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 sample size is not necessarily required to be fixed percentage of population size. The sample size should be determined based on desired level of confidence as well as variability of population.

   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 is a list of all the members of the population from which the sample is drawn, not just those who appear in the sample.

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

Ans: False (Not Always True). Larger surveys have the potential to provide a more accurate representation of the population. But accuracy depends on sampling method and degree of sample is representative of population. A smaller sample which is well designed and representative will be more accurate than poor designed large sample.

2. *PC Magazine* asked all 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 the reader of PC magazine who participated in survey.

   B. The parameter of interest

Ans: the average rating of the Kodak compact digital camera given by all readers of PC Magazine who participated in the survey

   C. The sampling frame

Ans: All the readers of PC magazine who are invited to participate in survey.

   D. The sample size

Ans: 225, all readers who rated Kodak compact digital camera.

   E. The sampling design

Ans: Sampling design is not specified in question. (Random Sampling maybe)

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

Ans: Non-response bias, if those who not respond had different than those who did. And selection bias if sample of 225 is not representative of population.

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

   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: False

   95% confidence interval means if we repeat the sampling process 100 times, we will capture population mean 95 times in our interval. It does not imply anything about the proportion being less than or greater than 50%.

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

   Ans: True

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

   A. ¼
   B. ½
   C. ¾
   D. 1

Ans: The information is not sufficient to answer this question.

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: If the sample were based on 2,000 users, we can construct a 95% confidence interval for the true proportion of users who use Mozilla Firefox:

$$p \pm z^* \left(\sqrt{\left(\frac{p(1-p)}{n} \right)} \right)$$

For the given sample proportion of 0.046 and sample size of 2000 and z* for 95% from normal distribution is 1.96, So we have:

$$0.046 \pm 1.96 \times \left(\sqrt{\left(0.046 \times \frac{1 - 0.046}{2000} \right)} \right)$$

$$0.046 \pm 0.009$$

which results in a confidence interval of (0.036, 0.056). Since the lower bound of this interval is greater than 0.05, 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 is the case, then can Microsoft conclude that Mozilla has a less than 5% share of the market?

Ans: It is highly unlikely to sample to be include data of all the daily internet users. Even though if claim of WebSideStory is true and their sample truly includes all daily Internet users, then Microsoft cannot conclude that Mozilla has a less than 5% share of the market based on this sample alone. However, it is still possible for Microsoft to conduct their own study and gather their own sample to try to estimate the true market share of Mozilla with a certain level of confidence.

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

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-interval is shorter
Since the t-distribution has heavier tails than the z-distribution, the t-interval will be wider than the z-interval for the same level of confidence.

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. 600

To find sample size we can use following formula,

$$n = \frac{(z^2 \times p \times q)}{E^2}$$

- z= 1.96 for 95% of confidence level
- p is the estimated proportion of businesses that plan to hire additional employees. For maximum variance we assume, p=0.5
- q = 1 − p = 0.5
- The maximum margin of error E = 0.04

By putting these values, we get,

$$n = \frac{(1.96^2 \times 0.5 \times 0.5)}{0.04^2} = 600.25$$

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. 848
By changing value of z and using same formula,
- z=2.33 for 98%

$$n = \frac{(2.33^2 \times 0.5 \times 0.5)}{0.04^2} = 848.26$$