AP® STATISTICS
2007 SCORING GUIDELINES

Question 5

Intent of Question

The primary goals of this statistical inference question are to assess a student's ability to: (1) distinguish an observational study from an experiment; (2) state the appropriate hypotheses for a research problem; (3) check the appropriate conditions for an inference procedure; and (4) interpret standard results for an inference procedure that is unfamiliar to students.

Solution

Part (a):

This is an experiment because the researchers imposed treatments by randomly assigning drivers to the two different conditions (simulated driving while talking on a cell phone versus simulated driving while talking to a passenger).

Part (b):

Let p_{cell} denote the proportion of drivers who miss an exit while using a cell phone and p_{pass} denote the proportion of drivers who miss an exit while talking to a passenger.

H_0: $p_{cell} = p_{pass}$
H_A: $p_{cell} > p_{pass}$

Part (c):

The conditions required for a two-sample z-test of equal proportions are:
(1) independent random samples or random assignment, and
(2) large sample sizes $\left[n_1 \hat{p}_1 \geq 10, n_1(1-\hat{p}_1) \geq 10, n_2 \hat{p}_2 \geq 10, n_2(1-\hat{p}_2) \geq 10 \right]$.
Random assignment is stated in the stem so the first condition is met. However, the numbers of successes ($n_{cell}\hat{p}_{cell} = 7$ and $n_{pass}\hat{p}_{pass} = 2$) are both smaller than 10, so the large sample condition is not met in this situation. Note: If the student uses the rule of thumb with 10 replaced by 5, then the number of successes for the second sample is still too small.

Part (d):

Interpretation: Assuming that talking on a cell phone and talking to a passenger are equally distracting (there is no difference in the two population proportions of drivers who will miss the exit), the p-value measures the chance of observing a difference in the two sample proportions as large as or larger than the one observed.

Conclusion: Since the p-value 0.0683 is larger than 0.05, we cannot reject the null hypothesis. That is, we do not have statistically significant evidence to conclude that using a cell phone is more distracting to drivers than talking to another passenger in the car.

Notice that if we increase the significance level to 0.1, then we could reject the null hypothesis and conclude that drivers are significantly more distracted when using a cell phone.

AP® STATISTICS
2007 SCORING GUIDELINES

Question 5 (continued)

Scoring

Parts (a) and (b) are scored as essentially correct (E) or incorrect (I). Parts (c) and (d) are scored as essentially correct (E), partially correct (P), or incorrect (I).

Part (a) is scored as essentially correct (E) if the student indicates that this is an experiment because treatments were imposed.

Part (a) is scored as incorrect (I) if no explanation is provided, or the student says that this is an observational study.

Part (b) is scored as essentially correct (E) if the student correctly identifies the two population proportions with the correct hypotheses. Nonstandard notation must indicate reference to population proportions.

Part (b) is scored as incorrect (I) if the student is *clearly* referring to the sample proportions.

Part (c) is scored as essentially correct (E) if the student provides both conditions and correctly comments on both.

Part (c) is scored as partially correct (P) if the student provides and correctly comments on only one of the conditions.

Part (c) is scored as incorrect (I) if conditions are provided but no correct comments are given.

Part (d) is scored as essentially correct (E) if the *p*-value is correctly interpreted *AND* the correct conclusion is provided *AND* context is given.

Part (d) is scored as partially correct (P) if:

i) either the *p*-value is correctly interpreted OR the correct conclusion is provided
AND
ii) context is given.

Part (d) is scored as incorrect (I) if *neither* a correct interpretation of the *p*-value in context *NOR* a correct conclusion in context is provided.

In part (d) if both an α and a *p*-value are given together, the linkage between the *p*-value and the conclusion is implied. If no α is given, the solution must be explicit about the linkage by giving a correct interpretation of the *p*-value or explaining how the conclusion follows from the *p*-value.

Note: Any choice of an α could have been made as long as the appropriate interpretation is made relative to that choice of α.

AP® STATISTICS
2007 SCORING GUIDELINES

Question 5 (continued)

Each essentially correct (E) response counts as 1 point; each partially correct (P) response counts as ½ point.

4 **Complete Response**

3 **Substantial Response**

2 **Developing Response**

1 **Minimal Response**

If a response is between two scores (for example, 2½ points), use a holistic approach to determine whether to score up or down depending on the strength of the response and communication.

