EDUC 6050

Name:_____

Take Home Exam 2 (Spring 2018)

Score:_____ /100

Directions: Complete all the questions. As the syllabus says you may use your "printed discussion points, homework, and other notes during examinations." Good luck!

Multiple Choice (2.5 pts each)

1. When do you use an ANOVA?

   A. When you have a categorical outcome.
   B. When your sample size is more than 100
   C. When you have 3 or more groups
   D. When you don't know how to use t-tests.

2. Which of the following interpretations of 95% confidence intervals is correct?

   A. A set of plausible values for a population parameter
   B. If the study were repeated, the population mean would be in the confidence interval 95% of the time
   C. If the mean is in the confidence interval, you can reject the null hypothesis
   D. All values within a confidence interval are the population value

3. Would you rather:

   A. Be able to read minds
   B. Both (this is cheating)
   C. Neither, they both sound ridiculous
   D. Impervious to pain and discomfort

4. ANOVA provides an omnibus result. What does omnibus mean here?

   A. It gives you an overall result for differences across groups
   B. It gives you an F statistic
   C. It gives you specific results for each group comparison
   D. It gives you interpretable results regarding group means

5. What is a "post hoc" test?

   A. It is not important to ever use.
   B. It is an analysis that provides a "hoc" estimate.
   C. It is an analysis you use after an omnibus result to find the specific differences
   D. It is an analysis that finishes an ANOVA F statistic.

6. You are doing a study about the influence of having children on depression levels. You do a regression and get an $R^2 = .50$. What is an appropriate interpretation of this result?

   A. Having children causes 50% of the depression found in parents
   B. Parents are 50% more likely to have depression than non-parents
   C. 50% of the variability in depression is accounted for by having children
   D. The results are not significant

7. We hypothesize that both time spent hiking and shoe-type (Jordans or Nike) predict resting heart rate. We want to test this hypothesis. What approach can we use?

   A. Multiple Regression
   B. Chi-square test of independence
   C. Pearson Correlation
   D. ANOVA

8. What should we use if we want to find the influence of a third variable on a relationship between two variables?

   A. Moderation
   B. Either mediation or moderation
   C. Neither mediation nor moderation
   D. Mediation

9. Which of the following is not an assumption for Chi Square?

   A. Homogeneity of variance
   B. Independence
   C. Expected frequency 5+
   D. Appropriated measurement

10. A correlational study between the number of shark attacks and the amount of ice cream purchased found an r of .361 and an R^2 of 0.10. What does this mean?

    A. There is no correlation between shark attacks and ice cream purchases.
    B. Ice cream causes more shark attacks
    C. There is a negative correlation between shark attacks and ice cream purchases.
    D. There is a positive correlation between shark attacks and ice cream purchases.

11. We want to know how the number of shark attacks and the amount of ice cream purchased are related. What approach could we use here?

    A. Chi Square Test of Indpendence
    B. Multiple Regression
    C. Pearson Correlation
    D. Not enough information

12. You decide to look at exam scores for students (with permission, of course). You found a positive association between note-taking and high performance on the exam. This association may be explained by hours studying. What type of variable is "hours studying"?

    A. Outcome
    B. A bad one
    C. Confounder
    D. Predictor

13. When using Pearson correlation, what scales of measurements should be used?

    A. Interval, Ratio, and Nominal
    B. Only Interval
    C. Only Ratio
    D. Interval and Ratio

14. See Output: What approach could we use?

    A. Chi Square Test of Independence
    B. Correlation
    C. ANOVA
    D. T-Test

15. See Output: What approach was used?

    A. Correlation
    B. Chi Square Test of Independence
    C. T-Test
    D. One-way ANOVA

16. See Output: Assuming the dependent variable is income and independent variable is biological sex, is the result significant and what does that mean?

    A. Not significant; there is not a difference between the sexes in their income
    B. Significant; there is a difference between the sexes in their income
    C. Not significant; there is a difference between the sexes in their income
    D. Significant;there is not a difference between the sexes in their income

17. See Output: What approach was used?

    A. T-Test
    B. Multiple Regression
    C. One-way ANOVA
    D. Chi Square Test of Independence

18. See Output: With the dependent variable being income and the independent variable being academic degree earned, is the result significant and what does that mean?

    A. Not significant; There is not a difference in income between the degrees earned
    B. Significant; There is a difference in income between the degrees earned
    C. Significant; There is not a difference in income between the degrees earned
    D. Not significant; There is a difference in income between the degrees earned

19. What are the steps of hypothesis testing that we've used in this class?

    A. 1. Examine Variables, 2. State Null/Research Hypotheses, 3. Define Critical Regions, 4. Compute Test Statistic, 5. Compute Effect Size, 6. Interpret Results
    B. 1. State Null/Research Hypotheses, 2. Define Critical Regions, 3. Compute Test Statistic, 4. Interpret Results
    C. 1. Examine Variables, 2. State Assumptions, 3. Calculate Effect Size, 4. Compute Test Statistic, 5. Interpret Results
    D. 1. Examine Variables, 2. State Assumptions, 3. Define Critical Regions, 4. Calculate Effect Size, 5. Compute Test Statistic, 6. Interpret Results

20. The independence assumption of linear regression means:

    A. That the individuals in the data are randomly selected
    B. That the individuals in the data are paying attention during the survey
    C. That the individuals in the data do not influence each other
    D. That the individuals in the data are not cheating

21. Rumor is that getting vaccinated can somehow cause problems in focus but that a dab of baking soda can suck out all the toxins of vaccines with the benefit of the immunization (obviously I just made this up… but it is similar to ridiculous rumors I've actually heard). You conduct a study by randomly assigning participants into two groups. One group is given regular vaccines while another has the baking soda added. What approach could we use here to test for differences among the two groups?

