

Name: _____

Section: 06 07 08 09

Read and Sign the Following Statement:

I understand that give or receiving help on this exam is a violation of academic regulations and is punishable by a grade of **F** in this course. This includes looking at other students' exams and/or allowing other students, actively or passively, to see answers on my exam. This also includes revealing, actively or passively, any information about the exam to any member of Professor Schroth-Glanz's STAT 218 class who has not yet taken the exam. The use of cell phones is strictly prohibited. **I pledge not to do any of these things.**

Signed: _____

Instructions.

1. Read and sign the honesty pledge at the top of this page. Your exam will not be graded unless the honesty pledge is signed!!
2. Attempt all questions and write legibly.
3. Show ALL the steps of your work clearly.
4. You have 50 minutes to complete this exam, so budget your time wisely.

Q1.[17] Does Vitamin C **improve** health? A study was carried out using 280 skiers at a ski resort. The skiers were all roughly of the same age and had similar nutrition during the study period. Half of the skiers were randomly assigned to take a Vitamin C pill each day, while the other half were given a placebo. At the end of the study period the researchers recorded whether or not each person was exhibiting any symptoms of a cold.

	Placebo	Vitamin C
Cold Symptoms	44	28
NO Cold Symptoms	96	112

a) Which of the following would be appropriate hypotheses for this investigation into people with cold symptoms? Circle all that apply.

- | | | |
|---|---|--|
| (i) $H_0 : \mu_{\text{placebo}} = \mu_{\text{vitC}}$
$H_A : \mu_{\text{placebo}} > \mu_{\text{vitC}}$ | (iv) $H_0 : \pi_{\text{placebo}} = \pi_{\text{vitC}}$
$H_A : \pi_{\text{placebo}} > \pi_{\text{vitC}}$ | (vii) $H_0 : \pi_{\text{placebo}} - \pi_{\text{vitC}} = 0$
$H_A : \pi_{\text{placebo}} - \pi_{\text{vitC}} < 0$ |
| (ii) $H_0 : \pi_{\text{placebo}} = \pi_{\text{vitC}}$
$H_A : \pi_{\text{placebo}} \neq \pi_{\text{vitC}}$ | (v) $H_0 : \mu_{\text{placebo}} - \mu_{\text{vitC}} = 0$
$H_A : \mu_{\text{placebo}} - \mu_{\text{vitC}} > 0$ | (viii) $H_0 : \mu_{\text{placebo}} - \mu_{\text{vitC}} = 0$
$H_A : \mu_{\text{placebo}} - \mu_{\text{vitC}} \neq 0$ |
| (iii) $H_0 : \mu_{\text{placebo}} = \mu_{\text{vitC}}$
$H_A : \mu_{\text{placebo}} \neq \mu_{\text{vitC}}$ | (vi) $H_0 : \pi_{\text{placebo}} - \pi_{\text{vitC}} = 0$
$H_A : \pi_{\text{placebo}} - \pi_{\text{vitC}} > 0$ | |

b) Which of the following would be the correct statistic for the given test? Circle one.

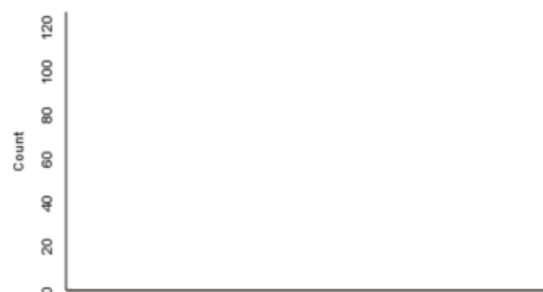
- (i) 28/140
- (ii) 44/140
- (iii) $(44/140) - (28/140) = 0.314 - 0.20 = 0.114$
- (iv) $(44/72) - (28/72) = 0.611 - 0.389 = 0.222$

c) A Chi-Squared Goodness-of-Fit Test would be more appropriate to implement here than the Chi-Squared Test for Independence. Circle one.

