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.

The manufacturer of a certain type of new cell phone battery claims that the average life span of the batteries is 500 charges; that is, the battery can be charged at least 500 times before failing. To investigate the claim, a consumer group will select a random sample of cell phones with the new battery and use the phones through 500 charges of the battery. The proportion of batteries that fail to last through 500 charges will be recorded. The results will be used to construct a 95 percent confidence interval to estimate the proportion of all such batteries that fail to last through 500 charges.

(a) Explain in context what it means to be 95 percent confident.

(b) Suppose the consumer group conducts its investigation with a random sample of 5 cell phones with the new battery, and 1 battery out of the 5 fails to last through 500 charges. Verify all conditions for inference for a 95 percent confidence interval for a population proportion. Indicate whether any condition has not been met. Do not construct the interval.

In some studies, large sample sizes are not feasible because of the time and expense involved. However, small sample sizes can create issues with inference procedures. A simulation can be used to investigate what might happen with intervals constructed from small sample sizes. Consider a population in which 30 percent of the population displays a certain characteristic. For each trial of the simulation, 5 observations are selected from the population and the sample proportion $\hat{p}$ is calculated, where $\hat{p}$ represents the proportion in the sample that display the characteristic. The following table shows the frequency distribution of $\hat{p}$ in the $1{,}000$ trials. Also shown are the upper and lower endpoints of a 95 percent confidence created from the value of $\hat{p}$, using the formula $\hat{p} \pm 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$. For example, the sample proportion of 0.4 occurred 309 times in the $1{,}000$ trials and produced a confidence interval of $(-0.029, 0.829)$.

$\hat{p}$	Frequency	Lower Endpoint	Upper Endpoint
0	168	0	0
0.2	360	-0.151	0.551
0.4	309	-0.029	0.829
0.6	133	0.171	1.029
0.8	28	0.449	1.151
1.0	2	1	1

(c) Based on the simulation, what proportion of the 95 percent confidence intervals capture the population proportion of 0.3 ? Explain how you determined your answer.

(d) For small sample sizes, an alternate method of constructing a confidence interval is available. To implement this method, first, include an additional 4 observations in the sample, 2 of which are successes and 2 of which are failures. Second, calculate a new sample proportion $\hat{p}_{\text{new}}$ for the new sample. Finally, construct a confidence interval using $\hat{p}_{\text{new}}$, with the formula $\hat{p}_{\text{new}} \pm 1.96\sqrt{\frac{\hat{p}_{\text{new}}(1-\hat{p}_{\text{new}})}{n+4}}$.

(i) For the cell phone batteries, consider a sample of 5 in which 1 battery fails to last through 500 charges. Using the alternate method described, what is the value of $\hat{p}_{\text{new}}$? Show your work.

The following table shows the results of the original simulation with revised 95 percent confidence intervals constructed with the alternate method.

Original $\hat{p}$	Frequency	Revised Lower Endpoint	Revised Upper Endpoint
0	168	-0.049	0.494
0.2	360	0.025	0.641
0.4	309	0.120	0.769
0.6	133	0.231	0.880
0.8	28	0.359	0.975
1.0	2	0.506	1.049

(ii) Based on the results of the simulation, is the alternate method better than the traditional method described in part (b) to construct a 95 percent confidence interval with a small sample size? 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 the confidence level for a method of calculating a confidence interval; (2) state and check the conditions for constructing a confidence interval for a population proportion; (3) use the results of a simulation to estimate how often a method of calculating a confidence interval will contain the population proportion; and (4) assess the quality of an alternative method of calculating a confidence interval for a population proportion.

Scoring

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

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*

· *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 ✓

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)

If the consumer group selects many random samples of the same size of cell phone batteries and constructs a 95% confidence interval using each sample, then about 95% of these intervals will capture the true proportion of all batteries of this type that fail to last through 500 charges.

Scoring

Part (a) is scored as follows:

Essentially correct (E) if the response provides a correct interpretation of the confidence level in the context of the study. A correct interpretation satisfies the following three components.

