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) **True - The survey should have a fixed percentage of population size that helps to analyze the data well in order to give 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.

   Ans) **False - The sample should contain only those items that have responded because data which is not true or irrelevant gives wrong results**

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

   Ans) **True: The larger conveys a more accurate impression of the population as larger surveys involve large sample size which gives correct results while analyzing , the more the data , the more accurate is the result.**

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) **p=x/n=225/9000=0.025**

   B. The parameter of interest

Ans) **sample size,average,scale**

   C. The sampling frame

Ans) **9000**

   D. The sample size

Ans) **225**

   E. The sampling design

Ans) **Random Sampling of 225 readers - as there is no mention of systematical , stratified or cluster strata sampling which has procedures to follow for collecting sample**

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

Ans) **Pre-screening or advertising bias- In this the sample collected by advertising may have some bias as only the persons who bought the magazine will take the survey targeting random audience and not all the customers who bought the products.**

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)True,

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

Ans) **False. With a large enough sample, the central limit theorem implies a normal sampling distribution regardless of the distribution**

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

    A. ¼

Ans)Sample mean to be greater than Population mean , the chances are very low as according to Central Limit Theorem(CLT) there should be more than 30 samples for the average of sample means to be nearer to population mean , as it is very expensive to get the data of population

    B. ½
    C. ¾
    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?

Ans) Let **p** = population proportion share of the market by Mozilla

So, **Null Hypothesis**, H0➔p>=5% means that Mozilla has more than or equal to 5% share of the market

Alternate Hypothesis, Ha➔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;**

z-test = p^-p/(sqrt(p(1-p)/n)~ N(0,1)

where, p^ = sample proportion of the share of the market grabbed by Mozilla in 2004 = 4.6%

n = sample of users = 2,000

So, the z-test = (4.6%-5%)/sqrt(5%(1-5%)/2000)

=(0.046-0.05)/sqrt(0.05(1-0.05)/2000)

= -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** st.norm.ppf(0.05)= -1.6448

for left-tailed test.

Since the value of our z-test -0.821>-1.6448 is more than the critical value of z, so we have insufficient evidence to reject our null hypothesis 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. 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) (II) **We are given that WebSideStory 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. ➔ **Incorrect**

Ans) **the 95% confidence interval for the size of the shipment is [250-45 , 250+45] books = [205,295] . So, there cant be all shipments as it is mentioned about 95% interval for the size of shipment**

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 , the range generated cannot be correct as population mean will be nearer to sum of sample means**

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

Ans) **95 % confidence interval for mean is shorter for z-interval because t-critical is greater than z-critical value. Yes, z-interval is always shorter because t-critical value cannot be smaller than z-critical value**

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

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

Ans)**Margin of error is 4%=0.04**
z critical value for 95% from table is z= 1.96
assuming p = 0.5
n= ((z/ME)^2)*(p(1-p))
n=((1.96/0.04)^2) * (0.5(1-0.5))
n= (2401)*(0.25)
n=600.25

B. 400
C. 550
D. 1000

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

Ans) **Margin of error is 4%=0.04**
z critical value for 98% from table is z= 2.326
assuming p = 0.5
n= ((z/ME)^2)*(p(1-p))
n=((2.326/0.04)^2) * (0.5(1-0.5))
n= (3381.42)*(0.25)
n= 845.35

D. 543