5. Researchers want to determine whether drivers are significantly more distracted while driving when using a cell phone than when talking to a passenger in the car. In a study involving 48 people, 24 people were randomly assigned to drive in a driving simulator while using a cell phone. The remaining 24 were assigned to drive in the driving simulator while talking to a passenger in the simulator. Part of the driving simulation for both groups involved asking drivers to exit the freeway at a particular exit. In the study, 7 of the 24 cell phone users missed the exit, while 2 of the 24 talking to a passenger missed the exit.

(a) Would this study be classified as an experiment or an observational study? Provide an explanation to support your answer.

This would be considered an experiment because treatments were imposed, talking on a cell phone, and talking to a passenger.

(b) State the null and alternative hypotheses of interest to the researchers.

I. P_1 = proportion of drivers who missed the exit while talking on a cellphone

P_2 = proportion of drivers who missed the exit while talking to a passenger.

H_0: $P_1 = P_2$

H_a: $P_1 > P_2$

(c) One test of significance that you might consider using to answer the researchers' question is a two-sample z-test. State the conditions required for this test to be appropriate. Then comment on whether each condition is met.

To use a 2-sample z-test the 2 samples must be 2 independent SRS, σ must be known, and the distribution must be normal.

① It was stated that the 2 samples were randomly assigned, but it did not say if participants were randomly selected. X

② The σ values of each sample are unknown. X

③ For a normal distribution, $n_1\hat{p} \geq 5$, $n_2\hat{p} \geq 5$, $n_1(1-\hat{p}) \geq 5$, $n_2(1-\hat{p}) \geq 5$

$24(0.078) = 0.1875 \not\geq 5$, $24(1-0.0078) = 23.81 \geq 5$, $24(0.0078) = 0.1875 \not\geq 5$

$24(1-0.0078) = 23.81 \geq 5$ → because they are not all met, the distribution is not normal. X

→ not all 3 conditions are met, therefore a 2-sample z test is not an appropriate test of significance.

(d) Using an advanced statistical method for small samples to test the hypotheses in part (b), the researchers report a p-value of 0.0683. Interpret, in everyday language, what this p-value measures in the context of this study and state what conclusion should be made based on this p-value.

The p-value calculates the probability that a more extreme value will occure, assuming the null hypothesis is true.

In this case, if we were testing at a 10% level, the p-value is smaller than 0.10 and we would have significant evidence to say people talking on the phone were more distracted than those talking to passengers.

However, if we were testing at the 5% level, the p-value is greater than 0.05 and we would not be able to say with significant evidence that there is a difference in distraction between talking on the cell phone and talking to a passenger.

GO ON TO THE NEXT PAGE.

5. Researchers want to determine whether drivers are significantly more distracted while driving when using a cell phone than when talking to a passenger in the car. In a study involving 48 people, 24 people were randomly assigned to drive in a driving simulator while using a cell phone. The remaining 24 were assigned to drive in the driving simulator while talking to a passenger in the simulator. Part of the driving simulation for both groups involved asking drivers to exit the freeway at a particular exit. In the study, 7 of the 24 cell phone users missed the exit, while 2 of the 24 talking to a passenger missed the exit.

(a) Would this study be classified as an experiment or an observational study? Provide an explanation to support your answer.

this study would be classified as an experiment because a treatment is being imposed on the subjects. subjects are put into a driving simulator. one group is given a cell phone to talk on and the other gets a passenger to talk to. if this were an observational study, the environment would not be manipulated.

(b) State the null and alternative hypotheses of interest to the researchers.

Null: The proportion of drivers who missed the exit while talking on the phone and talking to a passenger are the same ($H_0: P_1 = P_2$).

alternative: the proportion of drivers who missed the exit while talking on the phone is greater than the proportion who missed the exit while talking to a passenger ($H_a: P_1 > P_2$)

* P_1 = proportion of drivers who missed the exit while on the phone
P_2 = proportion of drivers who missed the exit while talking to a passenger

GO ON TO THE NEXT PAGE.

(c) One test of significance that you might consider using to answer the researchers' question is a two-sample z-test. State the conditions required for this test to be appropriate. Then comment on whether each condition is met.

- samples are independent of each other
  - different people are used in treatment groups, so chance they will miss the exit is independent
- samples are simple random samples
  - stated in problem that people were randomly assigned
- $n_1\hat{p}_1$, $n_1(1-\hat{p}_1)$, $n_2\hat{p}_2$, $n_2(1-\hat{p}_2)$ are all ≥ 5
  - $(24)(\frac{7}{24}) = 7 > 5$
  - $(24)(\frac{17}{24}) = 17 > 5$
  - $(24)(\frac{2}{24}) = 2 < 5$ ... this condition is not met because $n_2(\hat{p}_2) < 5$
  - $(24)(\frac{22}{24}) = 22 > 5$

(d) Using an advanced statistical method for small samples to test the hypotheses in part (b), the researchers report a p-value of 0.0683. Interpret, in everyday language, what this p-value measures in the context of this study and state what conclusion should be made based on this p-value.