· Mentions repeated sampling or "in the long run" or "using this method."

· Mentions that 95% or about 95% of the corresponding intervals will capture the population proportion.

· Describes the population proportion in context.

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

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

Note: If the interpretation is about the confidence interval, not the confidence level, the response would only satisfy component 3 and earn an I. For example, "We are 95% confident that the interval will capture the true proportion of all batteries that fail to last through 500 charges" is not an interpretation of confidence level.

Solution

Part (b)

One condition is that the consumer organization selected a random sample of cell phone batteries of this type. This condition is met. Another condition is that the sample size is less than 10% of the population size. This seems reasonable as there are very likely more than $10(5) = 50$ cell phone batteries of this type in the population. Another condition is that the $n\hat{p}$ and $n(1 - \hat{p})$ both must be at least 10. This condition is not met because both $n\hat{p} = 1$ and $n(1 - \hat{p}) = 4$ are less than 10.

Scoring

Part (b) is scored as follows:

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

· Verifies the random sample condition.

· Verifies the 10% condition.

· Provides numerical evidence that either $n\hat{p}$ or $n(1 - \hat{p})$ is not met.

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.

Notes:

· To satisfy component 3, the number of successes and/or number of failures must be compared to a boundary value such as 5 or 10.

· If an incorrect condition is included, such as $n < 30$ or that treatments are randomly assigned, a score of E is reduced to P. (A score of P is not reduced to I.)

· Attempts to define the parameter, calculate the interval, or interpret the interval can be ignored.

Solution

Part (c)

To capture the true proportion ($p = 0.3$), the lower boundary must be less than 0.30 and the upper boundary must be greater than 0.30. This occurs for $\hat{p}$ values of 0.2 because the interval is $(-0.151,\ 0.551)$, 0.4 because the interval is $(-0.029,\ 0.829)$, and 0.6 because the interval is $(0.171,\ 1.029)$. Because these $\hat{p}$ values occurred in $360 + 309 + 133 = 802$ of the 1,000 simulated trials, the capture rate is 80.2%.

$\hat{p}$	Frequency	Lower Endpoint	Upper Endpoint	Does the interval include 30%?
0	168	0	0	No
0.2	360	−0.151	0.551	Yes
0.4	309	−0.029	0.829	Yes
0.6	133	0.171	1.029	Yes
0.8	28	0.449	1.151	No
1.0	2	1	1	No

Scoring

Part (c) is scored as follows:

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

· Explains that the interval captures the true proportion if it includes 0.30.

· Provides an answer of $802/1000$ or 0.802 or 80.2%.

· Includes justification by referring to the appropriate values of $\hat{p}$ (0.2, 0.4, 0.6) explicitly or with clear reference to the table listing the actual intervals, or listing the values 360, 309, and 133.

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.

Notes:

· A response of $3/6 = 0.50$ (that is, three of the six possible intervals capture 0.30) is scored P if components 1 and 3 are satisfied.

· A response that says something like “less than 95% because the conditions aren’t met” is scored I because there is no reference to the results of the simulation.

· If the response provides a response of 802, not $802/1000$, component 2 is not satisfied. However, this could be viewed as a positive for holistic scoring.

Solution

Part (d)

(i) $\hat{p}_{\text{new}} = \frac{1+2}{5+4} = \frac{3}{9} = 0.333$

(ii) Because $168 + 360 + 309 + 133 = 970$ of the 1,000 simulated samples produced intervals that captured 0.30, the percentage for the alternate method is 97.0%. Because the percentage captured for the alternate method is much closer to 95% than the percentage captured for the traditional method (80.2%), the alternate method is better.

$\hat{p}$	Frequency	Revised Lower Endpoint	Revised Upper Endpoint	Does the interval include 30%?
0	168	*−0.049*	0.494	Yes
0.2	360	0.025	0.641	Yes
0.4	309	0.120	0.769	Yes
0.6	133	0.231	0.880	Yes
0.8	28	0.359	0.975	No
1.0	2	0.506	1.049	No

Scoring

Part (d) is scored as follows:

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

· In part (d-i), calculates the new value of $\hat{p}$ with work shown.

