AP® Statistics
2002 Sample Student Responses

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 of your results and explanation.

6. A survey given to a random sample of students at a university included a question about which of two well-known comedy shows, S or F, students preferred. The students were asked the question, "Do you prefer S or F ?" The responses are shown below.

Preference		
S	F	Total
185	139	324

(a) Based on the results of this survey, construct and interpret a 95% confidence interval for the proportion of students in the population who would respond S to the question, "Do you prefer S or F ?"

Assumptions (1-Proportion Z-Interval): $\hat{p} = .5710$

-SRS as stated in the problem ✓

- Population size of university is not known to be greater than 3240 (324·10) ✗

- $n\hat{p} > 10$ $- n(1-\hat{p}) > 10$

324(.5710) > 10   324(.4290) > 10

185 > 10 ✓   139 > 10

CI: $\hat{p} \pm Z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = .5710 \pm 1.96 \sqrt{\frac{.5710(.4290)}{324}} = \boxed{\langle .5171, .6249 \rangle}$

I am 95% confident that the proportion of students at this university who wou respond S to the question, "Do you prefer S. or F" is between .5171 and .6249. However, one assumption is not met, so the interval may be invalid.

(b) What is the meaning of "95% confidence" in part (a) ?

The meaning of 95% confidence is that if all possible confidence intervals ar constructed from all possible samples of 324, 95% of these intervals would capture the true proportion of people at the university who would answer S to the question asked.

GO ON TO THE NEXT PAGE.

(c) In a follow-up survey, a separate group of randomly selected students was asked "Do you prefer F or S ?" The responses are shown below.

Preference		
S	F	Total
68	88	156

Based on these two surveys, is there evidence that the stated preference depends on the order in which the comedy shows were listed in the survey question? Justify your answer.

Assumptions (2-Proportion Z-Test):

— 2 SRSs conducted as stated in problem ✓

— As stated in part (a), the population is not known to be greater than 3240 or 1560, so assumption is not met. ✗

— $\hat{p}_{pooled} = \frac{253}{480} = .52708$ $n_1 \hat{p}_{pooled} > 5$ $324(.52708) > 5$ $170 > 5$ ✓

$n_1(1-\hat{p}_{pooled}) > 5$ $324(.47292) > 5$ $153 > 5$ ✓

$n_2 \hat{p}_{pooled} > 5$ $156(.52708) > 5$ $82 > 5$ ✓ $n_2(1-\hat{p}_{pooled}) > 5$ $156(.47292) > 5$ $73 > 5$ ✓

$H_0: p_{S \, or \, F} = p_{F \, or \, S}$

$H_a: p_{S \, or \, F} \neq p_{F \, or \, S}$

$\hat{p}_1 - \hat{p}_2 = .1351$

-.1351   0   .1351

$\sigma_{\hat{p}} = \sqrt{\hat{p}_{pooled}(1-\hat{p}_{pooled})} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$

$= \sqrt{.52708(.47292)} \sqrt{\frac{1}{324} + \frac{1}{156}}$

$= .04866$

$z = \frac{\hat{p} - p}{\sigma_{\hat{p}}} = \frac{.1351 - 0}{.04866} = 2.776$ p-value = .0055
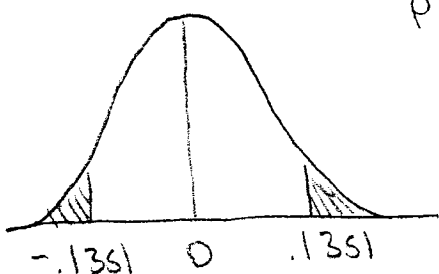
Reject H_0

If I assume the stated preference does not depend on the order in which the comedy shows were listed, the probability of seeing my results is .0055, which is low. Therefore, the evidence suggests that the stated preference depends on the order in which the comedy shows were listed. However, all of the assumptions had to be met to make this test valid.

GO ON TO THE NEXT PAGE.

(d) Suppose the test in part (c) indicates that the order in which the shows were listed does make a difference.

Is the pooled value $\dfrac{185 + 68}{324 + 156} = 0.527$ a reasonable estimate for the proportion of students at the university who would respond S ? If so, justify your answer. If not, what would be a more reasonable estimate? Explain why.

This is not a reasonable estimate for the proportion who would respond S. Since more people were asked in the order "S or F", these people had a greater impact on the pooled value than did those who were asked "F or S".

