Unit 6 Progress Check: FRQ

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

A recent survey collected information on television viewing habits from a random sample of 1,000 people in the United States. Of those sampled, 37 percent indicated that their favorite sport to watch on television was American football.

(a) Construct and interpret a 95 percent confidence interval for the proportion of all people in the United States who would indicate that their favorite sport to watch on television is American football.

(b) Based on your answer to part (a), is it reasonable to believe that 33 percent is the actual percent of people in the United States whose favorite sport to watch on television is American football? Justify your answer.

4-part Inference scoring

The primary goals of this question are to assess a student's ability to (1) construct and interpret a confidence interval for a population proportion; and (2) use a confidence interval to assess the plausibility of a particular claim about the value of a parameter.

Scoring

Part (a) has three scoring steps. Those scoring steps 1, 2, 3 and part (b) are each scored as essentially correct (E), partially correct (P), or incorrect (I). So, part (a) is worth 3 Es and part (b) is worth 1 E.

Each essentially correct (E) part counts as 1 point.

Each partially correct (P) part counts as ½ point.

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

Reasons to score up:

· *All notation is correct and clearly marked*

· *All explanations are clear*

· *No wrong information is included that was not part of the scoring (for example, saying sample size must be greater than 30 when that has nothing to do with the problem)*

· *No minor calculation errors are made, if they are not part of the scoring*

· *Interpretation parts are especially strong*

Reasons to score down:

· *Notation is not wrong, but is spotty and not clearly marked*

· *Explanations are not wrong, but are hard to follow*

Unit 6 Progress Check: FRQ

· *Wrong or extraneous information is included but not part of scoring*

· *Minor calculation errors that are not part of the scoring are made*

Interpretation parts are scored an E but are considered a weak E

0	1	2	3	4 ✓

Scoring steps 1, 2, 3 and part (b) sum to 4 points

OR

Scoring steps 1, 2, 3 and part (b) sum to $3\frac{1}{2}$ points AND a holistic approach is used to decide to score up

- ☐ **Scoring Step 1 - part (a)** Essentially correct
- ☐ **Scoring Step 1 - part (a)** Partially correct
- ☐ **Scoring Step 1 - part (a)** Incorrect
- ☐ **Scoring Step 2 - part (a)** Essentially correct
- ☐ **Scoring Step 2 - part (a)** Partially correct
- ☐ **Scoring Step 2 - part (a)** Incorrect
- ☐ **Scoring Step 3 - part (a)** Essentially correct
- ☐ **Scoring Step 3 - part (a)** Partially correct
- ☐ **Scoring Step 3 - part (a)** Incorrect
- ☐ **Part (b)** Essentially correct
- ☐ **Part (b)** Partially correct
- ☐ **Part (b)** Incorrect

Solution

Part (a)

Scoring step 1: Identify the correct procedure with conditions.

The appropriate procedure is the one-sample *z*-interval for a population proportion. The sample proportion is $\hat{p} = 0.37.$

Conditions:

· Data were collected from a random sample as stated.

· Sample size is large enough : $\begin{aligned} n\hat{p} &= 1,000(0.37) \geq 10 \\ n(1-\hat{p}) &= 1,000(0.63) \geq 10 \end{aligned}$

Unit 6 Progress Check: FRQ

The sampling method was probably conducted without replacement. It is reasonable to assume that the population of people in the United States was greater than 10 times the sample size of 1,000.

Scoring

Scoring step 1 is scored as follows.

Essentially correct (E) if the response satisfies the following three components

· The correct interval is identified, either by name or formula.

· The random sampling condition is checked.

· The sample size conditions are checked.

Partially correct (P) if the response satisfies only two of the three components

Incorrect (I) if the response does not meet the criteria for E or P.

Note: Checking the independence condition that $n \leq 10\%N$ is not necessary to score an E for scoring step 1, but this should be considered a plus for the purposes of holistic scoring.

Solution

Part (a)

Scoring step 2: Construct the interval.

For 95% confidence, the correct *z*-value is 1.96.

The interval is $0.37 \pm 1.96\sqrt{\frac{(0.37)(0.63)}{1{,}000}} = 0.37 \pm 0.03$, or $(0.34, 0.40)$.

Scoring

Scoring step 2 is scored as follows.

