BIO 206 Spring 2021

Assignment 14 Ch 13 “Handling violations of assumptions”

**Due Wed, May 19**

(77 pt)

* Please type your final answers to the problems on the main pages of your assignment. After the pages with typed answers, attach handwritten pages showing your calculations. On the pages of calculations, clearly indicate the problem numbers and final answers. For calculations using a contingency analysis, include a table of observed values, your calculations for expected values, and a table of expected values, on the handwritten pages. Don't use computers for the analyses.
* Directions for reporting results of statistical tests---on the typed pages, report:

--The means of each group

--n (sample size)

--degrees of freedom

--alpha (Threshold level of significance. Decide on the value before doing the statistical test. The default is to test at alpha = 0.05, meaning at the 5% significance level.)

--the name and value of the test statistic (Example: X2 value = 0.55)

--p-value (from a table of critical values in the back of the textbook)

--Describe your conclusion in a sentence. (Example 1: “Births are not evenly distributed over the days of the week.” Example 2: “The proportion of wild-type flies was significantly greater than expected by the 3:1 ratio.” Example 3: “The proportions of people who are interested in each of the five areas of psychology are significantly different.”

* For practice problems, please write an answer in your own words before checking your answer against the solutions in the back of the book. Then, compare the solution to your answer. If your answer wasn't correct, don't change it--write about how the solution in the book differed from your answer and why it's correct. In other words, show evidence that you are thinking independently and learning. In this way, you can get full credit even if you originally didn't have the correct answer.

Questions not from the book

1 (2 pt) What is the purpose of a data transformation? (Not the definition, but why might you do it?)

2 (2 pt) When is the log transformation most likely to be useful?

3 (2 pt) When is the arcsine transformation likely to be useful?

4 (2 pt) When is the square-root transformation likely to be useful?

5 (2 pt) How can you tell if a data transformation works?

Practice problems, Chapter 13

5a (4 pt)

8 Don’t perform the test—just answer these questions:

8a (2 pt) What are the variables that you should use in a statistical test?

8b (2 pt) For each variable(s), name its type.

8c (3 pt) To answer the question "does infanticide lad to a different mean delay to reproduction in females than when cubs die from other causes", which statistical test should you choose to answer the question and why? (From any in the chapter)

11 (a-e) (15 pt) Explain each choice.

12a (4 pt)

12c (2 pt) (skip 12b)

Question not from the book: For this problem, you'll be answering the question "Did cloud seeding increase the amount of rain that fell on rainy days?" using an unpaired (two-sample) t-test. Use the data file called "Rainfall after cloud seeding" that is on D2L.

Before using a two-sample t-test, you need to check whether the data for each category meet the assumptions of the test.

Part 1: (4 pt for each histogram; total = 8 pt.) To check whether data for each category meet the assumption of normality, create two histograms of frequency distributions using the website for the Histogram Maker. Look in Content>>Handouts>>”Directions for Histogram Maker for graphing freq dist”.

One of the histograms should show the frequency distribution of rainfall on unseeded days and one should show rainfall on seeded days. Label and attach each of the histograms to your assignment.

For each graph, remember to include labels and units on axes, a figure legend describing what the graph shows, and source of data. You can write the labels by hand along the axes.

Change the settings on the histogram to look at the data in different ways.  Try changing the bin numbers and look at different views. Also try changing the lowest class value.

Part 2: (4 pt each for 2 files; total = 8 pts) Turn in an Excel file with four labeled columns (unseeded rainfall, seeded rainfall, log unseeded rainfall, log seeded rainfall). Follow the directions in the file "Directions for Histogram Maker for graphing freq dist" to create the table in Excel. Print and attach the tables to the end of your assignment.

Part 3: (4 pt for each histogram; total = 8 pt.) To check whether log transforming the data for each category made the data meet the assumption of normality, create two histograms, one showing the frequency distribution of log transformed rainfall on unseeded days and one showing log transformed rainfall on seeded days. Label and attach those two histograms to your assignment.

Part 4: (3 pt) Is log transformation necessary for the data and do the data meet the assumptions of the two-sample t-test afterwards? Explain.

Note: you have to transform both groups of data (data for unseeded days and data for seeded days) in the same way before doing a t-test. For example, if the data for unseeded days needs to be log-transformed, then you also have to log-transform the data for seeded days so that the means are comparable in a t-test. However, if you're doing a correlation, you don't have to transform both variables (X and Y) before doing the analysis.

Part 5: (8 pt) Answer this question: Did cloud seeding increase the amount of rain that fell on rainy days? Perform a two-sample t-test using Vassarstats and data with the type of transformation (none or log) that you feel is the best choice to meet the assumptions of the test. Follow the directions above for statistical tests.

In Vassarstats, copy data from unseeded days into one column and data from seeded days into the other column. Remember a two-sample t-test is called "Independent" in Vassarstats.

===================

Extra question: At the end of your paper, please write “Number of hours I spent on this homework:” and add an estimate. Don’t include time that you spent reading the chapter. This number won’t affect your grade.