TRUE

FALSE

d) Suppose your chi-squared test-statistic turned out to be 4.786 thus giving a p-value of 0.029. Provide a rough sketch of what the null distribution of chi-squared statistics would look like indicating a shaded area for the p-value. You should be specific and clearly show where the shaded area will start.



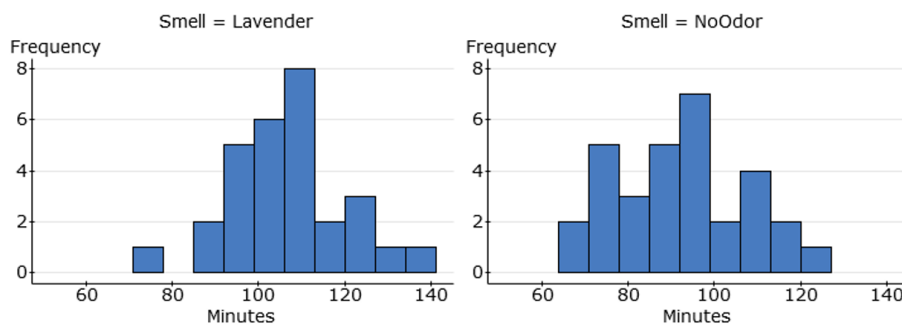
e) What was/were the main purpose(s) of the use of *randomness* in the actual study? Circle one.

- (i) To allow the researchers to generalize the results to a larger population.
- (ii) To allow the researchers to draw a cause-and-effect conclusion from the study.
- (iii) To simulate values of the statistic under the null hypothesis.
- (iv) To replicate the study and increase the accuracy of the results.
- (v) To get a larger sample size.
- (vi) None of the above.

f) Suppose I were to run the appropriate simulation for this situation. What would be the main purpose(s) of the use of *randomness* in this simulation? Circle one.

- (i) To allow the researchers to generalize the results to a larger population.
- (ii) To allow the researchers to draw a cause-and-effect conclusion from the study.
- (iii) To simulate values of the statistic under the null hypothesis.
- (iv) To replicate the study and increase the accuracy of the results.
- (v) To get a larger sample size.
- (vi) None of the above.

Q2.[32] Does a lavender odor encourage customers to stay longer in the store? One study of this question took place in a large department store on one Saturday afternoon. Similar products were placed in two sections of the store to emulate two stores. Each customer was randomly assigned to be in one of the two sections of the store. On one section, a relaxing lavender odor was used. On the other identical section no scent was used. (You can assume that the areas were distant enough that the smells were not mixed). For each customer, the time (in minutes) spent in the store was recorded. A total of 60 customers between 37-78 years old entered the store and came away with a purchase on that Saturday evening. Here is a description of the sample:



	n	Mean	SD
No Odor	31	91.29	14.68
Lavender	29	106.17	13.07

a) Is this an observational or experimental study? Circle one.

Observational Study

Experimental study

b) What is the response variable? What kind of variable is it (categorical or quantitative)?

For the following questions, please use the order: No Odor - Lavender

c) State the null and alternative hypotheses **in symbols**.

H_0 : _____ H_A : _____

Suppose the researcher decided to answer the research question by conducting a hands-on simulation.

d) How many cards would you need to conduct the simulation? _____

e) How will the cards be differentiated? Briefly explain.

f) Fill in the blanks below:

