Math 107 Final Exam

Prof. Adam Loy

March 16, 2015

Instructions:

- This exam is closed book and closed notes.
- No copying, cheating, collaborations, or cell phones (or similar devices) are allowed.
- You may use a scientific or graphing calculator.
- Show your work.
- Write complete and coherent answers.
- Incorrect answers unaccompanied by supporting reasoning will receive no partial credit.
- Please write as clearly and neatly as possible. If I cannot read your answers, I cannot give you any credit.
- If you want any work done on scratch paper to be graded, mention where to look (make this obvious by using big letters with a box around them) on the page with the question.

After you have completed the exam, please *reaffirm the Lawrence University Honor Code* in the space provided below, then print and sign your name.

Reaffirmation:

Name: _____

Signature: _____

Commonly used z^* values:

Confidence level	80%	90%	95%	98%	99%
z^*	1.282	1.645	1.960	2.326	2.576

Formulas:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} (x_i - \overline{x})^2}$$

$$\text{IQR} = Q_3 - Q_1$$

$$Q_1 - 1.5 \times \text{IQR}$$

$$Q_3 + 1.5 \times \text{IQR}$$

$$z = \frac{x - \mu}{\sigma}$$

$$x = z\sigma + \mu$$

$$r = \frac{1}{n-1}\sum_{i=1}^{n} \left(\frac{x_i - \overline{x}}{s_x}\right)\left(\frac{y_i - \overline{y}}{s_y}\right)$$

$$e_i = y_i - \widehat{y}_i$$

$$n \geq \left(\frac{z^*}{ME}\right)^2 p(1-p)$$

$$n \geq \left(\frac{z^* s}{ME}\right)^2$$

$$\text{statistic} \pm (\text{critical value}) \times \text{SE}$$

$$\text{test statistic} = \frac{\text{statistic} - \text{null value}}{\text{SE}}$$

Standard errors:

$$SE_{\widehat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

$$SE_{\overline{x}} = \frac{s}{\sqrt{n}}$$

$$SE_{\overline{x}_d} = \frac{s_d}{\sqrt{n_d}}$$

$$SE_{\overline{x}_1 - \overline{x}_2} = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Questions 1-7 concern the following scenario: You want to investigate a claim that women are more likely than men to dream in color. You take a random sample of men and a random sample of women (in your community) and ask whether they dream in color.

1. (2 pts) If the difference in the proportions (who dream in color) between the two groups turns out not to be statistically significant, which of the following is the best conclusion to draw?

   A. You have found strong evidence that there is no difference between the groups.

   B. You have not found enough evidence to conclude that there is a difference between the groups.

   C. Because the result is not significant, the study does not support any conclusion.

2. (2 pts) If the difference in the proportions (who dream in color) between the two groups does turn out to be statistically significant, which of the following is a valid conclusion?

   A. It would not be surprising to obtain the observed sample results if there is really no difference between men and women.

   B. It would be very surprising to obtain the observed sample results if there is really no difference between men and women.

   C. It would be very surprising to obtain the observed sample results if there is really a difference between men and women.

3. (2 pts) Suppose that the difference between the sample groups turns out not to be significant, even though your review of the research suggested that there really is a difference between men and women. Which conclusion is most reasonable?

   A. Something went wrong with the analysis

   B. There must not be a difference after all.

   C. The sample size might have been too small.

4. (2 pts) If the difference in the proportions (who dream in color) between the two groups does turn out to be statistically significant, which of the following is a *possible* explanation for this result?

   A. Men and women do not differ on this issue but there is a small chance that random sampling alone led to the difference we observed between the two groups.

   B. Men and women differ on this issue.

   C. Either A or B are possible explanations for this result.

5. (2 pts) Reconsider the previous question. Now think about *plausible* explanations rather than possible explanations. Which is the more plausible explanation for the result?

   A. Men and women do not differ on this issue but there is a small chance that random sampling alone led to the difference we observed between the two groups.

