BUS 210, Exam #2A, Fall 2018

Name KEY

Section _____

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:

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 range of silver is much smaller while the others are more similar, all appear skewed right. palladium and platinum both have multiple outliers. The largest spread is platinum.

The follow question refers to #2 from Part 1:

2. Consider the scatterplot of ~~money spent vs. number of children~~ time vs. % home ownership. Does the relationship appear to be linear or nonlinear? If it's nonlinear, describe the relationship. Which state did you choose? (8 points)

The relationship is not linear

I chose Alabama, but your state may vary

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 2007? (8 points)

Designated Hitter

4. Does the relationship between position and year appear to be strong? Explain. (8 points)

the relationship does appear to be strong over time biggest difference from year to year is 1st base

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)

linear - the variable is discrete but it is not strongly nonlinear as there is no clear curve

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

$$y = -203.27x + 1406.6$$

$$R^2 = 0.0494$$

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

$$r = -0.2223$$

weak

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

$593.52

So far. ceases to be reasonable at 7 kids

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

only 4.9% of the variability in amount spent can be explained by the # of children

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

10. For the outcome 7, state the proportion of your simulation that produced that result. Compare it to the true probability. (6 points)

0.136 (answers will vary)
fairly close to predicted 0.125 = 1/8

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 Job Status and Number of Children. The cell values should represent a count, not a sum. Use it to answer the following questions about a randomly selected person from the dataset:
    a. What is the probability the person has a full-time job (F)? (6 points)

307/750 = 0.409

    b. What is the probability that the person has 1 child? (6 points)

271/750 = 0.361

c. What is the probability that the person lives in a has a full-time job given that they have one child? (6 points)

$126/271 = 0.465$

d. What is the probability that the person both has a full-time job and has one child? (6 points)

$126/750 = 0.168$

e. What is the probability that the person either has a full-time job or has one child? (6 points)

$(271 + 307 - 126)/750 = 0.603$

f. What is the probability that someone does not have a full-time job? (6 points)

$1 - 0.409 = 0.591$

g. Are the variables Job Status and Number of Children independent? Why or why not? (10 points)

$P(A|B) = 0.465 \neq P(A) = 0.409$ dependent

$P(A \text{ and } B) = 0.168 \quad P(A) * P(B) = 0.148$

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)

percent stacked — non-working people more likely to have 2 kids

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

Classical probability calculated from general ideas about the world, but experimental probability based on looking at many repeated trials and estimating probability from proportions

13. A probability distribution is provided in the Excel file. Calculate the following:
    a. Find the expected value of the probability distribution. (6 points)

E(X)=3.02

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

4.066

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

2.016

14. A particular model of lie detector test has a 98% probability of correctly detecting someone who is lying, and 99.5% 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 20 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)

Not Lying Given Lie Detected 0.088
it's worth being strongly suspicious but not definitive (wrong 1 time in 12)

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.

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

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

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

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