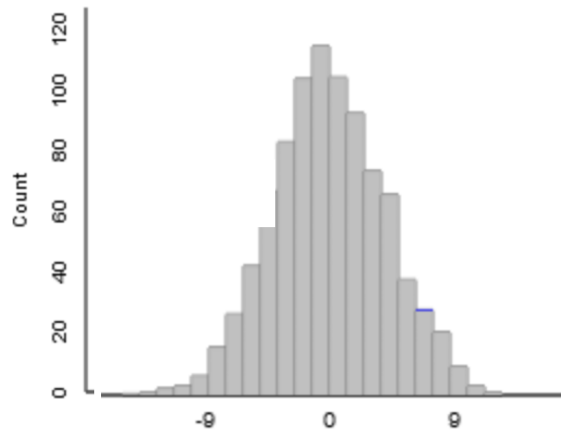
After you shuffle the cards, you should divide the cards into (*how many*) _____ piles. There are (*how many*) _____ cards in the first pile and (*how many*) _____ in the second pile.

The piles represent (*briefly describe below*)

g) What is the simulated statistic you should compute after the cards are divided in piles? Show this with symbols and keep in mind the order that was specified.

The researchers carried out a simulation with 1000 simulated samples using the order (No Odor - Lavender) and obtained the null distribution below.

h) In the distribution below, **clearly shade the area** that you would use to calculate the p-value. Make sure to identify (1) where you start the calculation of the p-value and (2) the direction(s) used. (Note: you will need to calculate your statistic from the study in order to do this. Keep in mind the order that was specified.)



i) Which of the following is the most likely value for the p-value?

- (i) p-value = 0.00 (ii) p-value = 0.03 (iii) p-value = 0.20 (iv) p-value = 1.00

j) Which of the following would be the best overall conclusion in context of the problem?

- (i) With such a small p-value, we have significant evidence to reject the null hypothesis that the average amount of time spent in the store is the same if a lavender odor is present compared to no scent present. We are therefore unable to conclude that the presence of a lavender odor encourages customers to stay longer in the store, on average.
- (ii) With such a small p-value, we do not have significant evidence to reject the null hypothesis that the average amount of time spent in the store is the same if a lavender odor is present compared to no scent present. We are therefore unable to conclude that the presence of a lavender odor encourages customers to stay longer in the store, on average.
- (iii) With such a large p-value, we have significant evidence to reject the null hypothesis that the average amount of time spent in the store is the same if a lavender odor is present compared to no scent present. We can therefore conclude that the presence of a lavender odor encourages customers to stay longer in the store, on average.
- (iv) With such a small p-value, we have significant evidence to reject the null hypothesis that the average amount of time spent in the store is the same if a lavender odor is present compared to no scent present. We can therefore conclude that the presence of a lavender odor encourages customers to stay longer in the store, on average.
- (v) With such a small p-value, we do not have significant evidence to reject the null hypothesis that the average amount of time spent in the store is the same if a lavender odor is present compared to no scent present. We can therefore conclude that the presence of a lavender odor encourages customers to stay longer in the store, on average.
- (vi) With such a large p-value, we do not have significant evidence to reject the null hypothesis that the average amount of time spent in the store is the same if a lavender odor is present compared to no scent present. We are therefore unable to conclude that the presence of a lavender odor encourages customers to stay longer in the store, on average.

k) In a different version of this study, the researcher obtained a p-value of 0.09 and a 95% confidence interval for the parameter of interest to be (-14.01, -5.28). Circle which of the following makes the following statement a true statement.

“In this different version of this study, the two results from the hypothesis test and the confidence interval.....

- (i) seem to agree with one another at the 5% significance level. Nine percent of the time we would obtain a statistic, like the one we saw, somewhere in the interval -14.01% to -5.28%.”
- (ii) are conflicting at the 5% significance level. With a p-value of 0.09 we do not have evidence to reject the null hypothesis, which would mean that our confidence interval would contain 0 (since we are looking a difference of values).”
- (iii) are conflicting at the 5% significance level. There’s a 95% chance that a statistic of 0.09 would end up in the interval (-14.01, -5.28).”
- (iv) seem to agree with one another at the 5% significance level. With a p-value of 0.09, we have evidence to reject the null hypothesis, thus indicating that 0 should be in our confidence interval (since we are considering a difference of values).”