Essentially correct (E) if the response calculates the correct interval with work.

Partially correct (P) if the response calculates the correct interval with no work;

OR

if the response gives an interval with a calculation error or with the wrong *z*-value.

Incorrect (I) if the response does not meet the criteria for E or P.

Solution

Part (a)

Scoring step 3: Interpret the interval.

Unit 6 Progress Check: FRQ

We are 95% confident that the proportion of all people in the United States who would indicate that American football is their favorite sport to watch on television is between 0.34 and 0.40.

OR

We are 95% confident that the percent of all people in the United States who would indicate that American football is their favorite sport to watch on television is between 34% and 40%.

Scoring

Scoring step 3 is scored as follows.

Essentially correct (E) if the response satisfies the following three components.

· A reasonable interpretation in context and details about the population the interval represents.

· The interpretation is clear that the interval estimates the population proportion.

· The interpretation is given with 95% confidence.

Partially correct (P) if the response includes only two of the three components.

Incorrect (I) if the response does not meet the criteria for E or P.

Solution

Part (b)

It is not reasonable to conclude that the population percent is 33% because 0.33 is not included in the interval.

Scoring

Part (b) is scored as follows.

Essentially correct (E) if the response satisfies the following two components.

· A correct conclusion that is consistent with the interval calculated in part (a).

· A correct justification.

Note: The justification must reference the interval.

Partially correct (P) if the response satisfies only one of the two components.

Incorrect (I) if the response does not meet the criteria for E or P.

Unit 6 Progress Check: FRQ

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

A fair die, with its faces numbered from 1 to 6, is one in which each number is equally likely to land face up when the die is rolled. On a fair die, the probability that the number 6 will land face up is $\frac{1}{6}$. A group of students wanted to investigate a claim about manipulating a fair die so that it favors one outcome. The claim states that if a fair die is put into an oven and baked at 200°F for 10 minutes, the inside of the die will begin to melt. When the die cools, the inside will be solid again, but with more weight toward the bottom. This shift in weight will cause the face that was up when the die cooled to land up more often that the other faces.

The students obtained a fair die and baked it according to the preceding directions. The die cooled with the number 6 face up. After the die cooled, they rolled the die 200 times, and the number 6 landed face up 43 times. Let p represent the population proportion of times the number 6 will land face up on the baked die if the die could be rolled an infinite number of times.

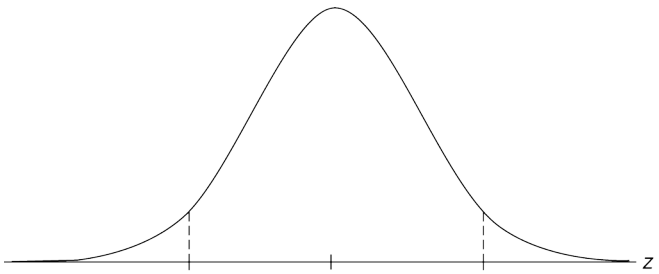
(a) Clarke, one of the students, constructed a 95 percent confidence interval for p as 0.215 ± 0.057. Does the interval provide convincing statistical evidence that the number 6 will land face up more often on the baked die than on a fair die? Explain your reasoning.

(b) Aurelia, another student, suggested they conduct a significance test to investigate the claim. She tested the hypotheses $\text{H}_0 : p = \frac{1}{6}$ versus $\text{H}_\text{a} : p > \frac{1}{6}$ at the significance level of $\alpha = 0.05$. She obtained a test statistic of 1.83 with a p-value of 0.033. Do the results of the significance test agree with the results of Clarke's confidence interval in part (a)? Explain your reasoning.

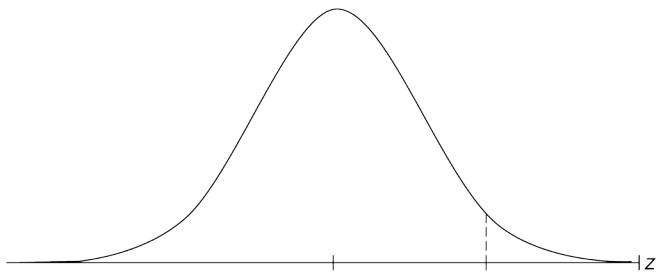
(c) Two standard normal curves are shown below, one for the confidence interval calculated in part (a) and one for the significance test conducted in part (b).

