

## 1 Part I: Pre-midterm Material

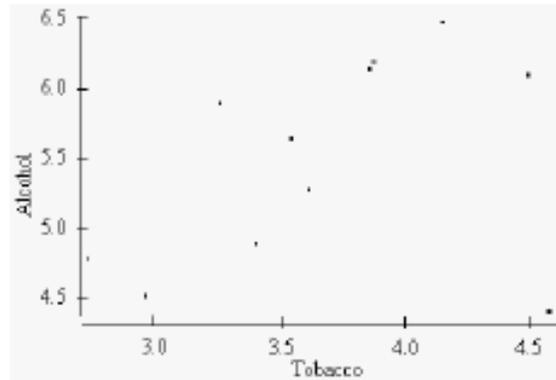
There will be a section of the final exam with questions on the material discussed before the chapter on correlation (Argumentation, Observation). The section will be worth 15 points and likely consist of multiple choice questions and one one-paragraph question.

## 2 Part II: Correlation

This part consists of multiple choice questions worth 20 points and two one-paragraph questions worth 15 points.

### 2.1 Answer the following multiple-choice questions (20 points)

1. To make a prediction for a new value of the predictor variable based on a correlation between two score values, you
  - (a) multiply the Pearson correlation coefficient by the new predictor value and add the value of the y-intercept.
  - (b) multiply the new predictor value by the slope of the regression line and add the value of the y-intercept.
  - (c) divide the new predictor value by the slope of the correlation line and add the value of the y-intercept.
  - (d) add the new predictor value to the slope of the correlation line and divide by the value of the y-intercept.
  
2. For SAT scores to be a legitimate factor in determining college admissions,
  - (a) colleges must align their curriculum with the questions used on the SAT
  - (b) SAT scores must be a direct (proximal) cause of success in college
  - (c) SAT scores must be a cause (direct or indirect) of success in college
  - (d) there must be a correlation between SAT scores and success in college
  
3. The diagram on top of the next page is an example of a
  - (a) histogram illustrating a lack of correlation between tobacco and alcohol
  - (b) scatterplot illustrating a perfect correlation between tobacco and alcohol
  - (c) scatterplot illustrating a positive correlation between tobacco and alcohol
  - (d) histogram illustrating a positive correlation between tobacco and alcohol
  
4. Which of the following choices best explains the concept of an ‘operational definition’?
  - (a) It supplies sufficient conditions for a variable



- (b) It provides necessary conditions for a variable
  - (c) It provides necessary and sufficient conditions for a variable
  - (d) It relates variables used in an hypothesis to measurable variables
5. Using the number of items a person can recall has high construct validity as a measure of memory if it
- (a) does not involve an operational definition of memory
  - (b) is only employed in a limited range of settings
  - (c) reliably assigns the same scores to the same individuals on different occasions
  - (d) actually measures memory
6. You have found a correlation of 0.42 between two variables in a sample of people. Why do you need to determine whether the correlation is statistically significant?
- (a) there could be a correlation in the sample but none in the actual population
  - (b) there could be a correlation in the actual population but none in the sample
  - (c) although there is a correlation, it might not be very important
  - (d) there may be very different reasons for the correlation in the actual population than in the sample
7. Which of the following makes no sense?
- (a)  $p < 0.10$
  - (b)  $r = 0.5$
  - (c)  $p = -0.05$
  - (d)  $r = -0.95$
8. Suppose that a study claimed  $p < 0.001$ . The probability of rejecting the null hypothesis when it is in fact true
- (a) cannot be determined from the information given
  - (b) depends upon whether there was a proper control group
  - (c) greater than 0.999

- (d) less than 0.001
9. In doing a test of statistical significance, you will have made a Type I error if you conclude that
- (a) there is no difference between means in the population when in fact there is a difference
  - (b) there is a difference between means in the population when in fact there is no difference
  - (c) the absence of a difference in the samples is due to chance
  - (d) any difference in the samples is due to chance
10. A test preparation company offers a money back guarantee that those whose verbal score was less than 250 on a previous SAT exam (which is scored between 200 and 800) will improve by at least 10 points after taking their course. Such a company
- (a) is likely to lose a lot of money since most students probably cannot improve that much
  - (b) can avoid losing a lot of money only by being very effective in teaching verbal skills
  - (c) is not likely to have to make many refunds due to the hot hand phenomenon
  - (d) is not likely to have to make many refunds due to the phenomenon of regression to the mean
11. Which of the following expressions makes no sense?
- (a) measured independent variable
  - (b) manipulated independent variable
  - (c) measured dependent variable
  - (d) manipulated dependent variable
12. How could there be a large difference between the means of two sample populations and it still be wrong to conclude that there is a difference in the means of the two actual populations
- (a) there were errors in the measurement of the difference in the actual population
  - (b) the difference between the sample populations was of the size that could have arisen by chance
  - (c) the difference between the sample populations was not statistically significant
  - (d) the nominal variable in terms of which the sample populations was divided was not operationally defined

## 2.2 Answer the following questions in a paragraph (15 points)

1. Explain why establishing correlation does not suffice to demonstrate causation. What are some ways in which variables  $A$  and  $B$  may be correlated without  $A$  causing  $B$ ?
2. Explain why, in a purely correlational study, one does not need to distinguish between independent and dependent variables.

### 3 Part III: Causation

This part consists of multiple choice questions worth 20 points and two one-paragraph questions worth 15 points.

