Question 6

Intent of Question

The primary goals of this question are to evaluate a student's ability to apply the concepts of significance testing to a new situation, in particular to: (1) state hypotheses for a parameter of interest given a research question; (2) recognize that a large sample test should not be used in this setting; (3) identify the possible values of a new test statistic and calculate the probability distribution for this new test statistic, assuming the null hypothesis is true; (4) use the probability distribution of the test statistic under the null hypothesis to identify possible significance levels; (5) conduct a significance test for a small set of data from an initial study; and (6) make a recommendation to improve on the initial study.

Solution

Part (a):

$H_0 : p = 0.5$ versus $H_a : p \neq 0.5$

Part (b):

The conditions for the large sample one-proportion *z*-test are not satisfied. $np = n(1 - p) = 8 \times 0.5 = 4 < 5$.

Part (c):

X will follow a binomial distribution with $n = 8$ and $p = 0.5$. The possible values of *X* and their corresponding probabilities are given in the table below.

X	Probability
0	0.00391
1	0.03125
2	0.10937
3	0.21875
4	0.27344
5	0.21875
6	0.10937
7	0.03125
8	0.00391

Part (d):

No, there is no possible test with a *p*-value of exactly 0.05.

The probability that none of the individuals $(X = 0)$ or all of the individuals $(X = 8)$ prefer Citrus Fresh is $2 \times 0.003906 = 0.0078$, which is less than 0.05.

The probability that one or fewer of the individuals $(X \leq 1)$ or seven or more of the individuals $(X \geq 7)$ prefer Citrus Fresh is $2 \times (0.003906 + 0.031250) = 0.070312$, which is greater than 0.05.

Question 6 (continued)

Part (e):

For the preference data provided, $X = 2.$ From the table of binomial probabilities computed in part (c), the probability that two or fewer of the individuals $(X \le 2)$ or six or more of the individuals $(X \ge 6)$ prefer Citrus Fresh when $p = 0.5$ is $2 \times (0.003906 + 0.031250 + 0.109375) = 0.289062$. Because the *p*-value of 0.289062 is greater than any reasonable significance level, say 0.070312, we would not reject the null hypothesis that $p = 0.5.$ That is, we do not have statistically significant evidence for a consumer preference between Citrus Fresh and Tropical Taste.

Part (f):

Increase the number of consumers involved in the preference test. More consumers will give you more data, and you will be better able to detect a difference between the population proportion of consumers who prefer Citrus Fresh and 0.5. The sample proportion in the initial study was only 0.25 (2/8), but we were not able to reject the null hypothesis that $p = \frac{1}{2}.$ By increasing the number of consumers, a difference of that magnitude would allow the null hypothesis to be rejected. For example, with $n = 80$ and $X = 20$ the large sample *z*-statistic would be $z = \frac{0.25 - 0.5}{\sqrt{\frac{(0.5)(0.5)}{80}}} = -4.47$ and the *p*-value would be approximately zero.

OR

The large sample test statistic is $z = \frac{\hat{p} - 0.5}{\sqrt{\frac{(0.5)(0.5)}{n}}} = \sqrt{n}\left(\frac{\hat{p} - 0.5}{0.5}\right)$, which would get larger as n is increased, for the same $\hat{p}$. Thus, by using more consumers it would eventually be possible to reject H_0 if there really is a difference and identify the preferred juice.

<u>Scoring</u>

Parts (a) and (b) are combined and scored as essentially correct (E), partially correct (P), or incorrect (I). Parts (c) and (d) are combined and scored as essentially correct (E), partially correct (P), or incorrect (I). Part (e) contains two parts, correct mechanics and conclusion, and is scored essentially correct (E), partially correct (P), or incorrect (I). Part (f) is scored as essentially correct (E), partially correct (P), or incorrect (I).

Parts (a) and (b) are essentially correct (E) if both parts are correct.

Parts (a) and (b) are partially correct (P) if one of the two parts is correct.

Notes for parts (a) and (b):

If a one-sided alternative is used in part (a), then the maximum score for the AB component is partially correct (P).

Question 6 (continued)

Other reasonable statements about the conditions for inference not being met are acceptable. Some examples are:

- Since $n = 8,$ the number of successes and the number of failures will both be less than 10.
- The sample size is too small to use the large-sample inference procedure.

A response in part (b) about a test for comparing proportions from two independent samples (one for Tropical Taste and one for Citrus Fresh) should be scored as incorrect (I).

