

Topics: Confidence Intervals

1. For each of the following statements, indicate whether it is True/False. If false, explain why.

I. The sample size of the survey should at least be a fixed percentage of the population size in order to produce representative results.

Ans: **False.** The representativeness of a survey sample depends on the selection method and its alignment with the population, not on fixed percentage of the population size. Proper sampling techniques ensure representative results regardless of the sample size in relation to the population. Sampling techniques such as random sampling or stratified sampling can help ensure representative results regardless of the sample size relative to the population size.

II. The sampling frame is a list of every item that appears in a survey sample, including those that did not respond to questions.

Ans: **False.** The sampling frame includes items in the population that have a chance of being selected, but it does not include non-respondents or those who did not participate in the survey. Non-respondents are different from the sampling frame as they are individuals who were selected to participate but did not provide a response.

III. Larger surveys convey a more accurate impression of the population than smaller surveys.

Ans: **True.** Larger surveys generally provide a more accurate representation of the population due to capturing more diversity, variability in the population and reducing sampling error. Other factors like sampling method, questionnaire design and data analysis techniques also impact accuracy.

2. *PC Magazine* asked all of its readers to participate in a survey of their satisfaction with different brands of electronics. In the 2004 survey, which was included in an issue of the magazine that year, more than 9000 readers rated the products on a scale from 1 to 10. The magazine reported that the average rating assigned by 225 readers to a Kodak compact digital camera was 7.5. For this product, identify the following:

A. The population

Ans: All readers of PC magazine who were asked to participate in the survey of their satisfaction with different brands of electronics.

B. The parameter of interest

Ans: The parameter of interest is the average rating assigned by all readers (the population) to a Kodak compact digital camera.

C. The sampling frame

Ans: It would be the list of all readers of PC Magazine who were asked to participate in the survey.

D. The sample size

Ans: The sample size mentioned in the question is 225 readers who rated the Kodak compact digital camera.

E. The sampling design

Ans: It can be assumed that the survey used a convenience sampling method, where all readers of PC Magazine were invited to participate.

F. Any potential sources of bias or other problems with the survey or sample

Ans: Potential sources of bias or problems: The survey may have non-response bias if only a specific group of readers responded, self-selection bias if participants had specific preferences, and sampling bias if the survey was not representative of all PC Magazine readers. Convenience sampling may introduce selection bias and limit the ability to generalize the results.

3. For each of the following statements, indicate whether it is True/False. If false, explain why.

I. If the 95% confidence interval for the average purchase of customers at a department store is \$50 to \$110, then \$100 is a plausible value for the population mean at this level of confidence.

Ans: True. If the 95% confidence interval for the average purchase at a department store is \$50 to \$110, cause \$100 falls within this range, it is a plausible value for the population mean at this level of confidence.

II. If the 95% confidence interval for the number of moviegoers who purchase concessions is 30% to 45%, this means that fewer than half of all movie goers purchase concessions.

Ans: False. The 95% confidence interval for the number of moviegoers who purchase concessions, ranging from 30% to 45%, it means 95% confident that the true proportion of moviegoers who purchase concessions falls within this range. This statement does not provide enough information to determine if fewer than half of all moviegoers purchase concessions, as the range includes values above and below 50%.

III. The 95% Confidence-Interval for μ only applies if the sample data are nearly normally distributed.

Ans: False. It does not require the sample data to be nearly normally distributed. From central limit theorem, it states that with a sufficiently large sample size, the sampling distribution of the sample mean will be approximately normally distributed, regardless of the distribution of the population. However, for small sample sizes (typically less than 30), it is important to check for normality assumptions or consider alternative methods for constructing confidence intervals.

4. What are the chances that $\bar{X} > \mu$?

- A. $\frac{1}{4}$
- B. $\frac{1}{2}$
- C. $\frac{3}{4}$
- D. 1

Ans: The chances that a sample mean is greater than the population mean is 50%. The sample mean is an unbiased estimator, equally likely to be greater or less than the population mean. The central limit theorem supports this, stating that as the sample size increases, the distribution of sample means approximates a normal distribution. Consequently, with a large number of samples, approximately half of the sample means will exceed the population mean.

5. In January 2005, a company that monitors Internet traffic (WebSideStory) reported that its sampling revealed that the Mozilla Firefox browser launched in 2004 had grabbed a 4.6% share of the market.

I. If the sample were based on 2,000 users, could Microsoft conclude that Mozilla has a less than 5% share of the market?