· In part (d-ii), states that the alternative method works better because a greater proportion of intervals contain the 30% ($97\% > 80.2\%$) and that value is closer to 95%.

· In part (d-ii), provides numerical justification for the decision (for example, by calculating the value of 97.0%)

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: If the response in part (c) was $3/6$, a response in part (d) of 4/6 satisfies component 3.

2. From a random sample of potential voters in an upcoming election, 47% indicated they intended to vote for Candidate R. A 95 percent confidence interval was constructed from the sample, and the margin of error for the estimate was 5%.

Which of the following is the best interpretation of the interval?

(A) We are 95% confident that the proportion who intend to vote for Candidate R from the random sample is between 42% and 52%.

(B) We are 95% confident that the proportion who intend to vote for Candidate R from the population is between 42% and 52%. ✓

(C) We are 95% confident that the proportion who intend to vote for Candidate R from the random sample is 47%.

(D) We are 95% confident that the proportion who intend to vote for Candidate R from the population is 47%.

(E) We are confident that 95% of the population intend to vote for Candidate R.

Answer B

Correct. Adding and subtracting the margin of error (5%) from the point estimate (47%) gives an interval from 42% to 52%. The interval is a statement about how confident we are that the interval has captured the population parameter.

3. A random sample of residents in city J were surveyed about whether they supported raising taxes to increase bus service for the city. From the results, a 95 percent confidence interval was constructed to estimate the proportion of people in the city who support the increase. The interval was $(0.46, 0.52)$.

Based on the confidence interval, which of the following claims is supported?

(A) More than 90 percent of the residents support the increase.

(B) More than 60 percent of the residents support the increase.

(C) More than 40 percent of the residents support the increase. ✓

(D) Fewer than 10 percent of the residents support the increase.

(E) Fewer than 25 percent of the residents support the increase.

Answer C

Correct. The claim is supported by the confidence interval. The interval represents plausible values for the population proportion of residents who support the increase and all values in the confidence interval are over 40 percent.

4. A random sample of 1,175 people in a certain country were asked whether they thought climate change was a problem. The sample proportion of those who think climate change is a problem was calculated, and a 95 percent confidence interval was constructed as $(0.146, 0.214)$.

Which of the following is a correct interpretation of the interval?

(A) We are 95 percent confident that any sample of 1,175 people will produce a sample proportion between 0.146 and 0.214.

(B) We are 95 percent confident that the proportion of all people in the country who think climate change is a problem is between 0.146 and 0.214. ✓

(C) We are 95 percent confident that the proportion of people in the sample who think climate change is a problem is between 0.146 and 0.214.

(D) The probability that 95 percent of all people in the country who think climate change is a problem is between 0.146 and 0.214.

(E) The probability is 0.95 that the proportion of all people in the country who think climate change is a problem is between 0.146 and 0.214.

Answer B

Correct. The interval is a statement about how confident we are that the interval has captured the population parameter—the proportion of all people in the country who think climate change is a problem.

5. Elly and Drew work together to collect data to estimate the percentage of their classmates who own a particular brand of shoe. Using the same data, Elly will construct a 90 percent confidence interval and Drew will construct a 99 percent confidence interval. Which of the following statements is true?

(A) The midpoint of Elly's interval will be greater than the midpoint of Drew's interval.

(B) The midpoint of Elly's interval will be less than the midpoint of Drew's interval.

(C) The width of Elly's interval will be greater than the width of Drew's interval.

(D) The width of Elly's interval will be less than the width of Drew's interval. ✓

(E) The width of Elly's interval will be equal to the width of Drew's interval.

Answer D

Correct. For the same sample, as the confidence level increases, the width of the interval increases. Elly's confidence level (90%) is less than Drew's (99%), so the width of her interval will be less than Drew's.

6. Consider a 90 percent confidence interval to estimate a population proportion that is constructed from a sample proportion of 66 percent. If the width of the interval is 10 percent, what is the margin of error?