Parts (c) and (d) are essentially correct (E) if both parts are correct.

Parts (c) and (d) are partially correct (P) if one of the two parts is correct.

Notes for parts (c) and (d):

Part (c) is correct if binomial probabilities are correctly calculated for each of the nine possible outcomes using $n = 8$ and $p = 0.5.$

Part (d) is correct if the response includes all three of the following:

- The student states that a 0.05 level test is not possible.
- The justification includes a correct description of the test with significance level 0.00782.
- The justification includes a correct description of the test with significance level 0.07032.

Part (e) is essentially correct (E) if both parts are correct.

Part (e) is partially correct (P) if one part is correct.

Notes for part (e):

The mechanics are essentially correct if the student uses the information from the response to part (c) to compute the appropriate *p*-value.

- If a large sample *z*-test, or a *t*-test, is presented, the mechanics are scored as incorrect.
- Using only 0.10937, the probability that *X* is 2 from the table in part (c), is incorrect.

To be essentially correct, a correct conclusion must be stated in the context of the problem with linkage between a *p*-value and a level of significance α .

- If the student refers to the rejection region identified in part (d), they do not need to restate the significance level from part (e).
- If the student conducts an $\alpha = 0.05$ level test by comparing the correct *p*-value to 0.05, even though they just said in part (d) that a 0.05 level test is not possible, the response should be scored as essentially correct.

Part (f) is essentially correct (E) if the provides a reasonable recommendation in the context of comparing preferences for the two juices with an appropriate statistical justification. For example, the student describes the impact of a larger sample size on the hypothesis testing procedure used in the initial study by saying that increasing the sample size will reduce the chance of making a Type II error, increasing the sample size will increase the power of the test used to detect a difference in juice preferences in part (d), or increasing the sample size will reduce the standard error for the estimated proportion of people who prefer Citrus Fresh.

Part (f) is partially correct (P) if the student provides a reasonable recommendation in the context of comparing preferences for the two juices but does not provides sufficient statistical justification that refers to the initial

Question 6 (continued)

study, e.g., a larger sample is needed to meet the conditions for using a *z*-test, larger samples are more representative of the population, or increasing the sample size makes the sample or population more normal
OR
Gives a weak recommendation and provides a good statistical justification, for example, providing the subjects with a cracker between juices to reduce possible carryover effects.

Incorrect (I) if the student fails to provide a reasonable recommendation in the context of comparing preferences for the two juices.

Each essentially correct response is worth 1 point; each partially correct response is worth 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.

STATISTICS
SECTION II
Part B
Question 6

Spend about 25 minutes on this part of the exam.

Percent of Section II grade—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Sunshine Farms wants to know whether there is a difference in consumer preference for two new juice products —Citrus Fresh and Tropical Taste. In an initial blind taste test, 8 randomly selected consumers were given unmarked samples of the two juices. The product that each consumer tasted first was randomly decided by the flip of a coin. After tasting the two juices, each consumer was asked to choose which juice he or she preferred, and the results were recorded.

   (a) Let p represent the population proportion of consumers who prefer Citrus Fresh. In terms of p, state the hypotheses that Sunshine Farms is interested in testing.

$H_0: p = .5$

$H_0: p \neq .5$

   (b) One might consider using a one-proportion z-test to test the hypotheses in part (a). Explain why this would not be a reasonable procedure for this sample.

A one-proportion z-test would not be appropriate because the sample size is too small and the standard deviation is unknown.

GO ON TO THE NEXT PAGE.

6A2

(c) Let X represent the number of consumers in the sample who prefer Citrus Fresh. Assuming there is no difference in consumer preference, find the probability for each possible value of X. Record the x-values and the corresponding probabilities in the table below.

x	$p(x)$
0	.00390625
1	.03125
2	.109375
3	.21875
4	.2734375
5	.21875
6	.109375
7	.03125
8	.00390625

(d) When testing the hypotheses in part (a), Sunshine Farms will conclude that there is a consumer preference if too many or too few individuals prefer Citrus Fresh. Based on your probabilities in part (c), is it possible for the significance level (probability of rejecting the null hypothesis when it is true) for this test to be exactly 0.05 ? Justify your answer.

No because this is a discrete random variable. In order for the significance level to be exactly 0.05 (according to the table above), the value of X would have to be between 7 and 8, and there are no values between 7 and 8.

GO ON TO THE NEXT PAGE.

