STAT303 Sec 502 and 503  
Spring 2015  
Exam #4  
Form A  

Instructor: Julie Hagen Carroll  
May 7, 2015

Some questions have been adapted from *Statistics: The Art and Science of Learning from Data*, 3rd ed, by Agresti & Franklin.

1. **Don’t even open this until you are told to do so.**

2. Remember to put your phone on AIRPLANE mode, not vibrate – it’ll still make noise! You may listen to your tunes as long as you do not disturb anyone.

3. Please turn your hats around backwards or take them off.

4. Please put your backpack and other things along the walls or at the front of the room.

5. You need a gray, 8½ × 11” scantron, pencil, calculator and you may have 5 sheets of notes.

6. There are 20 multiple-choice questions on this exam, each worth 5 points. There is partial credit. Please mark your answers clearly. Multiple marks will be counted wrong.

7. You will have 60 minutes to finish this exam.

8. If you have questions, please write out what you are thinking so that we can discuss it after I return it to you.

9. If you are caught cheating or helping someone to cheat on this exam, you both will receive a grade of zero on the exam. You must work alone. Do not discuss anything about the exam with the other section until grades are posted.

10. When you are finished please make sure you have filled in your name and marked your FORM (A or B) and 20 answers, and leave everything on your desk. Try to sit on the same area next time so that I can return your materials to you.

11. Good luck!
1. If our data is quantitative but separated into different groups, what type of test should we run?

A. an ANOVA test on means for the different groups  
B. a regression for each group since the data is quantitative  
C. one sample t-tests for each group  
D. $\chi^2$ test since the groups are categories, i.e., qualitative data  
E. We haven’t discussed how to analyze this since it’s a combination of both types of data.

2. The data above represent the number of calories and grams of sugar per serving for a random sample of energy bars. The first set is: $Sugar = 17.87 - 0.0146 \times Calories$ and the second: $Calories = 222.3 - 0.3999 \times Sugar$. Although the correlation, $r = -0.076$, is the same for both, the slope and intercept are different. Why?

A. Randomization allows for variability in the estimates.  
B. There is no relationship between the two equations because they are predicting different variables.  
C. The errors are minimized in the $y$-direction, so it matters which is $x$ and $y$.  
D. The correlation shouldn’t be the same since the slopes are different and the slope and correlation are related.  
E. It doesn’t matter because the P-values are the same.

3. Which of the following is the best conclusion for the previous output?

A. Since the correlation is weak, we shouldn’t use calories to predict sugar or vice versa.  
B. The P-value is large for both, so there is no relationship between calories and sugar.  
C. You shouldn’t just look at the calories if you’re counting carbs (sugar is a carbohydrate).  
D. All of the above are good conclusions.  
E. Two of the above are valid conclusions.

4. Why do all the test statistics (except the $\chi^2$) divide by the standard error or variance?

A. Variability affects the estimates so it must be included in the calculations.  
B. Differences in sample and hypothesized values are only significant relative to their variability.  
C. Differences in samples (due to randomization) must be included in the estimates.  
D. Variance is necessary to prove significance.  
E. Variability is a part of life.

5. Your family is in the market for a new vehicle. One basic criteria is safety. The Head Injury Criterion (HIC) is a measure of the likelihood of head injury arising from an impact. The data above is based on 7 models in three categories: large family car, passenger vans, and midsize SUV. Although the P-value is missing from the ANOVA, knowing what you know about what an F statistic represents, what should you conclude?

A. 0.399 is larger than even 10%, so we would not be able to say the means are all the same.  
B. 0.399 is larger than even 10%, so we would not be able to say the means are not all the same.  
C. 0.399 is less than 1, so the P-value is large, so we could not be able to say the type of vehicle affects the HIC.  
D. 0.399 is less than 1, so the P-value is large, so we could say the type of vehicle affects the HIC.  
E. Vans are for soccer moms only!

6. What are the missing values in the previous ANOVA table?

A. Vehicle df = 3, Error df = 18, MSE = 548011  
B. Vehicle df = 2, Error df = 20, MSE = 548011  
C. Vehicle df = 2, Error df = 20, MSE = 30445  
D. Vehicle df = 2, Error df = 18, MSE = 30445  
E. Vehicle df = 3, Error df = 20, MSE = 30445

7. Which of the following is the best interpretation of power in the previous F test?

A. the ability to detect a difference in HIC for the three types of vehicles (reject $H_0$) because one exists  
B. the ability to detect no difference in HIC for the three types of vehicles (reject $H_0$) because there isn’t one  
C. the ability to prove there is an effect even though one doesn’t exist  
D. the ability to prove there is not an effect because one doesn’t exist  
E. the ability to prove there is not an effect even though there is one
8. The graph above represents the same data. What should you conclude?