In order to get a more reasonable estimate, I will take the mean of the sample proportions because they would each have the same impact on the mean.

$$\text{Estimate} = \frac{\hat{P}_{S \text{ or } F} + \hat{P}_{F \text{ or } S}}{2} = \frac{\frac{185}{324} + \frac{68}{156}}{2} = \boxed{.5034}$$

GO ON TO THE NEXT PAGE.

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 of your results and explanation.

6. A survey given to a random sample of students at a university included a question about which of two well-known comedy shows, S or F, students preferred. The students were asked the question, "Do you prefer S or F ?" The responses are shown below.

Preference		
S	F	Total
185	139	324

(a) Based on the results of this survey, construct and interpret a 95% confidence interval for the proportion of students in the population who would respond S to the question, "Do you prefer S or F ?"

$= \dfrac{185}{324} = .57$

$\hat{p} =$ prop. of students responding "s" in the survey

☑ representative sample — yes b/c a random sample of students

☑ assume $324 < 10\%$ all of the students at the university

☑ $n\hat{p} = 185$ ⎫ more than 10
$n\hat{q} = 139$ ⎭ successes/ failures

conditions ok → I can use a 1 - proportion z - interval.

interval: $\hat{p} \pm (z^*)(SE(\hat{p}))$

$.57 \pm (1.96) \left(\sqrt{\dfrac{(.57)(.43)}{324}} \right)$

$[.517, .625] \rightarrow$ interval = .517 to .625

(b) What is the meaning of "95% confidence" in part (a) ?

I am 95% confident that the true proportion of students in the population who would respond "s" to the question is between .517 and .625

(c) In a follow-up survey, a separate group of randomly selected students was asked "Do you prefer F or S ?" The responses are shown below.

Preference		
S	F	Total
68	88	156

Based on these two surveys, is there evidence that the stated preference depends on the order in which the comedy shows were listed in the survey question? Justify your answer.

P_1 = prop. of students responding "S" when S is stated first.

P_2 = prop. of students responding "S" when F is stated first.

$H_0: P_1 = P_2$

$H_a: P_1 \neq P_2$

$\hat{P_1} = \frac{185}{324} = .57$ $\hat{P_2} = \frac{68}{156} = .44$

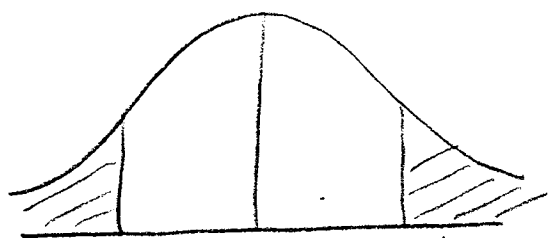$\hat{P}_P = \frac{185 + 68}{324 + 156} = \frac{253}{480} = .53$

☑ 2 representative samples — yes b/c both were randomly selected

☑ assume 324 < 10% all students at the university

assume 156 < 10% all students at the university

☑ $n\hat{P}_P = 170.775$
$n\hat{q}_P = 153.225$
$n\hat{P}_P = 82.225$
$n\hat{q}_P = 73.775$

both groups have more than 10 successes / failures

conditions ok → I can use a 2-proportion z-test for difference of proportions

$z = -2.78$

0 $\quad \hat{P_1} - \hat{P_2}$
$= .135$

$z = 2.78$

$P = 2 p(z > 2.78) = .0055$

Since $p = .0055$, which is far less than any reasonable level of significa[nce] I reject H_0. There is evidence of a difference in proportio[n] of students responding "S" depending on whether S or F is stated first, and there is evidence that a larger proportion of students respond

(d) Suppose the test in part (c) indicates that the order in which the shows were listed does make a difference. Is the pooled value $\frac{185 + 68}{324 + 156} = 0.527$ a reasonable estimate for the proportion of students at the university who would respond S ? If so, justify your answer. If not, what would be a more reasonable estimate? Explain why.

No, the pooled value is not a reasonable estimate when the order matters, because the pooled value assumes that the 2 proportions for students responding "S" are the same. ~~A more reasonable estimate would be the observed difference, $\hat{p_1} - \hat{p_2} = .135$ because this observed difference shows a significant difference in proportion based on order in which shows were listed.~~

A more reasonable estimate would be the average of the two observed proportions, because this shows half of each order.

GO ON TO THE NEXT PAGE.