

## Basic Statistics Level-2

Saikrishna Budi

Linkedin : <http://linkedin.com/in/sai-krishna-budi-3b0341209>

### 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.
  - II. The sampling frame is a list of every item that appears in a survey sample, including those that did not respond to questions.
  - III. Larger surveys convey a more accurate impression of the population than smaller surveys.

**Sol:**

- 1.) The sample size of the survey should at least be a fixed percentage of the population size in order to produce representative results.

The answer is False.

Reason:

Sample data can be decided by the size of the population data, that can be or cannot be adequate.

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

The answer is True

Reason:

The sampling frame lists all items in a survey, including non-responders, for an unbiased representation of the entire population.

- 3.) Larger surveys convey a more accurate impression of the population than smaller surveys.

The answer is True

Reason:

Larger surveys offer a more accurate picture of the population by minimizing random variations and capturing greater diversity.

## Basic Statistics Level-2

Saikrishna Budi

Linkedin : <http://linkedin.com/in/sai-krishna-budi-3b0341209>

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
  - B. The parameter of interest
  - C. The sampling frame
  - D. The sample size
  - E. The sampling design
  - F. Any potential sources of bias or other problems with the survey or sample

**Sol:**

**A.) The population**

Here, the population refers to **all the readers of PC Magazine** who were invited to participate in the survey of their satisfaction with different brands of electronics.

**B.) The parameter of interest**

The parameter of interest is **the average satisfaction rating (7.5)** given by PC Magazine readers to a Kodak compact digital camera in the survey.

**C.) The sampling frames**

Each and every reader that responded by rating **the products (around 9000).**

**D.) The sample sizes**

The sample **size is 225.**

**E.) The sampling designs**

The paragraph does not provide specific details about the sample design used in the survey (Volunteer Response).

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

Potential sources of bias in the survey include nonresponse bias, as surveys from magazines like PC Magazine may involve a specific group (its readers) rather than a representative sample of the entire population using various electronics brands.

## Basic Statistics Level-2

Saikrishna Budi

Linkedin : <http://linkedin.com/in/sai-krishna-budi-3b0341209>

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

Sol:

- 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.

The answer is True

Reason:

A 95% confidence interval of \$50 to \$110 implies that we are 95% confident the true average purchase falls within this range, making \$100 a plausible value for the population mean.

- 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.

The answer is False

Reason:

The above information implies only for 30 to 45% of moviegoers at 95% confidence further than this we do not have any knowledge. Therefore, we cannot be 100% sure for the above statement.

## Basic Statistics Level-2

Saikrishna Budi

Linkedin : <http://linkedin.com/in/sai-krishna-budi-3b0341209>

- iii.) The 95% Confidence-Interval for  $\mu$  only applies if the sample data are nearly normally distributed.

The answer is False

Reason:

The 95% confidence Interval for population mean can be applied to distribution that aren't normal, but they are easy to understand in symmetric distributions. We can use the central limit theorem to make a normal approximation. The rule where about 95% of observations are within 2 standard deviations of the mean is only approximately true. However, it holds very well for the normal distribution. The mean of data tends to be normally distributed when the sample size is sufficiently large.

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

- A.  $\frac{1}{4}$
- B.  $\frac{1}{2}$
- C.  $\frac{3}{4}$
- D. 1

Sol: The answer is B.

(This is only an assumption, because if we consider more than 50% for sample mean to be greater than the probability of getting a lower value gets overshadowed because sample mean has an equal chance to be lesser than population mean).

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?
  - 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?

## Basic Statistics Level-2

Saikrishna Budi

Linkedin : <http://linkedin.com/in/sai-krishna-budi-3b0341209>

Sol:

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

The answer is no.

Reason:

As the p Value > alpha for 95% confidence we failed to reject null hypothesis  $H_0$ : > 5% null hypothesis

- 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?

The answer is Yes.

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  $250 \pm 45$  books. Which, if any, of the following interpretations of this interval are correct?
- A. All shipments are between 205 and 295 books.
  - B. 95% of shipments are between 205 and 295 books.
  - C. The procedure that produced this interval generates ranges that hold the population mean for 95% of samples.
  - D. If we get another sample, then we can be 95% sure that the mean of this second sample is between 205 and 295.
  - E. We can be 95% confident that the range 160 to 340 holds the population mean.

Sol:

- A.) All shipments are between 205 and 295 books.

Incorrect, 95% of the time the size of the shipment will be around  $250 \pm 45$  books.

## Basic Statistics Level-2

Saikrishna Budi

Linkedin : <http://linkedin.com/in/sai-krishna-budi-3b0341209>

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

Reason

As we increase the range  $\pm 1$  sigma the % of confidence increases for a normal distribution it will be 97.5% that the mean will lie in between 160 to 340 ( $-3$  sigma to  $+3$  sigma).

7. Which is shorter: a 95% z-interval or a 95% t-interval for  $\mu$  if we know that  $\sigma = s$ ?

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

Sol:

The answer is A

Reason:

For 95% confident Z-interval is shorter (1.960), as for t-interval (1.962) it increases as degree of freedom decreases.

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.

## Basic Statistics Level-2

Saikrishna Budi

Linkedin : <http://linkedin.com/in/sai-krishna-budi-3b0341209>

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

**Sol:**

The answer is B

**Explanation:**

Margin of Error =  $Z \cdot \sqrt{p \hat{p} \cdot q \hat{q} / n}$  n = sample size, Z is score depends on the confidence level, p hat is the predictable value and q hat is the value to be predicted value

let's assume p hat and q hat as 0.5 and margin of error is given 0.04

$n = (z)^2 \cdot p \hat{p} \cdot q \hat{q} / ME^2$   $n = (1.65)^2 \cdot (0.5) (0.5) / (0.04)^2$   $n = 425$  sample size.

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

**Sol:**

The answer is C

**Explanation:**

Let's assume p hat and q hat as 0.5 and margin of error is given 0.04  
 $n = (z)^2 \cdot p \hat{p} \cdot q \hat{q} / ME^2$   $n = (2.32)^2 \cdot (0.5) (0.5) / (0.04)^2 = 848$  samples size.