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: The statement is **False**. The sample size of a survey should not be determined solely based on a fixed percentage of the population size. The appropriate sample size depends on various factors, including the level of precision required, desired confidence level, variability in the population, and the research objectives. A larger sample size can lead to more precise estimates, but it's essential to use proper statistical methods to determine the sample size, considering the specific requirements of the study.

    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: The statement is **False**. The sampling frame is a list of all the items (elements) that make up the target population from which the survey sample is drawn. It should ideally include all members of the population. The sample comprises the elements selected from the sampling frame to participate in the survey.

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

    Ans: The statement is **True**. Large sample = Less Standard Deviation

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: The population in this context refers to all the readers of PC Magazine who participated in the survey, i.e, **more than 9000**.

    B. The parameter of interest

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

    C. The sampling frame

    Ans: All the readers of PC magazine.

    D. The sample size

    Ans: The sample size in the survey was **225 readers.**

E. The sampling design

Ans: The information provided does not explicitly state the sampling design used in the survey. It's possible that the sample was a simple random sample, where each reader had an equal chance of being selected, or it could have been a convenience sample, where the participants were chosen based on their availability or willingness to respond.

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

Ans: Non-response bias, Self-selection bias, Sample size, Response bias

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: The statement is True.

If the 95% confidence interval for the average purchase of customers at a department store is $50 to $110, it means that we are 95% confident that the true population mean lies within this range. Since $100 is within this interval, it is considered 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 moviegoers purchase concessions.

Ans: The statement is False.

We cannot conclude that fewer than half of all moviegoers purchase concessions, as 45% is more than half (50%).

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

Ans: The Statement is False.

Confidence interval is dependent on many factors and not just on μ.

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

A. ¼
B. ½
C. ¾
D. 1

Ans: Based on the typical use of the Central Limit Theorem for large sample sizes, the probability that X̄ > μ is 0.5
The correct option is B.

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: To determine if Microsoft can conclude that Mozilla has a less than 5% share of the market, we need to perform a hypothesis test.

The null hypothesis (H0) would be that Mozilla's share of the market is equal to or greater than 5% ($p \geq 0.05$), and the alternative hypothesis (Ha) would be that Mozilla's share of the market is less than 5% ($p < 0.05$).

Given that the sample size is 2,000 and the proportion of users using Mozilla Firefox is reported as 4.6%, we can calculate the standard error of the proportion:

$SE = sqrt((0.046 * (1 - 0.046)) / 2000) \approx 0.0091$

$Z = (0.046 - 0.05) / 0.0091 \approx -0.4396$

Using a standard normal distribution table, the critical Z-value for a one-tailed test approximately -1.645.

Since the calculated Z-score (-0.4396) is greater than the critical Z-value (-1.645), we fail to reject the null hypothesis.

In other words, Microsoft cannot conclude that Mozilla has a less than 5% share of the market based on this sample of 2,000 users.

   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 the sample truly includes all the daily Internet users, it would be a census rather than a sample. In this case, Microsoft can confidently conclude that Mozilla has a 4.6% share of the market because the data represents the entire population of daily Internet users, and the reported proportion is based on an actual count rather than a sample-based estimate. **Microsoft can conclude that Mozilla has a 4.6% share of the market in this scenario.**

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.

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

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

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

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

Ans: C. The procedure generates intervals 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 σ =s?

    A. The z-interval is shorter
    B. The t-interval is shorter
    C. Both are equal
    D. We cannot say

Ans: A is the right option.

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: A. is the right option.
n = (1.96^2 * 0.5 * (1 - 0.5)) / (0.04^2)
n = (3.8416 * 0.5 * 0.5) / 0.0016
 n = 0.9604 / 0.0016
 n ≈ 600.25
Rounding up to the nearest whole number, the minimum sample size required is 600.

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: C. is the right option.
n = (2.33^2 * 0.5 * (1 - 0.5)) / (0.04^2)
n = (5.4289 * 0.5 * 0.5) / 0.0016
n = 1.3572 / 0.0016
n ≈ 848.25
Rounding up to the nearest whole number, the minimum sample size required is 849.
Therefore, 848 is correct option