**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.

Ans) False.

The results depend on the size(n) of the sample. The sample size should have at least 30 observations.

1. 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.

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

Ans) 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
3. The parameter of interest
4. The sampling frame
5. The sample size
6. The sampling design
7. Any potential sources of bias or other problems with the survey or sample

Ans) A: Readers of the magazine =9000

B: Rating of the camera(7.5)

C: Sampling frame: All readers of the issue where the survey was included.

D: 225

E: Voluntary response

F: It is possible that only those who were particularly pleased or only who are displeased with the product participated in the survey which can makes the results unreliable.

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.

Ans) True.

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

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.

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% confidance interval).

1. 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.

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

Ans) B.

This is pure assumption. There is a 50% chance that the sample mean() is greater than the population mean(µ).

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?

Ans) here,

=0.046, n=2000, = 1.96, q= 0.954

95% confidence interval for the proportion of web users using Mozilla is

± Z = ± 1.96 =0.046±0.00918 = 0.0368- 0.0551.

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?

Ans) In this case, 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.

Ans) Incorrect.

The interval of (205,295) is for 95% confidence not for 100%.

1. 95% of shipments are between 205 and 295 books.

Ans) Incorrect.

The interval doesn’t describe individual shipments.

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

Ans) Correct.

95% of intervals created in this way contain the true population mean.

1. 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.

The interval doesn’t describe the mean of another sample.

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

Ans) Incorrect.

The interval doesn’t correspond to a 95% confidence level.

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

Ans) A. the Z- interval is shorter.

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

Ans) here, n=number of employers, Assume =0.5,=0.5 Margin of Error=0.04

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

ME = Z \*

0.04 = 1.96 \*

n= = =600 =A

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

Ans) Z= 2.576

0.04 = 2.326 \*

n= = = 845.35 = C

1. Examine the following normal Quantile plots carefully. Which of these plots indicates that the data …
2. Are nearly normal?
3. Have a bimodal distribution? (One way to recognize a bimodal shape is a “gap” in the spacing of adjacent data values.)
4. Are skewed (i.e. not symmetric) ?
5. Have outliers on both sides of the center?



1. Skewed
2. Outliers
3. Normal
4. Bimodal
5. For each of the following statements, indicate whether it is True/False. If false, explain why.

The manager of a warehouse monitors the volume of shipments made by the delivery team. The automated tracking system tracks every package as it moves through the facility. A sample of 25 packages is selected and weighed every day. Based on current contracts with customers, the weights should have *μ* = 22 lbs. and *σ* = 5 lbs.

1. Before using a normal model for the sampling distribution of the average package weights, the manager must confirm that weights of individual packages are normally distributed.

Ans) False.

1. The standard error of the daily average SE() = 1.

Ans) True.

SE=5/ = 1.

1. Auditors at a small community bank randomly sample 100 withdrawal transactions made during the week at an ATM machine located near the bank’s main branch. Over the past 2 years, the average withdrawal amount has been $50 with a standard deviation of $40. Since audit investigations are typically expensive, the auditors decide to not initiate further investigations if the mean transaction amount of the sample is between $45 and $55. What is the probability that in any given week, there will be an investigation?
2. 1.25%
3. 2.5%
4. 10.55%
5. 21.1%
6. 50%

Ans) D.

Here, µ= 50, σ = 40, n = 100. Standard Error= 40/ =>4

> pnorm(55,50,4)-pnorm(45,50,4)

[1] 0.7887005

> 1-0.7887005

[1] 0.2112995 => 21.1%

1. The auditors from the above example would like to maintain the probability of investigation to 5%. Which of the following represents the minimum number transactions that they should sample if they do not want to change the thresholds of 45 and 55? Assume that the sample statistics remain unchanged.
2. 144
3. 150
4. 196
5. 250
6. Not enough information

Ans)D.

> qnorm(0.975)

[1] 1.959964

µ = ± Z

50 = 55+ 1.959964\*

-5=78.4/

= 78.4/-5 = 245

1. An educational startup that helps MBA aspirants write their essays is targeting individuals who have taken GMAT in 2012 and have expressed interest in applying to FT top 20 b-schools. There are 40000 such individuals with an average GMAT score of 720 and a standard deviation of 120. The scores are distributed between 650 and 790 with a very long and thin tail towards the higher end resulting in substantial skewness. Which of the following is likely to be true for randomly chosen samples of aspirants?
2. The standard deviation of the scores within any sample will be 120.
3. The standard deviation of the mean of across several samples will be 120.
4. The mean score in any sample will be 720.
5. The average of the mean across several samples will be 720.
6. The standard deviation of the mean across several samples will be 0.60

Ans) D. the average of the mean across several samples will be 720.