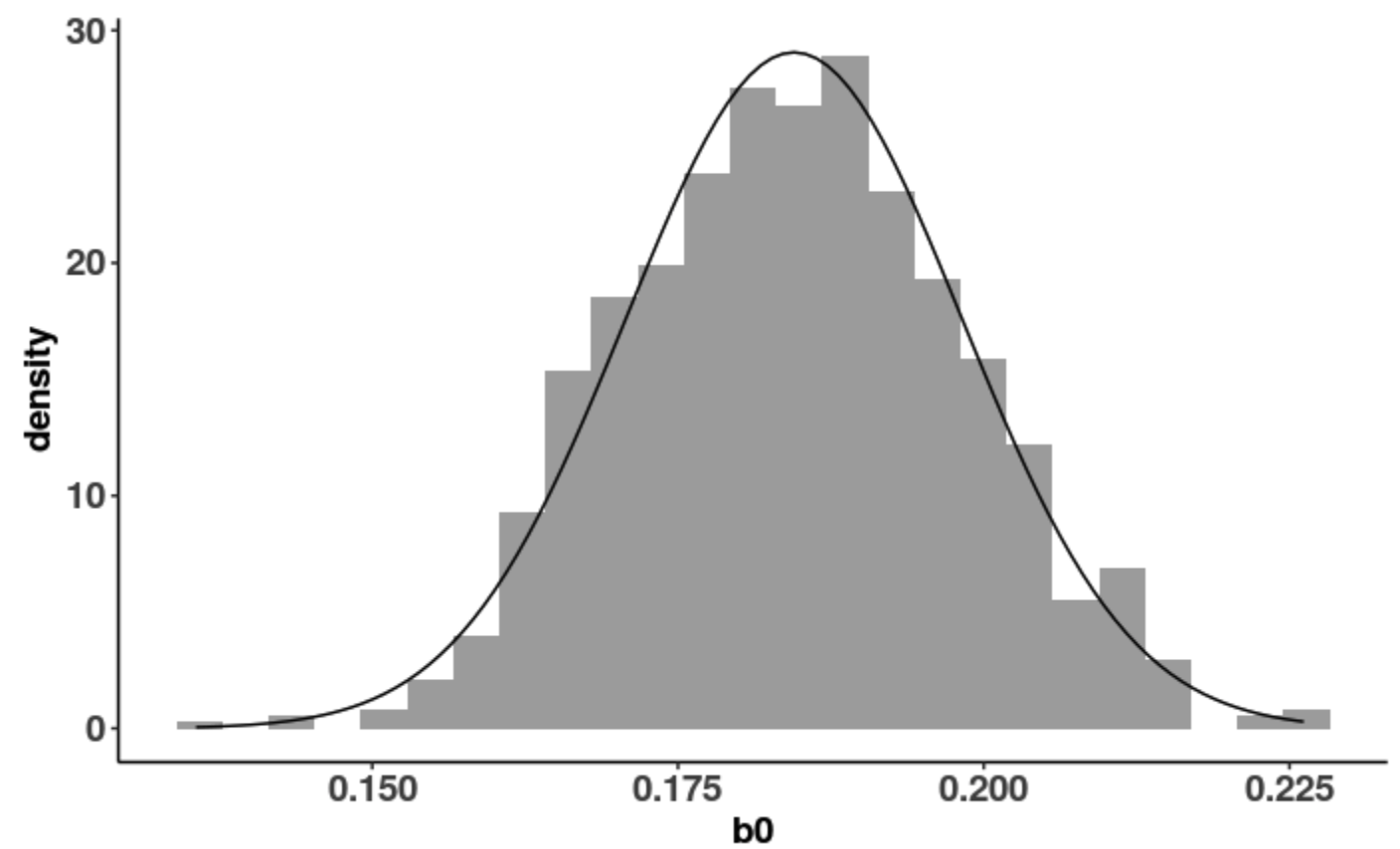
Shape –

Center –

Spread –

1. Which distributions are the CLT about? Sample? DGP? Or Sampling Distributions?

1. Roughly draw the mean and standard error of this sampling distribution that we created in Classwork 17. Then actually calculate the mean and standard error in RStudio.cloud. Does this follow the CLT?