   B. Men and women differ on this issue.

   C. They are equally plausible explanations.

6. (2 pts) Suppose that two different studies are conducted on this issue. Study A finds that 40 of 100 women sampled dream in color, compared to 20 of 100 men. Study B finds that 35 of 100 women dream in color, compared to 25 of 100 men. Which study provides stronger evidence that there is a difference between men and women on this issue?

   A. Study A

   B. Study B

   C. The strength of evidence would be similar for these two studies

7. (2 pts) Suppose that two more studies are conducted on this issue. Both studies find that 30% of women sampled dream in color, compared to 20% of men. But Study C consists of 100 people of each sex, while Study D consists of 40 people of each sex. Which study provides stronger evidence that there is a difference between men and women on this issue?

   A. Study C

   B. Study D

   C. The strength of evidence would be similar for these two studies

8. (3 pts) Briefly explain what the standard error represents.

9. (3 pts) Explain what it means to be 90% confident in a confidence interval.

10. (5 pts) An automobile manufacturer would like to know what proportion of its customers are dissatisfied with the service received from their local dealer. The customer relations department will survey a random sample of customers and compute a 90% confidence interval for the proportion that are dissatisfied. From past studies, they believe that this proportion will be about 0.15. Find the sample size needed if the margin of error of the confidence interval is to be no more than 0.02.

11. In a survey of 1430 randomly sampled undergraduate students, 76% of the students reported that they had one or more credit cards.

(a) (2 pts) Identify the population of interest.

(b) (2 pts) Identify the parameter of interest.

(c) (3 pts) Can we use the normal distribution to describe the sampling distribution of $\widehat{p}$? To answer this question check all of the conditions that must be met for the use of the normal distribution.

(d) (4 pts) Assuming that all necessary conditions are met, calculate a 95% confidence interval for the proportion of all college students who have at least one credit card.

(e) (4 pts) Interpret your 95% confidence interval in the context of the problem.

12. A company with a large fleet of cars hopes to keep gasoline costs down, and sets a goal of attaining a fleet average of at least 26 miles per gallon (mpg). To see if the goal is being met, they check the gasoline usage for 50 company trips, chosen at random, finding a mean of 25.02 mpg and a standard deviation of 4.83 mpg. Is there strong evidence that they have failed to attain their fuel economy goal?

(a) (3 pts) State the appropriate hypotheses.

(b) (3 pts) Are the necessary assumptions for inference satisfied? Briefly explain why or why not.

Regardless of your answer to (b), assume that the necessary assumptions are upheld for parts (c)–(e).

(c) (3 pts) Calculate the test statistic.

(d) (3 pts) The p-value is 0.079. Provide a sketch showing how you would find this using StatKey. Be sure to indicate what distribution you are using.

(e) (3 pts) State a conclusion within the context of the problem: At the $\alpha = 0.05$ significance level, there is

sufficient insufficient (circle one)

evidence to conclude that... (please complete this conclusion using the space provided below)

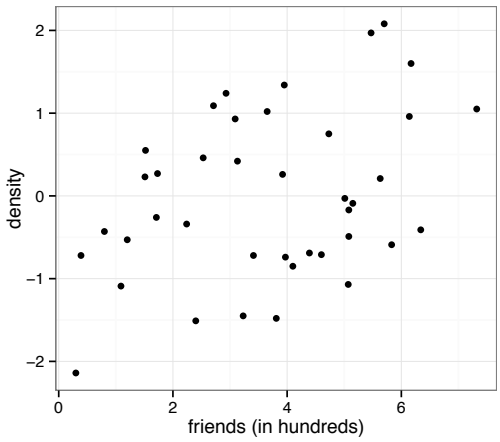
13. (21 pts) Selvi et al. (2012) conducted a study on university medical students in Turkey, to investigate whether and how sleeping habits were associated with having nightmares. During the study, these students were engaged in lecture-based learning with no hospital duties. All participants were given initial surveys, the results of which were used to determine who among them had "early bird" sleep habits and who had "night owl" sleep habits. Next, the students were given the Van Dream Anxiety Scale (VDAS) assessment, which involved questions regarding nightmare frequency and dream anxiety due to frightening dreams during the last 30 days. The summary statistics on nightmare frequency, separated by "early bird" and "night owl" status are given below. Note that higher scores indicate higher nightmare frequency.

