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 –

1. True: The representation of the survey results should have a sample size. The sample size must be a fixed percentage of the total population size of the survey.

2. False: The sampling frame refers to a list of an item which responds to the question and not the ones which do not respond to the questions.

3. True: The larger conveys a more accurate impression of the population as larger surveys involve large sample size which reduces the chances of error

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**
   - 9000
   B) The parameter of interest
   - μ, π, σ
   C) The sampling frame
   - 1 to 10
   D) The sample size
   - 225
   E) The sampling design
      - Define population, Determine sampling frame, Select sampling technique, Determine sample size, Execute sample process.

F) Any potential sources of bias or other problems with the survey or sample
- Marginal error, Confidence interval, confidence level

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**. Confidence interval identifies the collection of values for the population parameter that are consistent with the observed sample.

   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. We have evidence in that direction but we cannot confirm 100% based on this data. We have to consider the values out of this range (i.e. more than 95% confidence interval).

   III. **The 95% Confidence-Interval for μ only applies if the sample data are nearly normally distributed.**
   Ans -False. We should have a moderately large sample (usually at least larger than 30 for many cases), the central limit theorem implies that the sampling distribution is normal regardless of the data itself.

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

   A. ¼
   B. ½
   C. ¾
   D. 1

Ans –   **D.** 1

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?**

   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 - (I) Let **p** = population proportion share of the market by Mozilla

So, **Null Hypothesis**, H_0 : p $\geq$ 5% {means that Mozilla has more than or equal to 5% share of the market}

Alternate Hypothesis, H_A : p < 5% {means that Mozilla has a less than 5% share of the market}

The test statistics that will be used here is **One-sample z-test for proportions**;

$$\text{T.S.} = \frac{\hat{p}-p}{\sqrt{\frac{p(1-p)}{n}}} \sim N(0,1)$$

where, $\hat{p}$ = sample proportion of the share of the market grabbed by Mozilla in 2004 = 4.6%

   n = sample of users = 2,000

So, the test statistics = $\frac{0.046-0.05}{\sqrt{\frac{0.05(1-0.05)}{2,000}}}$

   = - 0.821

The value of z-test statistics is -0.821.

Since in the question we are not given with the level of significance so we assume it to be 5%. Now, at 5% level of significance the z table gives a critical value of -1.96 for left-tailed test**.**

Since the value of our test statistics is more than the critical value of z, so we have insufficient evidence to reject our null hypothes*i*s as it will not fall in the rejection region.

Therefore, we conclude that Mozilla has more than or equal to 5% share of the market.

(II) We are given that Web Side Story claims that its sample includes all the daily Internet users. This means that the 4.6% share of the market represents the whole population.

Hence, we can 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.**

   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 - A. All shipments are between 205 and 295 books. **INCORRECT**
   B. 95% of shipments are   between 205 and 295 books. **CORRECT**
   C. The procedure that produced this interval generates ranges that hold the population mean for 95% of samples. **CORRECT**
   D. If we get another sample, then we can be 95% sure that the mean of this second sample is between 205 and 295. **CORRECT**
   E. We can be 95% confident that the range 160 to 340 holds the population mean. **INCORRECT**

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. **The z-interval is shorter**

   Because it tell us difference between mean of distribution and data points in standard deviation

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 - n=number of employers,

Assume $\hat{P}$=0.5, $\hat{q}$=0.5

Margin of Error=0.04 For 95% confidence interval,

the critical value Z= 1.96 ME

= Z * √ $\hat{p}\hat{q}$ n 0.04

= 1.96 * √ 0.5*0.5 n n= (1.962*0.5*0.5)0.042 = (0.9604) 0.0016 =600

So option A 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 - Z= 2.576 0.04

= 2.326 * √ 0.5*0.5 n n= (2.3262*0.5*0.5) 0.042 = (1.3525 0.0016) = 845.35

So option C is correct.