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.

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

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

    Ans)

I. False. The statement is false because the sample size of a survey does not necessarily need to be a fixed percentage of the population size to produce representative results. The representativeness of a sample depends on various factors such as the sampling method used, the diversity within the population, and the level of precision required. Sometimes, smaller samples can still provide accurate representations of the population if they are selected properly through random sampling methods.

II. False. The statement is false because the sampling frame is not necessarily a list of every item that appears in a survey sample, including those that did not respond to questions. The sampling frame is actually a list of all the elements or individuals in the population that have a chance of being selected for the sample. It should ideally cover the entire population of interest and be free from biases to ensure that the sample is representative.

III. True. Generally, larger surveys tend to convey a more accurate impression of the population than smaller surveys. This is because larger samples reduce the margin of error and increase the precision of the estimates. However, it's important to note that the relationship between sample size and accuracy is not linear. There's a point of diminishing returns where increasing the sample size may not significantly improve accuracy, especially if the sampling method is flawed or if there are biases present in the sample.

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
    B. The parameter of interest
    C. The sampling frame
    D. The sample size
    E. The sampling design

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

Ans)

A. The population: The population in this scenario would be all consumers who have purchased or have the potential to purchase Kodak compact digital cameras.
P=x/n=225/9000=0.025

B. The parameter of interest: The parameter of interest would be the average satisfaction rating of all consumers who have purchased or have the potential to purchase Kodak compact digital cameras.smple size,average,scale.

C. The sampling frame: The sampling frame would be the list of all PC Magazine readers who participated in the survey.9000

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

E. The sampling design: The sampling design seems to be simple random sampling, where each PC Magazine reader who participated in the survey had an equal chance of being selected to rate the Kodak compact digital camera.voluntary response

F. Potential sources of bias or other problems with the survey or sample:

- Selection bias: The survey relies on PC Magazine readers, who might not represent the entire population of consumers interested in Kodak compact digital cameras.
- Response bias: Readers who choose to participate in the survey might have different opinions compared to those who choose not to participate.
- Self-selection bias: Only readers who are interested or have strong opinions about Kodak compact digital cameras might have participated in the survey, skewing the results.
- Social desirability bias: Participants might give higher ratings to Kodak products because they perceive it as socially desirable to endorse well-known brands.

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.

   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.

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

   Ans)

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

II. False. A 95% confidence interval for the proportion of moviegoers who purchase concessions ranging from 30% to 45% does not necessarily mean that fewer than half of all moviegoers purchase concessions. It means that we are 95% confident that the true proportion of moviegoers who purchase concessions lies within this interval, but it doesn't give a precise indication of whether it's less than or greater than 50%.

III. False. The 95% confidence interval for the population mean (μ) does not necessarily require the sample data to be nearly normally distributed. The central limit theorem states that as the sample size increases, the sampling distribution of the sample mean tends to be normally distributed, regardless of the distribution of the population. Therefore, the 95% confidence interval for the population mean is valid even if the sample data are not exactly normally distributed, as long as the sample size is sufficiently large (usually n ≥ 30).

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

   A. ¼
   B. ½
   C. ¾
   D. 1

Ans) b
This is pure assumption. There is a 50% chance that the sample mean(x) is greter than the population mean(mu)

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)

- Sample proportion (p̂) = 4.6% = 0.046 (proportion of Firefox users in the sample)
- Sample size (n) = 2000

We can use the formula for a confidence interval for a proportion:

Confidence Interval=*p*^ ±*Z*×*np*^ (1−*p*^)

Where:

- Z is the Z-score corresponding to the desired confidence level. For a 95% confidence level, $Z \approx 1.96$.
- $p\text{^}$ is the sample proportion.
- n is the sample size.

Substituting the given values:

Confidence Interval=0.046±1.96×20000.046(1−0.046) Confidence Interval=0.046±1.96×2000 0.046×0.954

Confidence Interval=0.046±1.96×20000.043924

Confidence Interval=0.046±1.96×0.000021962

Confidence Interval=0.046±1.96×0.004688

Confidence Interval=0.046±0.00919

Confidence Interval=(0.03681,0.05519)

So, at a 95% confidence level, we can be confident that the true proportion of Firefox users in the population lies between 3.681% and 5.519%. Since this interval includes 5%, Microsoft cannot conclude that Mozilla has a less than 5% share of the market based on this sample

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) Sample proportion ($\hat{p}$) = 4.6% = 0.046 (proportion of Firefox users in the sample)
- Sample size (n) is not specified because it's a census.

With a census, we don't need to calculate confidence intervals since we have data for the entire population. If WebSideStory's claim is true and the sample includes all daily Internet users, then the proportion of Firefox users in the population is indeed 4.6%. Therefore, Microsoft still cannot conclude that Mozilla has a less than 5% share of the market.

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) the interval of (205,295) is for 95% confidence not for 100%.

B. 95% of shipments are between 205 and 295 books.
Ans) the interval doesnot describe individual shipments

C. The procedure that produced this interval generates ranges that hold the population mean for 95% of samples.
Ans) 95% of intervals created in this way contain the true population mean.

D. If we get another sample, then we can be 95% sure that the mean of this second is between 205 and 295.
Ans) the interval does not describe the mean of another sample.

E. We can be 95% confident that the range 160 to 340 holds the population mean.
Ans) the interval doesnot correspond to a 95% confidence interval level.

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) a. the Z-internal 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) a

For a 95% confidence level and a margin of error of 4%, we have:

$Z\approx 1.96$

$E=0.04$

We don't have the estimated proportion p, but to guarantee the maximum margin of error, we use $p=0.5$ (maximum variance).

$n=1.962\cdot\ 0.5\cdot\ (1-0.5)0.042n=0.0421.962\cdot\ 0.5\cdot\ (1-0.5)$

$n=3.8416\cdot\ 0.250.0016n=0.00163.8416\cdot\ 0.25$ $n\approx 240.1$

So, the minimum sample size required is approximately 240.1. Since you cannot have a fraction of a respondent, you round up to the nearest whole number. Therefore, the answer 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

For a 98% confidence level and a margin of error of 4%, we have:

$Z \approx 2.33$

$E = 0.04$

$p = 0.5$

$=2.332 \cdot 0.5 \cdot (1-0.5)0.042 n = 0.0422.332 \cdot 0.5 \cdot (1-0.5)$

$=5.4289 \cdot 0.250.0016 n = 0.00165.4289 \cdot 0.25$

$n \approx 848.18$

So, the minimum sample size required is approximately 848.18. Rounding up to the nearest whole number, the answer is 848