	Sample size	Sample avg.	Sample SD
"Diagnosed" early bird	67	1.23	0.93
"Diagnosed" night owl	59	2.10	0.99

Analyze the data to determine whether and how early birds differ from night owls, with respect to average nightmare frequency. The data are not strongly skewed within either group. In your conclusion you should answer the research question and include comments on causation and generalizability.

To answer this question you will need to use R. There are two computers in the front of the room that you can use. When you are ready to use R (no calculations should be performed at the computer), please come to the front of the room and form a line. When you are done using R, make sure that you clear the history and the console.

14. Researchers used MRI's to examine areas of the brain that are involved with social interaction, memory, and emotional responses (Kanai, Bahrami, Roylance, and Rees, 2011). Their subjects were 40 university students in London. They examined five areas of the brain that have previously been linked to social perception and associative memory, and they compared each to the number of Facebook friends that the subject reported. The results from each brain area were quite similar. In this problem we will look at one set of brain density measurements focusing on the left middle temporal gyrus (MTG), which has been linked in other studies to face recognition. These results are shown below. Note that the number of friends is given in units of 100 friends, so, for example, 0.30 = 30 friends and 1.09 = 109 friends.



(a) (4 pts) Describe the association between the number of Facebook friends and brain density. Be sure to mention form, direction, strength, and outliers.

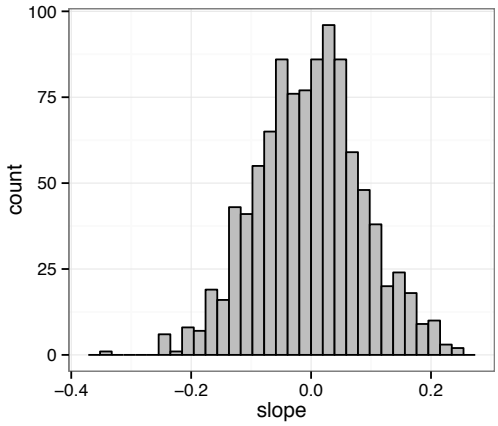
The equation for the fitted least squares regression line is $\widehat{y} = -0.74 + 0.20x$

(b) (3 pts) Interpret the y-intercept of the least squares regression line in the context of the problem.

(c) (5 pts) Interpret the slope of the least squares regression line in the context of the problem.

(d) (3 pts) What is the predicted brain density for an individual with 600 Facebook friends?

(e) The main research question in the paper is: Does brain density tend to increase as a person's number of Facebook friends increases? To answer this question, the researchers used a permutation (randomization) test to test the null hypothesis that the slope of the regression line is equal to zero. The test statistic was the estimated slope of the regression line. The permutation distribution was created by shuffling the y-values (the density) between the x-values (the number of Facebook friends), re-fitting the regression model, and extracting the slope of the regression line. A histogram of the resulting permutation distribution, assuming $H_0 : \beta_1 = 0$ is true, is given below. To the right of the figure are the largest 20 and smallest 20 observations. There are a total of 1000 values in the permutation distribution.



Smallest slopes:

-0.340, -0.253, -0.248, -0.241, -0.240,
-0.237, -0.237, -0.228, -0.212, -0.208,
-0.205, -0.205, -0.203, -0.198, -0.198,
-0.198, -0.195, -0.190, -0.188, -0.187

Largest slopes:

0.184, 0.187, 0.188, 0.190, 0.192,
0.198, 0.200, 0.201, 0.204, 0.204,
0.206, 0.206, 0.207, 0.207, 0.210,
0.218, 0.224, 0.226, 0.244, 0.247

i. (2 pts) State the appropriate alternative hypothesis.

ii. (3 pts) Calculate the p-value for this hypothesis test.

iii. (3 pts) Based on the p-value you found in part (ii), what can you conclude about the association between brain density and the number of Facebook friends a person has?