(e) The preference data for the 8 randomly selected consumers are given in the table below.

Individual	Juice Preference
1	Tropical Taste
2	Citrus Fresh
3	Tropical Taste
4	Tropical Taste
5	Tropical Taste
6	Citrus Fresh
7	Tropical Taste
8	Tropical Taste

Based on these preferences and your previous work, test the hypotheses in part (a).

$\frac{2}{8}$ people prefer Citrus fresh, thus $X = 2$. The probability that one observes a value at least as extreme as this, assuming the null hypothesis is true, equals:

$P(X=0) + P(X=1) + P(X=2) = .1445$

Because this is a two-tailed test, the p-value is double, thus:

$p = 2(.1445) = .2890$

A significance level of .2890 is insufficient to conclude that there exists a preference for either juice product.

6A21

(f) Sunshine Farms plans to add one of these two new juices—Citrus Fresh or Tropical Taste—to its production schedule. A follow-up study will be conducted to decide which of the two juices to produce. Make one recommendation for the follow-up study that would make it better than the initial study. Provide a statistical justification for your recommendation in the context of the problem.

One suggestion would be to significantly increase the number of customers sampled. If $n \geq 30$, then a z-test could be used. A larger sample size would also increase the power of the test, increasing the probability that the null hypothesis will be rejected when it is false, thus allowing Sunshine Farms to reach a conclusion as to which juice to produce.

STOP

END OF EXAM

THE FOLLOWING INSTRUCTIONS APPLY TO THE COVERS OF THE SECTION II BOOKLET.

- **MAKE SURE YOU HAVE COMPLETED THE IDENTIFICATION INFORMATION AS REQUESTED ON THE FRONT AND BACK COVERS OF THE SECTION II BOOKLET.**
- **CHECK TO SEE THAT YOUR AP NUMBER LABEL APPEARS IN THE BOX(ES) ON THE COVER(S).**
- **MAKE SURE YOU HAVE USED THE SAME SET OF AP NUMBER LABELS ON ALL AP EXAMS YOU HAVE TAKEN THIS YEAR.**

6B1

STATISTICS
SECTION II
Part B
Question 6
Spend about 25 minutes on this part of the exam.
Percent of Section II grade—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Sunshine Farms wants to know whether there is a difference in consumer preference for two new juice products —Citrus Fresh and Tropical Taste. In an initial blind taste test, 8 randomly selected consumers were given unmarked samples of the two juices. The product that each consumer tasted first was randomly decided by the flip of a coin. After tasting the two juices, each consumer was asked to choose which juice he or she preferred, and the results were recorded.

   (a) Let p represent the population proportion of consumers who prefer Citrus Fresh. In terms of p, state the hypotheses that Sunshine Farms is interested in testing.

   H_0: $p = 0.5$.
   H_a: $p \neq 0.5$.

   (b) One might consider using a one-proportion z-test to test the hypotheses in part (a). Explain why this would not be a reasonable procedure for this sample.

   Because the sample size is not large enough
   $np = (8)(0.5) = 4 < 5$ (X)

GO ON TO THE NEXT PAGE.

6B2

(c) Let X represent the number of consumers in the sample who prefer Citrus Fresh. Assuming there is no difference in consumer preference, find the probability for each possible value of X. Record the x-values and the corresponding probabilities in the table below.

x	$p(x)$
0	0.0039
1	0.03125
2	0.109375
3	0.21875
4	0.273438
5	0.21875
6	0.109375
7	0.03125
8	0.0039

I would use binomial theorem because
1) there are n trials
2) there are two possible outcomes
3) two samples are random and independent. ✓

(d) When testing the hypotheses in part (a), Sunshine Farms will conclude that there is a consumer preference if too many or too few individuals prefer Citrus Fresh. Based on your probabilities in part (c), is it possible for the significance level (probability of rejecting the null hypothesis when it is true) for this test to be exactly 0.05 ? Justify your answer.

No, the sums of the probabilities of consumers preferring particular product either exceeds or does not reach 0.05. Therefore it cannot be exactly 0.05.

GO ON TO THE NEXT PAGE.

(e) The preference data for the 8 randomly selected consumers are given in the table below.

Individual	Juice Preference
1	Tropical Taste
2	Citrus Fresh
3	Tropical Taste
4	Tropical Taste
5	Tropical Taste
6	Citrus Fresh
7	Tropical Taste
8	Tropical Taste

Based on these preferences and your previous work, test the hypotheses in part (a).