    A. Paired samples t-test
    B. Independent sample t-test
    C. Chi Square Goodness of Fit
    D. One-way ANOVA

22. You have a continuous outcome, a continuous predictor, and covariates. What approach could you use here?

    A. One-way ANOVA
    B. Paired samples t-test
    C. Chi Square Goodness of Fit
    D. Multiple Regression

23. We want to know the way chocolate consumption affects voting behavior. We send out a survey that asks how much chocolate they consume and how often they vote. We run a simple regression model that shows that for every additional ounce of chocolate consumed, voting goes up by .5 units. Is this relationship causal?

    A. Yes, the difference shows that more chocolate, more voting. To be more certain, we could always control other variables.
    B. Yes, but I'm not sure how
    C. No, but I'm not sure why.
    D. No, there could be confounders since we did not control for anything and did not use random assignment

24. Which of the following coefficients from a regression would make sense with the following situation: The predictor significantly positively predicts the outcome.

    A. 0
    B. None of the above
    C. 2.1
    D. -2.1

25. Which of the following is NOT an the assumption of linear regression?

    A. Independence of Data
    B. All variables that predict the outcome should be in the model
    C. Homoskedasticity
    D. Linear relationships

26. What is the intercept in a regression model?

    A. The average of the predicted values.
    B. The average of the outcome when all predictors are zero
    C. The average of your favoriate movies
    D. The average of the predictors when the outcome is zero

27. What does "controlling for a variable" mean in regression?

    A. Statistically we make everyone equal on the controlled for variables
    B. Statistically we make minus the controlled for variables from the outcome
    C. The controlled for variables are erased from the data
    D. We assume that those people are the same on those controlled for variables

28. If two variables are collinear, that means:

    A. They are related in some linear fashion
    B. They measure the same thing
    C. They interact with each on on their effect on the outcome
    D. They are independent

29. How would you interpret the confidence interval of: the mean is 1.5 (95% CI: 0.5, 2.5)?

    A. This is a toughy. Give me some and I'll get back to you.
    B. There is a 95% chance the population mean is between 0.5 and 2.5.
    C. 95% of the population means will be between 0.5 and 2.5
    D. We are 95% confident that the population mean is between 0.5 and 2.5.

30. See Output: What approach could you use with the data shown in the figure?

    A. Independent sample t-test
    B. Chi Square Goodness of Fit
    C. Chi Square Test of Independence
    D. Pearson's Correlation

Interpretation (25 pts)

Let's say you are interested in understanding how watching The Office influence individuals' views of regarding awkward social interactions. You ask individuals how much they've watched The Office and present them with awkward social interactions (via video) to have them rate the awkwardness. You go through each step of the hypothesis test, you see no problems with the assumptions and have decided on an alpha of .05 for concluding significance. You think there may be an interaction between hours watched and whether they are a boss at a company. You also decide to control for age in the analysis. You assess the data and find the following results.

- What analysis did you use?
- What does the R^2 mean here?
- Is the interaction significant? What does the interaction mean in general terms?
- Using the plot, what is the effect of watching the office on awkwardness?

Model Fit Measures

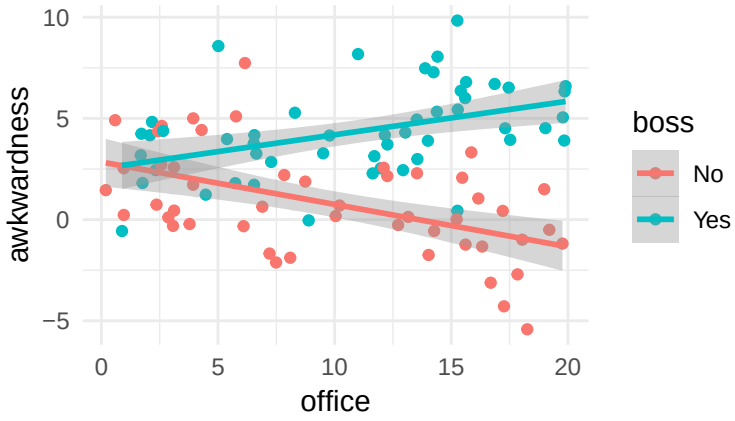
Model	R	R^2
1	0.798	0.636

MODEL SPECIFIC RESULTS

MODEL 1

Model Coefficients

Predictor	Estimate	SE	t	p
Intercept	-0.4304	0.7382	-0.583	0.561
office	-0.2182	0.0424	-5.145	< .001
age	0.0755	0.0128	5.910	< .001
boss	-0.1271	0.7432	-0.171	0.865
office:boss	0.3761	0.0620	6.063	< .001



Output Page

Question 14

Contingency Tables

degree	Female	Male	Total
<HS	111	85	196
Associates	63	53	116
Bachelors	105	109	214
Graduate	60	41	101
HS	344	301	645
Total	683	589	1272

Question 15

		statistic	df	p
income06	Student's t	-3.81	1112	< .001

Question 16

		statistic	df	p
income06	Student's t	-3.81	1112	< .001

Question 17

	Sum of Squares	df	Mean Square	F	p
degree	3.74e+11	4	9.36e+10	61.1	< .001
Residuals	1.70e+12	1109	1.53e +9		

Question 18

	Sum of Squares	df	Mean Square	F	p
degree	3.74e+11	4	9.36e+10	61.1	< .001
Residuals	1.70e+12	1109	1.53e +9		

Question 30