A. The means (medians) are not all equal.
B. The variances are so large that the difference in means (medians) is insignificant.
C. The F statistic should be large since the variances are large.
D. The F statistic is not valid since the variances are so large.
E. The must be some lurking variable that makes the variances so large.

9. Which of the following is true about the $\chi^2$ test of independence of two categorical variables?

A. Both variables must have at least 30 observations.
B. The test compares whether joint probabilities are equal to the product of the two (appropriate) conditional probabilities.
C. The greater the dependency (more related) the two variables, the smaller the test statistic value.
D. Two of the above are true.
E. None of the above are true.

10. What is the correct range of the P-value for testing whether $x$ is a good predictor for $y$ or not given the following confidence intervals for the true slope?

90% CI: (0.7143, 2.674)
95% CI: (0.4896, 2.899)
99% CI: (-0.0192, 3.408)

A. P-value > 0.10
B. 0.10 > P-value > 0.05
C. 0.05 > P-value > 0.01
D. P-value < 0.01
E. The hypothesized value is not given.

11. Since weighing bears is rather risky, is it possible to estimate their weight with their length (assuming you can get their length through some safe method)? The output above is from 12 black bears, the most common bear on the planet. Based on this data

A. we can conclude that weighing bears is not necessary since length is a good predictor.
B. we can conclude at the 1% level that length is not a good predictor for weight.
C. we can conclude at the 5% level that length is a better predictor than weight.
D. All of the above are valid.
E. Two of the above are valid.

12. The P-value for the intercept is missing (not that we use it). What is the correct range for it?

A. It’s 0.0106 like given in the ANOVA table.
B. It’s not possible to have a negative $t$ value.
C. 0.02 > P-value > 0.01
D. 0.10 > P-value > 0.05
E. 0.05 > P-value > 0.025

13. What would be the consequence of a Type II error?

A. trying to weigh black bears (and maybe getting eaten) when you should have just used their length
B. using the length black bears and misrepresenting their actual weights
C. wasting time and money looking for other ways to find the weight of black bears when their length was sufficient
D. using the length of black bears for all types of bears
E. Two of the above are plausible Type II errors.
14. Which of the following is the best interpretation of the previous P-value = 0.0106?

A. 10.6% of the time we will get at least this good of a predictor when there really isn’t any relationship between x and y.
B. 1% of the time we won’t be able to predict y even though there is a strong relationship between x and y.
C. A linear relationship between x and y that’s at least this strong will happen only about 1% of the time if x and y are independent.
D. 1% of the linear relationships between x and y will be at least this strong.
E. You should avoid black bears 99% of the time.

15. The plot above is based on data an obstetrician gathered on births from eight randomly chosen days for each week day. If you were to run an ANOVA on this data, what result should you expect?

A. It’s doubtful that the true variances are equal, so an F test is not appropriate.
B. An F test is not the correct procedure since it’s obvious that Monday different from the other days.
C. The P-value would be very small because it’s obvious that Monday different from the other days.
D. The number of births and days of the week have a strong relationship.
E. Looks like Mondays are slow days even in obstetrics!

16. To improve the power of an ANOVA F test we should

A. increase the number of groups (means)
B. increase the size of each group
C. reduce the variability between the groups (means)
D. All of the above will increase the power.
E. Only two of the above would increase the power.

17. Do males and females drop courses for different reasons? The table above shows reasons for dropping: didn’t like/need the class, personal issues, had to work more/change schedule by gender. What conclusion(s) are valid for this data?

A. Yes, we can conclude at the 10% level that their reasons are different, there is some relationship between reason and gender.
B. At the 1 and 5% levels, we can say that there is no difference in reasons for dropping by gender.
C. Since not all of the cells have at least 5, the χ² test is invalid.
D. Two of the above are correct.
E. None of the above are correct.

18. Which of the following is a Type I error for the previous χ² test?

A. claiming males and females are different when they are really the same
B. claiming males and females are independent when they are really related
C. claiming males and females are related when they are really independent
D. claiming the reason for dropping isn’t important to either males or females when it actually is
E. None of the above are correct.

19. What is the expected number of males who dropped a course for work reasons if gender and reasons are independent?

A. 13
B. 9
C. 16
D. 29
E. 4