Stat 221 Exam 3 November 21, 2008

Note: Your questions may have appeared in a different order.

1. (16 pts, 4 pts each) For the following questions, identify which of the five types of inference procedures would be appropriate. Choose from the following options:

A. Test about a population proportion
B. Test about a population mean
C. Test for comparing two proportions
D. Test for comparing two means (not matched pairs)
E. Test for comparing two means (matched pairs)

(Report only the appropriate letter. Do not bother to explain.)

a) Researchers played half an hour of joyful music to a group of ten healthy volunteers. They also played half an hour of anxious-sounding music to the same 10 volunteers. The question was whether the subjects' blood vessel diameters would be larger while listening to the joyful music.

E. Test for comparing two means (matched pairs)

b) Are smokers who take a daily nicotine lozenge more likely to quit smoking than smokers who take a placebo lozenge?

C. Test for comparing two proportions

c) Based on a random sample of 50 Cal Poly students, do we have strong evidence that the average IQ score of a Cal Poly student is greater than 105?

B. Test about a population mean

d) Did more than half of Cal Poly students vote in the election this month?

A. Test about a population proportion

2. (21 pts, 3 pts each) 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. For each of the following, circle the best answer among the options presented. Do not bother to explain.

Note: A "statistically significant" difference is one that provides convincing evidence (such as, a small p-value) of a difference between men and women on this issue.

a) 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.

x 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.

b) 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 interpretation?

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

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

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

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

A) A calculation error occurred in the analysis.

B) There must not be a difference after all.

x C) The sample sizes might have been too small.

d) 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 probability that random chance alone led to the difference we observed between the two groups.

B) Men and women differ on this issue.

x C) Either (a) or (b) are possible explanations for this result.

e) Reconsider the previous question. Now think about not possible explanations but *likely* explanations. Which is the more likely 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.

x B) Men and women differ on this issue.

C) They are equally likely explanations.

f) 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

g) 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

3. (25 pts) A psychology study investigated whether people display more creativity when they are thinking about intrinsic motivations or about extrinsic motivations. The subjects were 47 people with extensive experience with creative writing. They were randomly assigned to one of two groups: One group of 24 answered a survey about intrinsic motivations for writing (such as the pleasure of self-expression), and the other group of 23 answered a survey about extrinsic motivations (such as public recognition). Then all subjects were instructed to write a Haiku poem, and these poems were evaluated for creativity with a numerical score determined by a panel of judges. The researchers conjectured that subjects who were thinking about intrinsic motivations would tend to display more creativity than subjects who were thinking about extrinsic motivations.

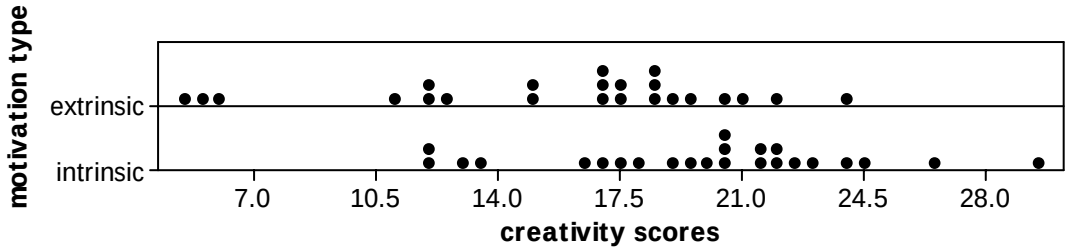
a) (4 pts) State the appropriate null and alternative hypotheses (in symbols) for testing the researchers' conjecture (as stated in the last sentence).

H_0: $\mu_{intrinsic} = \mu_{extrinsic}$

H_a: $\mu_{intrinsic} > \mu_{extrinsic}$

Many students mistakenly thought that this study involved comparing two *proportions* rather than two *means*. The response variable here is creativity score, which is quantitative. One clue to this is the phrase "with a numerical score." Another big clue is the fact that the response variable is displayed with dotplots below.

Consider the following graphs:



b) (6 pts) State the technical conditions required to conduct the appropriate test, and explain whether or not they appear to be satisfied.

The first condition is that the data be collected either from independent randomly samples or from random assignment. We're told that the subjects were randomly assigned to a group, so this condition is satisfied.

The second condition is either that the sample sizes are large or the populations are normally distributed. The sample sizes are not very large here (24 and 23), but the dotplots of creativity scores look roughly normal, so this condition is also satisfied.