(A) 2.5 percent

(B) 5 percent ✓

(C) 10 percent

(D) 20 percent

(E) 45 percent

Answer B

Correct. The margin of error is one-half of the total width of the confidence interval, and one-half of 10 percent is 5 percent.

7. Based on findings from a recent study on women's health, researchers created a 90 percent confidence interval of $(0.42, 0.48)$ to estimate the percent of all women who do not find time to focus on their own health. Based on the confidence interval, which of the following claims is not supported?

(A) Less than half of all women do not find time to focus on their own health.

(B) More than 40 percent of all women do not find time to focus on their own health.

(C) Approximately 45 percent of all women do not find time to focus on their own health.

(D) More than 45 percent of all women do not find time to focus on their own health. ✓

(E) More than 25 percent of all women do not find time to focus on their own health.

Answer D

Correct. The claim is not supported by the confidence interval. The percentages contained in the interval are from 42 percent to 48 percent. Percentages less than 45 percent are also plausible values for the population parameter.

8. Consider a 90 percent confidence interval for a population proportion p. Which of the following is a correct interpretation of the confidence level 90 percent?

(A) There is approximately a 90 percent chance that p is contained in the interval.

(B) There is approximately a 90 percent chance that a randomly selected proportion $\hat{p}$ will be contained in the interval.

(C) Approximately 90 percent of all possible sample proportions $\hat{p}$ will be contained in the interval.

(D) In repeated samplings with the same sample size, approximately 90 percent of the intervals created will capture the population proportion p. ✓

(E) In repeated samplings with the same sample size, approximately 90 percent of the intervals created will capture the sample proportion $\hat{p}$.

Answer D

Correct. The confidence level of 90 percent reflects the percent of all possible intervals that will capture the population parameter p.

9. A recent study on the way that people talk indicated, with 95 percent confidence, that between 35 percent and 41 percent of all adults find the word "whatever" to be the most annoying word in conversation.

Based on the confidence interval, which of the following claims is supported?

(A) Less than 25 percent of all adults find the word "whatever" to be the most annoying word in conversation.

(B) More than 30 percent of all adults find the word "whatever" to be the most annoying word in conversation. ✓

(C) More than 45 percent of all adults find the word "whatever" to be the most annoying word in conversation.

(D) More than half of all adults find the word "whatever" to be the most annoying word in conversation.

(E) At least 95 percent of all adults find the word "whatever" to be the most annoying word in conversation.

Answer B

Correct. A claim that the actual percent is greater than 30 percent is supported by the confidence interval. The interval represents plausible values for the population proportion and all values contained in the interval are greater than 0.3.

10. Lila and Robert attend different high schools. They will estimate the population percentage of students at their respective schools who have seen a certain movie. Lila and Robert each select a random sample of students from their respective schools and use the data to create a 95 percent confidence interval. Lila's interval is $(0.30, 0.35)$, and Robert's interval is $(0.27, 0.34)$. Which of the following statements can be concluded from the intervals?

(A) Lila's sample size is most likely greater than Robert's sample size. ✓

(B) Robert's sample size is mostly likely greater than Lila's sample size.

(C) Lila and Robert will both find the same sample proportion of students who have seen the movie.

(D) Lila's interval has a greater degree of confidence than that of Robert.

(E) Robert's interval has a greater degree of confidence than that of Lila.

Answer A

Correct. Both Lila and Robert use the same level of confidence, but Lila's interval is narrower with a width of 5 percent $(0.35 - 0.30 = 0.05)$ as opposed to Robert's interval width of 7 percent $(0.34 - 0.27 = 0.07)$. Lila's $\hat{p}$ value is 0.325 (the midpoint of her interval), and Robert's $\hat{p}$ value is 0.305 (the midpoint of his interval), so the value of $\hat{p}(1 - \hat{p})$ in the calculation for the margin of error is very close for Lila's interval and Robert's interval. Therefore, the difference in the confidence interval width is most likely due to Lila's sample size being greater.