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

   There should be sufficient data to draw an inference. If the sample size is small, we use T statistics & if sample size is more than 30 then we can use Z statistics.

   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 items in the target population from which the sample is selected. Sampling with no value has no meaning or does not add value hence we remove such entries.

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

   Ans: True

   SE inversely proportional to the sample size, hence more the sample size will have less SE and vice versa.

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 : Readers of PC Magazine (9000)
   B. The parameter of interest: Avg rating of Electronics Brand (Rating of Camera)
   C. The sampling frame: All readers of the issue where the survey was included
   D. The sample size : 225
   E. The sampling design: - Voluntary Response (Simple Random Sampling)
   F. Any potential sources of bias or other problems with the survey or sample:

Ans: There is chance of Voluntary Response Bias. (those who were particularly or only who are not happy with the product participated in the survey which can makes the results unreliable.)
Selection Bias - This survey will only talk about people who reads PC Magazine not all the people who does not read PC Magazine but uses the electronics product.

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

   The 95% confidence interval for the average purchase of customers at a department store is $50 to $110 means there is 95% possibility that the population mean lies between $50 to $110.

   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

   If the 95% confidence interval for the number of moviegoers who purchase concessions is 30% to 45%. This means that there is 95% probability NOT 100% that less than 50% moviegoers purchase concessions.

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

   Ans: False

   Not only Normally distribution but sample size also should be sufficiently large (>=30). Even while applying central limit theorem, we assume that the mean sampling avg is normally distributed.

4. What are the chances that $\overline{X} > \mu$?
   A. ¼
   B. ½
   C. ¾
   D. 1

In a bell shape curve area to the right side of mean is $\overline{X} > \mu$ so Chances are ½.
Sampling mean is around the population mean (less or greater) this we give by giving CI.

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 : False

$\sigma = ?$

n = 2000,

H0 : $\mu >= 0.05$

Ha : $\mu < 0.05$

X=0.046

One Sample – One Tail test - lets assume CI = 95% , Z= +/- 2

Considering the binomial distribution =

(0.046-0.05) / *sqrt(0.05*(1-0.05)/2000) should be less than Z score*

P value (20.58%)> alpha (5%), hence accept the null Hypothesis.

Mozilla has market share greater than 5%

   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?

Yes.

Since it claims that its sample includes all the daily internet users hence the population. So, the inference of Mozilla Firefox browser had grabbed a 4.6% share of the market applies to whole population.

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: Incorrect

   95% Probability that the size of shipments are between 205 to 295 books for shipments.

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

   Ans: Incorrect.

   Interval tells about size of shipment not about shipment.

   Weather it is 1 shipment or many, that shipment has 95% possibility that size is 250 ± 45 books.

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

   Ans: Correct.

   95% chance that Population mean is contained in that interval.

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

   Ans: Incorrect

   Interval doesn't describe the mean of another sample. **It describes Population Mean.**

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

   Ans: Incorrect

   - as range is increased probability will also increase it will be more.
   - This range is not for 95% CL

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

Though: CI = mean + Z*Std Error

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

Standard deviation of binomial distribution =sqrt [p*(1-p)/n]

Margin of Error = Z * sqrt [p*(1-p)/n]

ME < 0.04 , Z for 95% CI = 1.96

σ^2 = p(1-p) = 0.5 x(1-0.5) = 0.25
n = (1.96 * 1/0.04)^2 x 0.25 = 600.25

Hence min sample size is 600 for CI of 95%.

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

Standard deviation of binomial distribution =sqrt [p*(1-p)/n]

Margin of Error = Z * sqrt [p*(1-p)/n]

ME < 0.04 , Z for 98% CI = 2.33

σ^2 = p(1-p) = 0.5 X (1-0.5) = 0.25

n = (2.33 * 1/0.04)^2 X 0.25 = 848.26

Hence min sample size is 848 for CI of 98%.