(i) For the confidence interval curve, label the critical values for the 95% confidence level and shade the area that represents values in the outer 5%.

(ii) For the significance test curve, label the critical value for the 5% significance level and shade the area representing the values of z that would lead to a rejection of the null hypothesis in part (b).



Confidence Interval



Hypothesis Test

(d) Joachim, a third student, noted that the confidence interval in part (a) gives plausible values of the parameter as an interval between two values. He suggested that they develop a one-sided confidence interval because they were only concerned with whether the number 6 was landing face up more often than expected, not less often. The one

Unit 6 Progress Check: FRQ

sided-interval will determine a value L such that all plausible values of p are greater than L. The formula for L is $L = \hat{p} + z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

(i) Determine the values of z^* needed to create the one-sided 95 percent confidence interval. Then calculate the value of L.

(ii) Do the results of Joachim's one-sided confidence interval agree with results of Aurelia's significance test in part (b)? Explain your reasoning.

Part A, B, C, and D

The primary goals of this question are to assess a student's ability to (1) interpret a 95% confidence interval, (2) explain the relationship between the confidence interval and an alpha level, (3) find a critical value that corresponds to a one-sided confidence interval, and (4) recognize that the one-sided confidence interval has the same results as a one-sided significance test.

Each essentially correct (E) part counts as 1 point.

Each partially correct (P) part counts as ½ point.

Scoring

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

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

Reasons to score up:

· All notation is correct and clearly marked

· All explanations are clear

· No wrong information is included that was not part of the scoring (for example, saying sample size must be greater than 30 when that has nothing to do with the problem)

· No minor calculation errors are made, if they are not part of the scoring

· Interpretation parts are especially strong

Reasons to score down:

· Notation is not wrong, but is spotty and not clearly marked

· Explanations are not wrong, but are hard to follow

· Wrong or extraneous information is included but not part of scoring

· Minor calculation errors that are not part of the scoring are made

Interpretation parts are scored an E but are considered a weak E

Unit 6 Progress Check: FRQ

0	1	2	3	4 ✓

Parts (a) through (d) sum to 4 points

OR

Parts (a) through (d) sum to $3\frac{1}{2}$ points AND a holistic approach is used to decide to score up

- ☐ Part (a) essentially correct
- ☐ Part (a) partially correct
- ☐ Part (a) incorrect
- ☐ Part (b) essentially correct
- ☐ Part (b) partially correct
- ☐ Part (b) incorrect
- ☐ Part (c) essentially correct
- ☐ Part (c) partially correct
- ☐ Part (c) incorrect
- ☐ Part (d) essentially correct
- ☐ Part (d) partially correct
- ☐ Part (d) incorrect

Solution

Part (a)

We can be 95% confident that the proportion of times the 6 will land face up on the baked die is between 0.158 and 0.272. Because $\frac{1}{6}$ (or ≈ 0.1667) is included in the interval, we do not have convincing statistical evidence to support the claim that the 6 will land face up more often than on a fair die.

Scoring

Part (a) is scored as follows:

Essentially correct (E) if the response satisfies the following three components.

- Correct calculation of the endpoints of the interval AND recognition that $\frac{1}{6}$ is contained in the interval.

¨ States the interval does not support the claim.

¨ Response refers to the parameter (the proportion of times 6 will land face up).

Partially correct (P) if the student satisfies only two of the three components.

Unit 6 Progress Check: FRQ

Incorrect (I) if the response does not meet the criteria for E or P.

Note: If a computation mistake is made in calculating the endpoints of the interval causing $\frac{1}{6}$ to fall outside the interval and the claim to be supported, the response is scored P if the reasoning is correct and in context.

Solution

Part (b)

At the significance level of 0.05, the null hypothesis is rejected, because $0.033 < 0.05$. There is convincing statistical evidence that the proportion of times 6 will land face up in the baked die is greater than $\frac{1}{6}$. This does not agree with the confidence interval in part (a) which suggested that $\frac{1}{6}$ was a plausible value for the proportion of times 6 will land face up in the baked die.

Scoring

Part (b) is scored as follows:

Essentially correct (E) if the response satisfies the following four components.