The probability that 2 consumers would prefer citrus Fresh when there is no consumer preference is 0.109375.
I would use $\alpha = 0.05$, $0.109375 > 0.05$
Because the p-value is greater than α, there is not enough evidence to reject the null hypothesis that there is no consumer preference.

GO ON TO THE NEXT PAGE.

6B4.

(f) Sunshine Farms plans to add one of these two new juices—Citrus Fresh or Tropical Taste—to its production schedule. A follow-up study will be conducted to decide which of the two juices to produce. Make one recommendation for the follow-up study that would make it better than the initial study. Provide a statistical justification for your recommendation in the context of the problem.

Use a sample with larger sample size. Large sample size will allow researchers to use proportion z-test and reduce the confidence intervals.
Because Standard deviation of proportion is $\sqrt{\frac{p_1(1-p_1)}{n_1}}$, larger sample size will reduce the interval. Furthermore with larger sample size, Sunshine Farms can meet the conditions for using proportion z-test; $np > 5$
$n(1-p) > 5$.

STOP

END OF EXAM

THE FOLLOWING INSTRUCTIONS APPLY TO THE COVERS OF THE SECTION II BOOKLET.

- **MAKE SURE YOU HAVE COMPLETED THE IDENTIFICATION INFORMATION AS REQUESTED ON THE FRONT AND BACK COVERS OF THE SECTION II BOOKLET.**
- **CHECK TO SEE THAT YOUR AP NUMBER LABEL APPEARS IN THE BOX(ES) ON THE COVER(S).**
- **MAKE SURE YOU HAVE USED THE SAME SET OF AP NUMBER LABELS ON ALL AP EXAMS YOU HAVE TAKEN THIS YEAR.**

6 C I

STATISTICS
SECTION II
Part B
Question 6
Spend about 25 minutes on this part of the exam.
Percent of Section II grade—25

Directions: Show all your work. Indicate clearly the methods you use, because you will be graded on the correctness of your methods as well as on the accuracy and completeness of your results and explanations.

6. Sunshine Farms wants to know whether there is a difference in consumer preference for two new juice products—Citrus Fresh and Tropical Taste. In an initial blind taste test, 8 randomly selected consumers were given unmarked samples of the two juices. The product that each consumer tasted first was randomly decided by the flip of a coin. After tasting the two juices, each consumer was asked to choose which juice he or she preferred, and the results were recorded.

(a) Let p represent the population proportion of consumers who prefer Citrus Fresh. In terms of p, state the hypotheses that Sunshine Farms is interested in testing.

H_0 $p^1 = p^2$: There is no differens in the preferchce of the two new juic products

H_a $p^1 \neq p^2$: There is a differens in the preffference of the two new juice products.

(b) One might consider using a one-proportion z-test to test the hypotheses in part (a). Explain why this would not be a reasonable procedure for this sample.

this woud not be reasonable because this is a 2 ~~sa~~ proportion test. if we use a 1 proportion we wive not bet able to compare to two products.

GO ON TO THE NEXT PAGE.

6C2

(c) Let X represent the number of consumers in the sample who prefer Citrus Fresh. Assuming there is no difference in consumer preference, find the probability for each possible value of X. Record the x-values and the corresponding probabilities in the table below.

x	$p(x)$
~~0~~	~~0.0039~~
1	0.03125
2	0.10937
3	0.21875
4	0.27343
5	0.21875
6	0.10937
7	0.03125
8	0.0039

(d) When testing the hypotheses in part (a), Sunshine Farms will conclude that there is a consumer preference if too many or too few individuals prefer Citrus Fresh. Based on your probabilities in part (c), is it possible for the significance level (probability of rejecting the null hypothesis when it is true) for this test to be exactly 0.05 ? Justify your answer.

yes, but very unlikely

GO ON TO THE NEXT PAGE.

6C3

(e) The preference data for the 8 randomly selected consumers are given in the table below.

Individual	Juice Preference
1	Tropical Taste
2	Citrus Fresh
3	Tropical Taste
4	Tropical Taste
5	Tropical Taste
6	Citrus Fresh
7	Tropical Taste
8	Tropical Taste

$\frac{6}{8}$ $\frac{2}{4}$

Based on these preferences and your previous work, test the hypotheses in part (a).