Q3.[8] A student conducted a research study in which the research question was whether financial incentives can improve performance on video games. The subjects in his study were a random sample of 80 university students. In the study, students were randomly assigned to one of two groups. The first group, with 40 students, was offered \$5 for a score above 100 and the other group, with 40 students, was simply told to “do your best.” Each of the 80 students played the video game and their score was recorded.

a) Based on the data from the study, the research student computed a 95% confidence interval: (-12.56, -3.92) points. The order used was “Do Your Best - \$5 Incentive.” Here is the student’s interpretation of the confidence interval:

I am 95% confident that the proportion of video game scores for students who would be told to “do your best” is between 3.92 to 12.56 times higher than the proportion of video game scores for students who would receive the \$5 incentive.

Identify the **three** mistakes committed and fix it. Be brief but clear in your description.

Mistake 1: _____

Fix: _____

Mistake 2: _____

Fix: _____

Mistake 3: _____

Fix: _____

b) The p-value is about 0.03. Select the choice below which best fills in the blanks in the following:

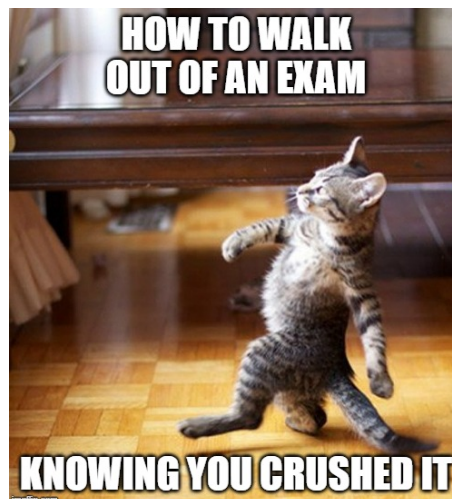
It would [blank 1] surprising to obtain the observed sample results, or more extreme results, if there is really [blank 2] between the dollar incentive and score one achieves on a video game.

- (i) (blank 1) be, (blank 2) no association
 - (ii) (blank 1) not be, (blank 2) no association
 - (iii) (blank 1) be, (blank 2) an association
 - (iv) (blank 1) not be, (blank 2) an association
-

Q4.[5] Are textbooks more expensive on campus than online? To investigate this question, 73 UCLA courses were randomly sampled in Spring 2010. For each course, the price of the primary textbook was measured both at UCLAs bookstore and at Amazon. Two students (A and K) analyzed the data in two separate ways. They used the same data; they just analyzed it differently. These analyses are summarized on the next page. The order used was Bookstore Amazon. Is A's analysis or K's analysis more appropriate for this problem? Briefly describe why. The next page has their analyses available.

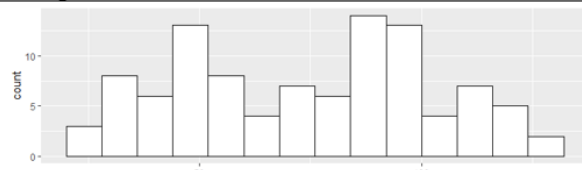
END OF EXAM

There is too much space left, so enjoy this meme:

