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**

Calculating the right sample size is crucial to gaining the accurate information. The sample size of survey should not be fixed percentage of the population size.

Survey’s Confidence level and Margin of error almost solely depends on the number of responses you received.

Confidence level describes how sure you can be that your results accurate. As per the market standards it is 95%.

Margin of Error shows the range the surveys result would fall between if your confidence level true. As per the market standards it is 5%.

The lower sample size, higher margin of error and lower confidence level then it is less reliable.

Theoretically speaking a sample can Never be Too high. It is too expensive and time taking process.

Higher confidence level and lower margin of error with normal sample size are easier to achieve the reliable result.

Sample size should be neither too low Not too high. It should be appropriate.

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 list of required items that appear in a survey sample not required to include those that did not respond to questions. Always Survey questions should be very to Understand, Close end questions like Yes/No. Most important question should be very simple no complex, logical questions.

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

**Ans: True**

Explanation mentioned above.

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 **: (F)** - Any potential sources of bias or other problems with the survey or sample

The PC Magazine was included in an issue of Kadak that might be reason to rate average rate (7.5) be 225 readers. Here the issue was look like Biased and influenced the readers in the survey. Survey should not be influenced by anything.

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 : Z = X – *μ /* σ

*μ* = $100

C.I with 95% = [$50 , $110]

If it’s True Both Higher and Lower level standard deviation should match. Assume it is Normal distribution taking Z values with respect to 95% Confidence level

Lower Level - Z value at .025% = -1.97

σ = 50-100/-1.97 = 25.51

Higher Level - Z value at .97.5% = 1.97

σ = 110-100 / 1.97 = 10/1.97 = 5.10

Both are not equal means, **Hence it is False.**

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**

The Confidence Interval tells you the probability of population mean falling within that interval, it does not give you anymore information.

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

**Ans : True**

The Confidence-Interval for *μ* only applies if the sample data are normally distributed. Otherwise data skewed. That means it is not normally distributed.

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

**Ans : D ( 1 )** – Take a sample from the population of interest and calculate the mean of the sample giving the sample mean (X Bar). Now of course the sample mean will not equal to the population mean. But if the sample mean is SIMPLE Random sample, the sample means in an Unbiased estimate of the population mean. This mean that the sample mean is not systematically smaller or larger than the population mean or part another way**. If we repeatedly take lots of lots ( infinity number ) of samples then that mean of sample would equals to population mean.**

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 : No** , because sample were based on 2000 users. It might be biased. So Microsoft could not conclude that statement.

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: Yes,** its sample includes all the daily Internet users. It is not biased. So Microsoft conclude that above statement.

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.

**Ans : B** - 95% of shipments are between 205 and 295 is correct.

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 :** When we know the sigma ( s ) we go for the t- Distribution, if we know the (σ) then we go for Z- Distribution. As per the given statement that σ=s then Both are **Equal(C ).**

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.

Ans : The below two business issues solved by using the Margin of Error ( ME).

Margin of Error ( ME) = Z (1-Alpha) \* SQRT ( P \* Q) / n

P = Probability of Success

Q = 1- P

ME = Margin of Error.

Z (1-Alpha) = Z score at 1-Alpha confidence Level.

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 - true**
3. 400
4. 550
5. 1000

ME = 4% = 0.04

Z (95% of Confidence Interval ) = 1.96

Assume of P =50% = 0.5 and Q = 50% = .0.5

.04 = 1.96 \* SQRT(( .5 \* .5 )/n)

n = 0.9604/ 0.0016 = 600

n = 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 - TRUE**

ME = 5% = 0.05 ( Considered as per the Standard)

Z (98% of Confidence Interval ) = 2.326

Assume of P =50% = 0.5 and Q = 50% = .0.5

.05 = 2.326 \* SQRT(( .5 \* .5 )/n)

n = 1.3525/ 0.0025 = 541.02

**n = 542**

1. Examine the following normal Quantile plots carefully. Which of these plots indicates that the data?
2. Are nearly normal?

Ans : **Diagram C** is nearly Normal.

1. Have a bimodal distribution? (One way to recognize a bimodal shape is a “gap” in the spacing of adjacent data values.) -

**Ans : Diagram D**

1. Are skewed (i.e. not symmetric)?

Ans : **Diagram A**

1. Have outliers on both sides of the center?

Ans : **Diagram B**



1. 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 : **True :** The central Limit Theorem states that if you have population with mean and standard deviation takes sufficient large random samples from the population with replacement , then the distribution of sample mean will be approximately Normal Distribution. This will hold true regardless if whether the source population is Normal or Skewed, Provide the sample size sufficiently large ( n >=30). If the population is Normal then theorem states hold true even the sample size smaller than 30. In fact it holds true even the population is Binomial.

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

Ans: **True :** Standard Error = Standard Deviation / SQRT of (n)

In the given details are

n = 25

σ = 5 lbs.

SE() = 5 / SQRT(25) = 5/5 = 1

The above given statement is TRUE.

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 : 21.1%**

As the given details in the problem statement.

σ = $40 μ = $50 and n=100

Control Interval = [Lower Limit, Upper Limit]

= [45 , 55 ]

CI = [Xbar +or – Z score(σ/SQRT(n)]

CI = [$50 + or – Z score(40/SQRT(100)]

CI = [$50 + of – Z score (4)]

In order to match with given control interval [$45 , $55] then Z score will be 1.25

Hence both should match.

**If you look at 1.25 value in the Z- table it will a 21.13%**

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 thresholdsof 45 and 55? Assume that the sample statistics remain unchanged.
2. 144
3. 150
4. 196
5. 250
6. Not enough information

Ans : D : 250

The given problem statement has below details.

σ = $40 μ = $50 and n=?

Maintain the probability of investigation = 5% = .05

Lower limit of Z 47.5% = -0.06270677794321385

CI = [Xbar +or – Z score(σ/SQRT(n)]

45 = [50 – (-0.06270677794321385) (40/SQRT(n)]

45 -50 = 2.496 / SQRT(n)

-5 = 2.496/ SQRT(n)

SQRT(n) = -.4992 =**250**

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.

False – Any sample deviation will not be 120 because it has very long and thin tail toward the higher end.

1. The standard deviation of the mean of across several samples will be 120.

True – for several will be 120.

1. The mean score in any sample will be 720.

False – Any mean will not be 720 because it has very long and thin tail toward the higher end.

1. The average of the mean across several samples will be 720.

True – for several will be 720.

1. The standard deviation of the mean across several samples will be 0.60

Flase – Given standard deviation is 120 for population. If we do the average the it should very closer value to 120 not like ,60.

**Ans : B and D**