There is a probability of .0683 that 7/24 cellphone users and 2/24 passenger talkers will miss their exit or even more extreme proportions will occur. At $\alpha = .05$, there isn't enough evidence to reject the null hypothesis. The proportion of drivers who missed the exit while talking on a cell phone and talking to a passenger is the same. However, there may be some error in the conclusion because one assumption is not met.

5. Researchers want to determine whether drivers are significantly more distracted while driving when using a cell phone than when talking to a passenger in the car. In a study involving 48 people, 24 people were randomly assigned to drive in a driving simulator while using a cell phone. The remaining 24 were assigned to drive in the driving simulator while talking to a passenger in the simulator. Part of the driving simulation for both groups involved asking drivers to exit the freeway at a particular exit. In the study, 7 of the 24 cell phone users missed the exit, while 2 of the 24 talking to a passenger missed the exit.

(a) Would this study be classified as an experiment or an observational study? Provide an explanation to support your answer.

This study would be classified as an observational study. There was no control group present so it was not an experiment

(b) State the null and alternative hypotheses of interest to the researchers.

H_0 p of missing exit on phone = p of missing exit talking to passanger

H_a p of missing exit on phone > p of missing exit talkin to passenger

p = probability

(c) One test of significance that you might consider using to answer the researchers' question is a two-sample z-test. State the conditions required for this test to be appropriate. Then comment on whether each condition is met.

A two sample z test requires that $n\hat{p}$ for each sample be > 5. It is not for the group talking to the passenger. Also, sample sizes must be large, > 30. This was not met for either.

(d) Using an advanced statistical method for small samples to test the hypotheses in part (b), the researchers report a p-value of 0.0683. Interpret, in everyday language, what this p-value measures in the context of this study and state what conclusion should be made based on this p-value.

In the context of the study, this p-value measures the probability that, judging from the data, drivers in the two samples are equally likely to miss the exit. Because the p value is so small, & about 7%, it should be concluded that drivers are more likely to miss the exit while talking on the phone than while talking to a passenger.

AP® STATISTICS
2007 SCORING COMMENTARY

Question 5

Overview

The primary goals of this statistical inference question were to assess a student's ability to: (1) distinguish an observational study from an experiment; (2) state the appropriate hypotheses for a research problem; (3) check the appropriate conditions for an inference procedure; and (4) interpret standard results for an inference procedure that is unfamiliar to students.

Sample: 5A
Score: 4

Each part of this response is complete, concise, and clearly communicated. In part (a) the response clearly and correctly classifies the study as an experiment and explains that treatments were imposed. In part (b) the response correctly states the null and alternative hypotheses and gives standard notation for the null and alternative hypotheses for the two-sample test of proportions. In part (c) the response correctly lists conditions (random assignment and the numerical checks for approximate normality) and correctly comments on each of the conditions. The response uses the pooled proportions estimate to conduct the numerical checks for approximate normality. In part (d) the response gives a correct interpretation of the *p*-value and correctly interprets the conclusion in context. This is a complete response.

Sample: 5B
Score: 3

Three of the four parts are complete and clearly communicated. In part (a) the response clearly and correctly classifies the study as an experiment and gives as an explanation that treatments were imposed and that the subjects were randomly allocated into two groups. In part (b) the response correctly states the null and alternative hypotheses for the two-sample test of proportions using standard notation. The response distinguishes and defines the two population proportions. In part (c) the response correctly lists conditions (random allocation and the numerical checks for approximate normality) and correctly comments on each of the conditions. In part (d) the response gives an incorrect interpretation of the *p*-value and incorrectly gives the conclusion as a definitive statement on the equality of the two proportions, in effect accepting the null hypothesis. This is a substantial response.

Sample: 5C
Score: 2

One of the four parts of this response is complete and clearly communicated, and two parts are partially correct. In part (a) the response incorrectly classifies the study as an observational study and gives as an explanation that there was no control. This is an incorrect response for part (a). In part (b) the response correctly states the null and alternative hypotheses for the two-sample test of proportions using standard notation. The response also distinguishes the two population proportions by defining them. In part (c) the response correctly lists and checks the sample-size requirements for approximate normality but fails to recognize the necessity of random assignment. This is a partially correct response for part (c). In part (d) the response gives an incorrect interpretation for the *p*-value but a correct conclusion based on the observed *p*-value. Overall, this is a developing response.