

Midterm 1 - Fall 2025

ECON 4753 — University of Arkansas

1. Say you have a sample of observations of some variable. You want to summarize the variable to a stakeholder.
 - (a) Describe two ways you can help someone understand the distribution of a variable and what function you might use in R to do this.
 - (b) In your own words, describe what the concept of the sampling distribution of a statistic is. Why is it helpful to know about the sample distribution of a statistic?
 - (c) Typically, we report 95% confidence intervals. Give an example where someone might want to use a higher level of confidence (e.g 99% or 99.9%)

2. This question is based on our review of statistics. Say you observe a sample of workers from a firm with sample size $n = 100$. You observe their wages w_i and want to estimate the average wage at the firm. You estimate the following statistics in your sample: $\bar{w} = 17.53$ and $\text{var}(w) = 4.2$.
 - (a) Given this information what is the (approximate) sample distribution of the sample mean?
 - (b) Form a 95% confidence interval for your sample mean. Interpret this in words.
 - (c) Another student claims the average worker earns \$17. Using your confidence interval, would you reject this null with a 5% significance level?

Below is the result of a two regressions using data on nutrition information on Starbucks' food items.

3. First, we will look at this regression of the number of calories in the food item on indicators for each food type with 'bakery' beign the omitted group.

```
OLS estimation, Dep. Var.: calories
Observations: 73
Standard-errors: IID
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	368.78049	12.9273	28.527289	< 2.2e-16 ***
type::bistro box	8.71951	31.9934	0.272541	7.8603e-01
type::hot breakfast	-43.78049	31.9934	-1.368422	1.7568e-01
type::petite	-191.00271	30.4699	-6.268568	2.8685e-08 ***
type::sandwich	26.93380	33.8516	0.795644	4.2901e-01

- (a) What is the average amount of calories for food items in the 'petite' type?
- (b) Which food type has the largest number of calories on average?
- (c) What is the difference in average amount of calories for 'bistro box' foods relative to sandwiches? How would you modify this regression to test if the difference is statistically significant?

4. Second, we regress the number of calories in the item on the amount of fat in each item (in grams).

```
OLS estimation, Dep. Var.: calories
Observations: 73
Standard-errors: Heteroskedasticity-robust

             Estimate Std. Error   t value   Pr(>|t|)
(Intercept) 183.2400    20.98025   8.73393 7.2892e-13 ***
fat          11.2768     1.09564  10.29240 1.0109e-15 ***
```

- (a) Interpret the coefficient on 'fat' in words. Comment on its statistical significance.
- (b) Predict the number of calories in a food item with 14g of fat.
- (c) Construct a 95% confidence interval around the slope coefficient. What is the smallest slope that you can not reject with a 5% level of significance?

Finally, we add the number of carbs in that item as a second explanatory variable.

```
OLS estimation, Dep. Var.: calories
Observations: 73
Standard-errors: Heteroskedasticity-robust

             Estimate Std. Error   t value   Pr(>|t|)
(Intercept) 39.10401   13.709974   2.85223 0.0057027 **
fat          9.98090    0.476424  20.94964 < 2.2e-16 ***
carb         3.62562    0.202739  17.88319 < 2.2e-16 ***
```

- (d) Why do you think the slope coefficient on 'fat' went down after controlling for 'carb'?