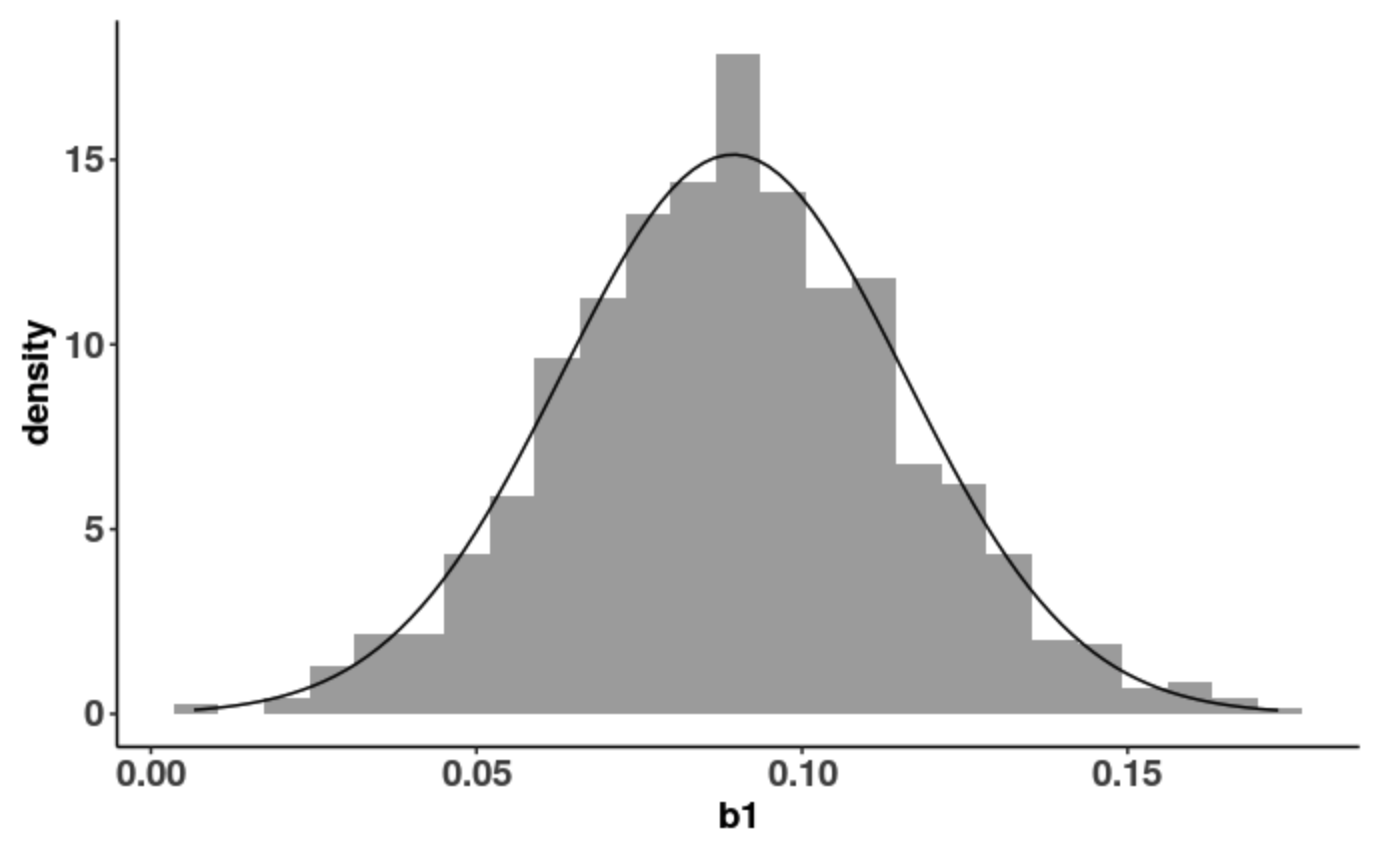


1. Last time we imagined a hardcore skeptic who was like, “NULL MODEL ALL THE WAY!” I want to give you a little bit of a sense of what we mean by “imagining different DGPs” and “pushing around the sampling distributions.” Let’s expand here on the distribution triad with our cut outs. [mini-lecture]
2. Now imagine David Yeager and his team. They are probably like, “Hey, just using a little bit of respectful language might make a difference in how much vegemite students eat. I think the DGP is . What’s your best estimate of and ?” What would you say?
3. Remember this question: When we look at a visualization and think – yeah, this one looks like some of the variation is explained – which parameter estimate best represents the “shift”: or ?
4. So, which parameter in the DGP should we be super duper concerned about?
5. Let’s go back to our best efforts in question 5. We are doing the best we can. What are our chances of being right? Explain your thinking.
6. If we had a similar but slightly different sample, what kind of variation is this? Explained? Unexplained? Random variation? Sampling variation? Variation caused by respect?
7. Now imagine if we really could pluck 1000 random samples that were roughly similar to ours. Could we use that to learn what the true in the DGP is?
8. Here is a new function to try out: b1(veg\_eaten ~ respect\_condition, YeagerData)

\*\*Note: We made **veg\_eaten**, a variable, way back in Classwork 12. You may have named it something slightly different.

What do you think this function does?

1. Now I want you to imagine that the world, “the population,” is just like our sample. Imagine doing this experiment again but we just randomly select a slightly different group of 160 people. How would you use the function **resample()** to do this? Write the code.
2. If we fit the respect model onto this new resampled data, would we come up with the exact same ? Why or why not?
3. Try to use the two functions **b1()** and **resample()** together somehow to select a new group of 160 people and calculate a new . Write the code here.
4. What is result of running this code? What does that number stand for? Does that mean the “shift” in our sample changed?
5. Let’s use the **do()** function to run that code 10 times. What does the **do()** function do? What does this function give you? A single number? A variable? A data frame?
6. What do you notice about these s? How much do you think the s vary?
7. If we ran this code 1000 times, using the **do()** function, what kind of distribution do we get: a sample, the DGP, or sampling distribution?
8. If we visualized this distribution, what do you think the shape, center, and spread of this distribution would be?
9. Go ahead and try to create a visualization of these s. Write the code here.
10. What is depicted in this histogram? What is the shape, center, and spread of this distribution?



1. What is the standard error of the s?
2. Why are the s roughly normally distributed? Is every statistic going to have a normally distributed sampling distribution?
3. Take a look at that one bootstrapped sample right near .15 in the histogram in question 20. What would the best fitting respect model be if that bootstrapped sample was our real sample? Write this in GLM notation.
4. Take a look at that one bootstrapped sample right near 0 in the histogram in question 20. What would the best fitting respect model be if that bootstrapped sample was our real sample? Write this in GLM notation.