Ans: Based on a sample of 2000 users, additional information is needed. The reported sample of 4.6% share does not provide enough context to make a definitive conclusion about the entire market. It would require a confidence interval or hypothesis test to assess the statistical significance of the difference.

II. WebSideStory claims that its sample includes all the daily Internet users. If that's the case, then can Microsoft conclude that Mozilla has a less than 5% share of the market?

Ans: If WebSideStory claims that its sample include all daily internet users, it would imply a census rather than a sample. In this case, if the reported 4.6% share is based on a comprehensive data collection of all internet users, Microsoft can reasonably conclude that Mozilla has a less than 5% share of the market. However, it is important to verify the accuracy and representativeness of the data collection process before drawing final conclusions.

6. A book publisher monitors the size of shipments of its textbooks to university bookstores. For a sample of texts used at various schools, the 95% confidence interval for the size of the shipment was 250 ± 45 books. Which, if any, of the following interpretations of this interval are correct?

A. All shipments are between 205 and 295 books.

Ans: Incorrect. The correct interpretation of the confidence interval is not that all shipments are between 205 and 295 books. It represents the range within which we are 95% confident that the population mean falls.

B. 95% of shipments are between 205 and 295 books.

Ans: Incorrect. The correct interpretation of the confidence interval is not that 95% of shipments are between 205 and 295 books. It represents the range within which we are 95% confident that the population mean falls, not individual shipments.

C. The procedure that produced this interval generates ranges that hold the population mean for 95% of samples.

Ans: Correct. This interpretation is correct. The confidence interval is obtained using a procedure that, on average, produces ranges that contain the population mean for 95% of samples. However, it does not guarantee that the specific interval from this sample captures the population mean.

D. If we get another sample, then we can be 95% sure that the mean of this second sample is between 205 and 295.

Ans: Incorrect. The correct interpretation of the confidence interval does not pertain to a second sample. It represents the range within which we are 95% confident that the population mean falls for the current sample.

E. We can be 95% confident that the range 160 to 340 holds the population mean.

Ans: Incorrect. The correct interpretation of the confidence interval is not that the range 160 to 340 holds the population mean. The provided confidence interval of 250 ± 45 books is the appropriate range within which we can be 95% confident that the population mean falls.

Therefore, the correct interpretations are C. The procedure that produced this interval generates ranges that hold the population mean for 95% of samples.

7. Which is shorter: a 95% z-interval or a 95% t-interval for μ if we know that $\sigma = s$?
- A. The z-interval is shorter
 - B. The t-interval is shorter
 - C. Both are equal
 - D. We cannot say

Ans: When the standard deviation(σ) is known, the z-interval is used, and it tends to be shorter than the t-interval. The z-interval is based on the standard normal distribution, which has a narrower distribution compared to the t-distribution used in the t-interval. Therefore, when σ is known, the z-interval provides a more precise estimate, resulting in a shorter interval compared to the t-interval.

So, A. The z-interval is shorter.

Questions 8 and 9 are based on the following: To prepare a report on the economy, analysts need to estimate the percentage of businesses that plan to hire additional employees in the next 60 days.

8. How many randomly selected employers (minimum number) must we contact in order to guarantee a margin of error of no more than 4% (at 95% confidence)?
- A. 600
 - B. 400
 - C. 550
 - D. 1000

Ans: To determine the minimum sample size required to guarantee a margin of error of no more than 4% at a 95% confidence level, we can use the formula

$$n = (z^2 * p * (1-p)) / e^2$$

Where,

n = sample size

z = z-score corresponding to the desired confidence level (given, 95%)

p = estimated proportion (not given, we can use 0.5 for a conservative estimate)

e = margin of error (4% or 0.04)

So,

$$n = (1.96^2 * 0.5 * (1-0.5)) / 0.04^2$$

$$n = 600.2499999999999 \sim 600$$

Therefore, the minimum number of randomly selected employers we must contact is 600.

So, A. 600 is correct.

9. Suppose we want the above margin of error to be based on a 98% confidence level. What sample size (minimum) must we now use?
- A. 1000
 - B. 757
 - C. 848
 - D. 543

Ans: If same margin or error but 98% confidence level, we need to find the z-score corresponding to find the z-score corresponding to a 98% confidence level. The z-score is approximately 2.33.

Using same formula as above, with the new z-score,

$$n = (2.33^2 * 0.5 * (1-0.5)) / 0.04^2$$
$$n = 848.265625 \sim 848$$

Therefore, the minimum sample size we must use is 848.

So, C. 848 is correct.