#### 3.1 Answer the following multiple-choice questions (20 points)

1. Mill's methods allow us to:
  - (a) Identify actual effects from among possible effects
  - (b) Identify actual causes from among possible causes
  - (c) Avoid performing costly experiments
  - (d) Determine what variables are correlated with others
2. A \_\_\_\_\_ cause is a cause near the beginning of a chain of causation
  - (a) Proximate
  - (b) Ultimate
  - (c) Partial
  - (d) Necessary
3. Which of the following is true of a necessary cause:
  - (a) If they fail to occur the effect will not occur
  - (b) If they occur then the effect will also occur
  - (c) Every effect has only one necessary condition
  - (d) If they fail to occur the effect might still occur
4. You have read that owning a cell phone is a sufficient cause of getting dates. What would show you that this is incorrect?
  - (a) getting dates but not owning a cell phone
  - (b) not owning a cell phone and not getting dates
  - (c) owning a cell phone and not getting dates
  - (d) getting a fancy car and getting dates
5. In order to determine what caused some of her puppies to die and others to live, Ali decided to employ one of Mill's methods. She constructed the table below and from the results concluded it was leaving the puppies with her brother. Which of Mill's methods was Ali employing?

had enough food	had the proper shots	stayed with her brother	puppies died
Yes	No	Yes	Yes
No	Yes	Yes	Yes
Yes	No	No	No
Yes	No	No	No

- (a) agreement

- (b) difference
- (c) residues
- (d) concomitant variation

6. In the causal diagram



- (a) driving intoxicated is a common cause of the accident and dying
  - (b) the accident is a causal intermediate between driving intoxicated and dying
  - (c) the accident is a common cause of driving intoxicated and dying
  - (d) driving intoxicated is the proximate cause of dying
7. You hypothesize that an observed association between heroin use (A) and AIDS (C) is due to the effects of a causal intermediary, the HIV virus (B). Assume that you can measure all three (A, B, and C) and can manipulate A and B. Which of the following would count as a test of your hypothesis that B is the causal intermediary?
- (a) block A and check whether A is still correlated with B
  - (b) block B and check whether B is still correlated with A
  - (c) block A and check whether B is still correlated with C
  - (d) block B and check whether A is still correlated with C
8. If there are differences in the way the subjects in the experimental and control group are treated, other than the manipulation of the independent variable being investigated, these are considered
- (a) subject variable confounds
  - (b) procedural variable confounds
  - (c) dependent variable confounds
  - (d) independent variable confounds
9. Confounding variables are a threat to the
- (a) the logical validity of a study
  - (b) external validity of a study
  - (c) internal validity of a study
  - (d) statistical significance of the result
10. Assume that in a laboratory study saccharin has been demonstrated to cause cancer in rats. Whether or not it causes cancer in humans is a question of:
- (a) what the  $p$ -value of the correlation is
  - (b) internal validity
  - (c) external validity
  - (d) the strength of the correlation

11. A within-subjects experimental design
  - (a) eliminates any concern with procedural confounds
  - (b) increases the risk of carry-over effects between conditions
  - (c) increases the risk of experimenter bias
  - (d) requires many more subjects than a between subjects design
  
12. When a researcher identifies possible subject confounds while doing a prospective study of differences due to shoe size, a good strategy is to
  - (a) randomize the assignment of subjects so as to break any correlation between them and the independent variable
  - (b) screen off the confounds so that they do not influence the dependent variable
  - (c) match subjects in the different groups in terms of these variables
  - (d) lock the confounding variables so that they cannot affect the outcome
  
13. Whenever on the morning of a game I put on my lucky socks, the Chargers win that afternoon. This is an instance of
  - (a) the genetic fallacy
  - (b) ignoring a common cause.
  - (c) post hoc, ergo propter hoc
  - (d) confusing cause and effect

### **3.2 Answer the following questions in a paragraph (15 points)**

1. Explain what is wrong with the following inference as a way to confirm a causal claim and explain what inference strategy should be used instead to confirm the causal claim. Explain in clear prose what is different about the relation between premise and conclusion advanced in the two cases:

If watering my plants causes them to grow, then there will be a statistically significant difference in resulting growth when I water some and don't water others. I watered some of my plants and not others and there was a statistically significant difference in their growth. So I conclude that watering my plants causes them to grow.

2. Many psychological studies performed in the U.S. use samples composed entirely of undergraduates in psychology classes. Discuss the implications of this in terms of the external validity of the experiments conducted. Under what circumstances will the experiments remain externally valid and when will external validity be compromised?

## 4 Part IV: Paradoxes of confirmation

This section will be worth 15 points and only include material covered in the last lecture and the slides that go with it. It will likely consist of multiple choice question and a one-paragraph question.

1. What is the core of Reichenbach's 'pragmatic vindication' of Hume's problem of induction?
  - (a) A replacement of our inductive practices with a purely deductive methodology.
  - (b) A decision-theoretic argument that concludes that we have nothing to lose but plenty to gain from applying induction.
  - (c) A psychological explanation of how repeated similar experiences create a propensity to expect a similar train of experience in the future.
  - (d) A proposal that we award a privileged status to 'projectable' predicates.
  
2. What is the core of the curve-fitting problem?
  - (a) That it shows that scientists only use projectable predicates in forming their functional hypotheses.
  - (b) That functional hypotheses are underdetermined by the data.
  - (c) That there is always an infinity of mutually contradictory functional hypotheses.
  - (d) 1, 2, and 3.
  - (e) Only 2 and 3.