**Topics: Confidence Intervals**

1. For each of the following statements, indicate whether it is True/False. If false, explain why.
2. The sample size of the survey should at least be a fixed percentage of the population size in order to produce representative results.

***False***.

The sample size depends on required margin of error and confidence interval. The sample size should have at least 30 observations, so that t- value is equivalent to the z score, and available from statistical tables

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

***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.

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

***True***.

Large sample size will result in less standard deviation compared to small sample size. Thus, we can say larger sample is more accurate.

1. *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:
2. The population:

Readers of the PC magazine i.e., 9000.

1. The parameter of interest:

Reader’s satisfaction with different brands of electronics.

1. The sampling frame:

Rating for Kodak compact digital camera.

1. The sample size:

225

1. The sampling design:

Cluster sampling method.

[Cluster sampling](https://www.questionpro.com/blog/cluster-sampling/) is a method where the researchers divide the entire population into sections or clusters that represent a population. Clusters are identified and included in a sample based on demographic parameters like age, sex, location, etc. This makes it very simple for a survey creator to derive effective inference from the feedback.

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

Readers who don’t know about the product can give rating which results in loss of reliable insights. If some of readers are in favor of company and didn’t know about product can make survey basis towards the product favor or increase average of rating product and vice versa for those displeased with company.

1. For each of the following statements, indicate whether it is True/False. If false, explain why.
2. 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.

***True***.

Confidence interval identifies the collection of values for the population parameter that are consistent with the observed sample. So, mean of population can be in between $50 to $110.

1. 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.

***False.***

The 95% confidence interval for the number of moviegoers who purchase concessions is 30% to 45%, this means that there is a 95% chance that only 30 to 45 % of moviegoers purchase concessions, which is less than 50%. But, when we take 100% confidence, we cannot make an inference that fewer than half of all moviegoers purchase concessions as range of moviegoers who purchase concessions can increase from (30%, 45%) to (35%, 50%).

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

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.

1. What are the chances that ?
2. ¼
3. ½
4. ¾
5. 1

The probability of getting a sample mean greater than μ (population mean) is **50%**, as long as your sampling distribution follows a normal distribution (this occurs if the population distribution is normal or the sample size is large)

1. 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.
2. If the sample were based on 2,000 users, could Microsoft conclude that Mozilla has a less than 5% share of the market?

Sample = 2000

Null hypothesis, H0 is p> =0.05, has more than or equal to 5% share in the market

Alternate hypothesis, H1 is p < 0.05, has less than 5% share in the market

Taking test statistics, one sample z test for proportion,

Z = - 0.821

Z statistics for 5% is -1.96. so, test statistic value is greater than z critical value. So, null hypothesis is true.

1. 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?

We have data on the entire population and the sample value accurately reflects the population number. Thus, we can conclude that the share is less than 5%

1. 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?
2. All shipments are between 205 and 295 books.
3. 95% of shipments are between 205 and 295 books.
4. The procedure that produced this interval generates ranges that hold the population mean for 95% of samples.
5. If we get another sample, then we can be 95% sure that the mean of this second sample is between 205 and 295.
6. We can be 95% confident that the range 160 to 340 holds the population mean.

It’s clearly stated that the confidence interval is 95%. So, we are sure that 95% of the shipment is between 205 and 295.

1. Which is shorter: a 95% *z*-interval or a 95% *t*-interval for *μ* if we know that σ =s?
2. The z-interval is shorter
3. The t-interval is shorter
4. Both are equal
5. We cannot say

z-interval is always shorter because t-critical value cannot be smaller than z-critical value. It tells the 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.

1. 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)?
2. 600
3. 400
4. 550
5. 1000

Given

n=number of employers, Margin of Error=0.04

Assume =0.5, 𝑞̂=0.5

For 95% confidence interval, the critical value Z= 1.96

Margin of Error =

0.04 =

n = = 600.25 ≈ 600

1. Suppose we want the above margin of error to be based on a 98% confidence level. What sample size (minimum) must we now use?
2. 1000
3. 757
4. 848
5. 543

Given

n=number of employers, Margin of Error=0.04

Assume =0.5, 𝑞̂=0.5

For 98% confidence interval, the critical value Z= 2.326

Margin of Error =

0.04 =

n = = 845.35 ≈ 845

minimum sample size is 845