

Instructions: This exam is in two parts: Part I is to be completed partly at home using the materials posted on Blackboard for Part I and you will answer questions about that work in class below; Part II is to be completed entirely in class. You may not use cell phones, and you may only access internet resources you are specifically directed to use. You may access your data file for Part I of the exam in Blackboard. You may access the data files posted to Blackboard for the Exam part II. Be sure you are using the data file that matches the exam version you are given.

Part I: At Home

This part was completed at home. You can upload the Excel file for Part I to the Part I folder in Blackboard for use during the Exam period. However, this submission will not be graded in this location.

Part II: In Class

1. Use the work done at home to answer the Part I questions.
2. Open the file from the in-class portion of the final posted on Blackboard that corresponds to the version of the exam you have. This is Exam B.
3. Answer the questions corresponding to the data file, and any additional calculation in Excel required.
4. When you have finished answering questions on the exam, and all your answers have been recorded on the paper test for grading, upload both the take home Excel file and the in-class Excel file to the same in-class Exam folder in Blackboard for grading. Only those files submitted to the correct folder will be graded.
5. Turn in your paper copy of the exam to your instructor.
6. Enjoy your break!

Part I:

The following question refers to #1 from Part 1:

1. Describe what you see in the comparative box plots. Are they about the same? Do any seem dramatically different than the others? Is the spread about the same? (10 points)

The follow question refers to #2 from Part 1:

2. Consider the scatterplot of time vs. percent of homeownership. Does the relationship appear to be linear or nonlinear? If it's nonlinear, describe the relationship. Which state did you choose? (8 points)

The following questions refer to problem #3 from Part I:

3. Consider the pivot table you created. Which position appears to be paid the most on average in 2009? (8 points)
4. Does the relationship between position and year appear to be strong? Explain. (8 points)

The following questions refer to problem #4 from Part I:

5. Consider your scatterplot of Children vs. Money. Does the relationship appear to be linear or nonlinear? Explain. [Note: Children is a discrete variable. Do not confuse that with a nonlinear relationship!] (8 points)

6. State the linear trendline equation and its R^2 value for the relationship. (8 points)

7. State the correlation r value. Is the relationship positive or negative? Is the relationship strong, moderate or weak? (8 points)

8. Use the equation you found to predict the average amount of money spent if the family has five children, if the trend continues. Does this seem reasonable? (8 points)

9. Describe the meaning of the R^2 value in context. (6 points)

The following question refers to problem #5 from Part I:

10. For the outcome 3, state the proportion of your simulation that produced that result. Compare it to the true probability. (6 points)

Calculations in Excel: (1) 15 points, (2) 15 points, (3) 15 points, (4) 30 points, (5) 25 points.

Part II:

11. Create a Pivot Table from the data in the Excel file for #11 comparing Type and Preference. Use it to answer the following questions about a randomly selected person from the dataset:
- What is the probability the person is black (B)? (6 points)
 - What is the probability that the person prefers left (L)? (6 points)
 - What is the probability that the person is black given that they prefer left? (6 points)

d. What is the probability that the person both is black and they prefer left? (6 points)

e. What is the probability that the person either is black or they prefer left? (6 points)

f. What is the probability that someone do not prefer left? (6 points)

g. Are the variables Type and Preference independent? Why or why not? (10 points)

h. Create a stacked column graph of the data. What do you notice overall? Did you choose a count stacked or a percent stacked chart? (10 points)

12. Explain the difference between a personal (subjective) probability and an experimental (observational) probability. (8 points)

13. A probability distribution is provided in the Excel file. Calculate the following:

a. Find the expected value of the probability distribution. (6 points)

b. Calculate the variance of the probability distribution. (8 points)

c. What is the standard deviation of the distribution? (4 points)

14. A particular model of lie detector test has a 97% probability of correctly detecting someone who is lying, and 99.1% probability of correctly detecting someone who is not lying. A human police officer interviews suspects using the lie detector, hoping to catch the 1 of the 200 suspects worth investigating further who they believe to be lying. If someone tests positively on the lie detector for lying, what is the probability that the person is actually lying? Construct a tree diagram to model the situation. Should the police continue to use the lie detector, and if so, is it safe to assume the person is definitely guilty? (15 points)

Upload your completed Excel files (**both of them!**) to the Exam #2 **to be graded** submission box in Blackboard and submit your completed paper exam to your instructor. You may not modify anything once the exam is submitted.

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

$$P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$$

$$\mu = E(X) = \sum x_i p(x_i)$$

$$\sigma^2 = Var(X) = \sum (x_i - \mu)^2 p(x_i)$$