¨ A statement that the null hypothesis will be rejected or that there is convincing evidence for the alternative hypothesis.

¨ Justification for the decision by stating that the *p*-value is less than α .

¨ Reference to the parameter (the proportion of times 6 will land face up) in context.

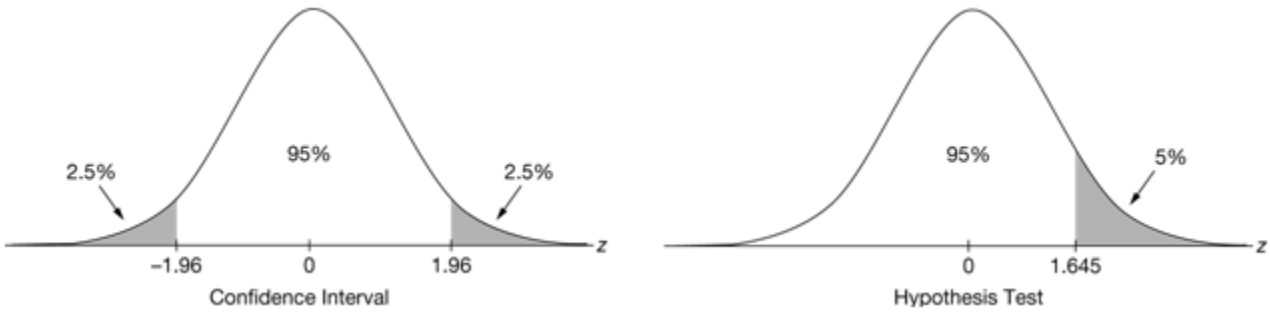
¨ States that results of the significance test do not agree with the results of the confidence interval in part (a).

Partially correct (P) if the response satisfies only two or three of the four components.

Incorrect (I) if the response does not meet the criteria for E or P.

Solution

Part (c)



Scoring

Part (c) is scored as follows:

Essentially correct (E) if the response satisfies the following four components.

Unit 6 Progress Check: FRQ

¨ For part (c-i), the values -1.96 and 1.96 are labeled on the confidence interval curve.

¨ For part (c-i), the confidence interval curve has shading in both tails.

¨ For part (c-ii), the value 1.645 (or 1.64 or 1.65) is labeled on the significance test curve.

¨ For part (c-ii) the significance test curve has shading in the right tail.

Partially correct (P) if the response satisfies only two or three of the four components.

Incorrect (I) if the response does not meet the criteria for E or P.

Note:

· Using a t-distribution to find the critical values does not satisfy components 1 or 3.

· Clearly labeling the area (e.g., 0.025, 0.05) in both curves should be considered a plus for the purposes of holistic scoring.

Solution

Part (d)

Because we want a 95% interval with no upper limit, the critical value z^* is the boundary that separates the smallest 5% from the largest 95% in a standard normal distribution. This value is $z^* = -1.645$. Therefore, the value of L is $\frac{43}{200} + (-1.645)\sqrt{\frac{\left(\frac{43}{200}\right)\left(\frac{157}{200}\right)}{200}} \approx 0.1672.$

According to the one-sided interval, we can be 95% confident that the proportion of times that 6 will land face up is greater than 0.1672. Because the value of $\frac{1}{6}$ is not contained in the interval, the results agree with the significance test in part (b). The one-sided confidence interval is equivalent to a one-sided significance test.

Scoring

Part (d) is scored as follows:

Essentially correct (E) if the response satisfies the following two components.

¨ In part (d-i), states the correct value for z^* $(-1.645,\ -1.64,\ -1.65)$

¨ In part (d-i), states the correct value for L.

¨ In part (d-ii), makes a correct statement about the equivalence of the one-sided confidence interval and the one-sided significance test that is consistent with the value of L calculated in part (d-i)

Partially correct (P) if the response satisfies only two of the three components.

Incorrect (I) if the response does not meet the criteria for E or P.

Note:

· A response that uses a critical value of $z^* = -1.96$ should be scored incorrect.

Unit 6 Progress Check: FRQ

· If a response clearly indicates that the sample proportion $\frac{1}{6}$ is rounded to 0.17, component 3 can still be satisfied if the response correctly compares 0.17 to the value of L calculated in part (d-i)