$z = 2$

$p = 0.0455$

Since the p-value is very small; this suggest that there is a different between the two juice products. Thus we reject ~~our~~ null hypothesis at ~~1~~5 percent.

6C4

(f) Sunshine Farms plans to add one of these two new juices—Citrus Fresh or Tropical Taste—to its production schedule. A follow-up study will be conducted to decide which of the two juices to produce. Make one recommendation for the follow-up study that would make it better than the initial study. Provide a statistical justification for your recommendation in the context of the problem.

One reccomendation would be to increase the sample size. Using 8 people to justify your result is not reasonable. Therefore by increase your ~~s~~ sample size it would decrease the standard deviation which helps us to ~~[illegible]~~ ~~the~~ ~~a~~ find the true proportion of the preference between the two products.

STOP

END OF EXAM

THE FOLLOWING INSTRUCTIONS APPLY TO THE COVERS OF THE SECTION II BOOKLET.

- **MAKE SURE YOU HAVE COMPLETED THE IDENTIFICATION INFORMATION AS REQUESTED ON THE FRONT AND BACK COVERS OF THE SECTION II BOOKLET.**
- **CHECK TO SEE THAT YOUR AP NUMBER LABEL APPEARS IN THE BOX(ES) ON THE COVER(S).**
- **MAKE SURE YOU HAVE USED THE SAME SET OF AP NUMBER LABELS ON ALL AP EXAMS YOU HAVE TAKEN THIS YEAR.**

AP® STATISTICS
2006 SCORING COMMENTARY (Form B)

Question 6

Sample: 6A
Score: 4

This essay recognizes that a sample of eight subjects is too small to use a *z*-test and accurately applies the small sample binomial test suggested in the question. The appropriate null and alternative hypotheses are identified in part (a). Part (b) is correct, but it would be strengthened by including a statement that for $p = 0.5$ and $n = 8$, both np and $n(1 - p)$ are too small to use a one-sample *z*-test. Binomial probabilities with $p = 0.5$ and $n = 8$ are accurately computed for the nine possible outcomes in part (c), and part (d) recognizes that a significance level of 0.05 cannot be achieved. This essay uses the table of probabilities computed in part (c) to obtain the correct *p*-value and provides an appropriate conclusion in the context of the juice-tasting study. Increasing the number of consumers randomly selected for the study is the improvement discussed in part (d); and the resulting increase in the power of the test to reject the null hypothesis of equal preference, when it is false, provides the requested statistical justification. The additional note that a larger sample size would justify the use of a *z*-test is also a valid point, but it is not as important as the increase in the power of the test (or the decrease in the standard error of the estimated proportion) provided by the increased sample size.

Sample: 6B
Score: 3

In parts (a) and (b) this essay correctly identifies the null and alternative hypotheses and shows that the sample size is too small to use a *z*-test. Binomial probabilities are correctly computed in part (c), but in part (d) it is not clear if the student is thinking about the sums of tail probabilities that correspond to possible type I error levels of 0.0079 and 0.0703. Realizing that the sample size is too small to apply a *z*-test from the previous parts of the problem, this essay refers to the table of binomial probabilities in part (c) but fails to present the appropriate *p*-value. The essay only displays the probability that exactly two consumers would prefer Citrus Fresh instead of the sum of the appropriate tail probabilities. An appropriate conclusion is reached, however, in the context of the problem. Part (d) provides statistical justification for increasing the sample size by addressing the issue of reducing the standard error of the estimated proportion and decreasing confidence interval width. It also addresses the issue raised in part (b) of obtaining a large enough sample size to use a large sample *z*-test. Since $p = 0.5$ for the null hypothesis, a sample size of $n = 20$ is needed to satisfy the common rule that both np and $n(1 - p)$ are larger than 10.

Sample: 6C
Score: 2

This essay correctly identifies the appropriate null and alternative hypotheses in part (a) but inappropriately suggests a two-sample *z*-test in part (b). Binomial probabilities are correctly computed in part (c), but an incorrect answer with no explanation is given in part (d). Although no work is shown, the value of an inappropriate two-sample *z*-test is presented in part (e) along with the *p*-value for that test. The conclusion is incorrect for the given data, but it is consistent with the *p*-value for the inappropriate two-sample *z*-test, and it is stated in the context of the juice-tasting study. Although it could be communicated better, the response to part (f) addresses the concept that increasing the sample size will decrease the standard deviation of the estimated proportion and lead to more precise inference about the proportion of the population that prefers Citrus Fresh.