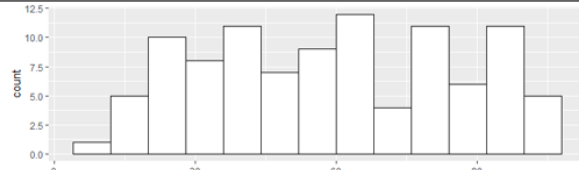


A'S ANALYSIS

Sample Distributions

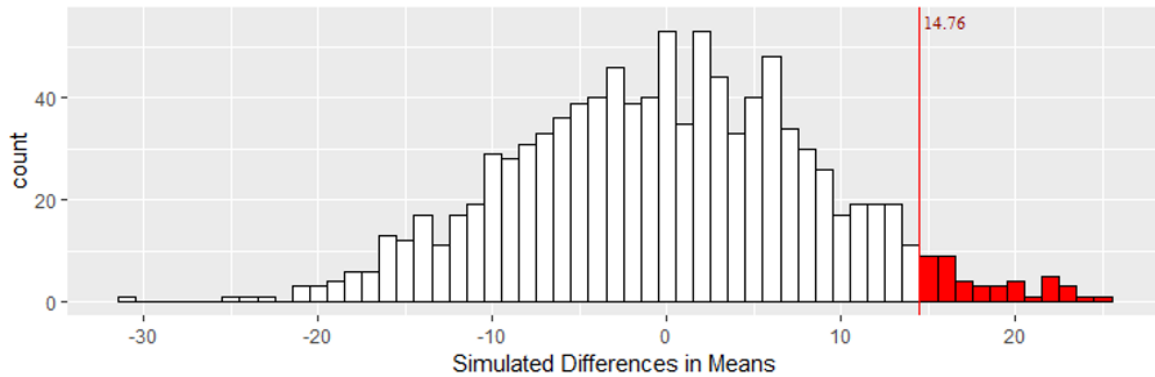


Prices in Bookstore
 $n = 73$, mean = 74.22, sd = 59.66



Prices in Amazon
 $n = 73$, mean = 59.46, sd = 49.00

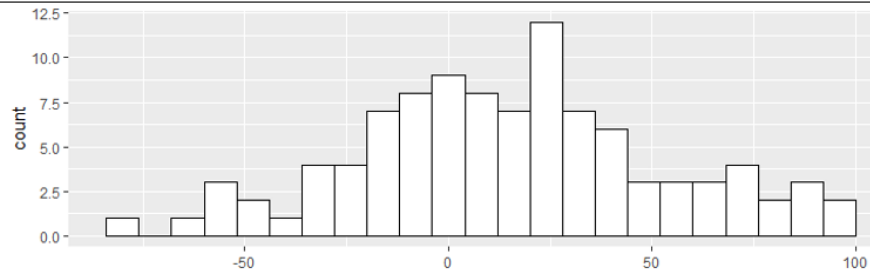
Null distribution using simulation-based approach



p-value = 0.077

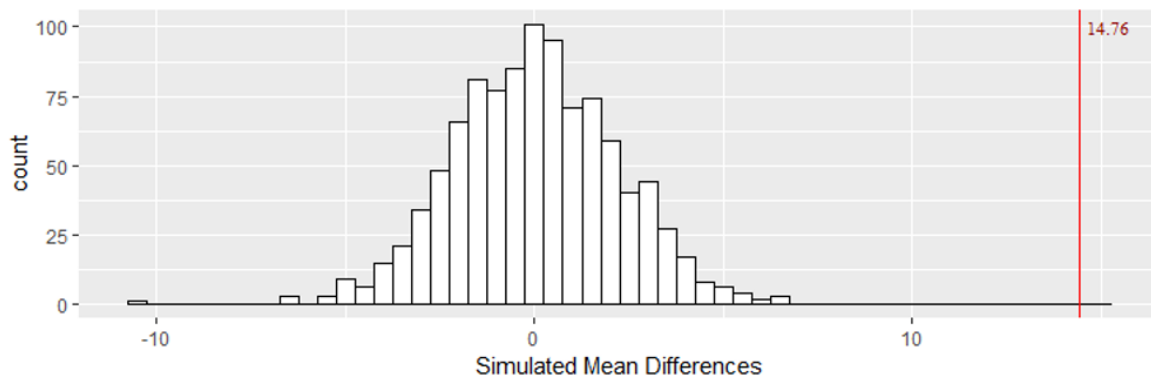
K'S ANALYSIS

Sample Distribution



$n = 73$, mean = 14.76, sd = 37.68

Null Distribution using simulation-based approach



p-value = 0.000