The test statistic turns out to equal 2.92.

c) (4 pts) Use this test statistic value (do not bother to recalculate it) to determine the p-value of the test (as accurately as possible with your table).

Looking under 22 degrees of freedom, this p-value is between .005 and .001.

d) (3 pts) Based on your p-value, is the difference between the two groups' performance statistically significant at the $\alpha = .05$ level? (Answer yes or no. Do not bother to explain.)

Yes (because the p-value is less than .05).

e) (8 pts) Summarize the conclusion that you would draw for the researchers. Be sure to address not only the issue of statistical significance but also the question of whether a cause-and-effect conclusion is warranted.

The data provide strong evidence that subjects in the intrinsic motivation group tend to display significantly higher creativity scores (on average) than those in the extrinsic motivation group. Because subjects were randomly assigned to a group, we can conclude that the intrinsic motivation causes greater creativity.

4. (22 pts) In a study conducted in 2003, researchers investigated whether children might be as tempted by toys as by candy for Halloween treats. Test households in five Connecticut neighborhoods offered two bowls to trick-or-treaters: One bowl contained lollipops or fruit candy, and the other bowl contained small, inexpensive Halloween toys, like plastic bugs that glow in the dark. Of the 283 children whose reactions were observed in this study, 148 (52.3%) chose candy and 135 (47.7%) chose toys.

a) (4 pts) What is the parameter of interest here? Provide the appropriate symbol for this parameter, and also describe the parameter in a sentence.

The parameter is the proportion of all trick-or-treaters in these Connecticut neighborhoods who would choose a toy over candy. The symbol for a population proportion is π.

Many students mistakenly wrote the research question (are children equally likely to choose toy or candy?) or the population (trick-or-treaters in these neighborhoods) as the parameter. Many did not mention that the parameter is a proportion. Many who mentioned proportion did not say "proportion who choose the toy." You could also have focused on the proportion who would choose candy.

b) (10 pts) Determine a 95% confidence interval for this parameter.

A 95% CI for the population proportion who would choose the toy is: $\hat{p} \pm z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$, which gives $.477 \pm 1.96\sqrt{\frac{.477 \times .523}{283}}$, which is .477 ± .058, which is the interval (.419, .535).

If instead you find a 95% CI for the population proportion who would choose the candy, you get .523 ± .058, which is the interval (.465, .581).

c) (4 pts) Think about conducting a significance test of whether the relevant population proportion equals .5. Based solely on your confidence interval (without actually conducting the test), what can you say about how the p-value of the test would turn out? Explain your answer.

The 95% confidence interval for p does include the value .5, so we would not reject that value with the $\alpha = .05$ significance level. The p-value is therefore greater than .05.

d) (4 pts) Suppose that someone else used the same sample data and mistakenly calculated a confidence interval to be (.514, .631). Explain why this confidence interval cannot possibly be correct.

This interval does not have the sample proportion (.523 or .477) as its midpoint.

5. (16 pts) I recently read an article in *Science* magazine that described an experiment concerning people's willingness to be organ donors. All of the subjects in the experiment were told that to imagine that they have moved to a new state and have applied for a driver's license, and they must make a decision about whether to become an organ donor. Some of the subjects were randomly assigned to be told that the default option is not to be a donor, and the rest of the subjects were told that the default option is to be a donor. All subjects were then given the choice of whether or not to become an organ donor. The researchers suspected that a higher proportion of people are willing to be donors when the default option is to be a donor.

a) (4 pts) Identify the explanatory and response variables in this study.

Explanatory: the default option presented

Response: whether or not the person chose to be an organ donor

b) (4 pts) State the appropriate null and alternative hypotheses (in symbols) for testing the researchers' suspicion.

H_0: $\pi_{\text{def yes}} = \pi_{\text{def no}}$ H_a: $\pi_{\text{def yes}} > \pi_{\text{def no}}$

where π_i is the population proportion who would agree to become an organ donor if presented with default option i.

c) (4 pts) Which kind of test is appropriate here? (Circle your answer. Do not bother to explain.)

A. Test about a population proportion
B. Test about a population mean
C. Test for comparing two proportions
D. Test for comparing two means (not matched pairs)
E. Test for comparing two means (matched pairs)

d) (4 pts) The article reported that 42% of the subjects in the "default is not to be a donor" group decided to become a donor, compared to 82% in the "default is to be a donor" group. What further information would you need in order to conduct the test (i.e., to calculate a test statistic and p-value)?

You need to know the number of people in each group.