

Topic: 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: True.

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 must only contain those values which has response. The ones that have no response only increases the size of the sample and makes no scene to the analyst.

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

Ans: True.

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

Ans: 9000

B. The parameter of interest

Ans: 7.5 out of (1 – 10)

C. The sampling frame

Ans: All the readers of the magazine (225)

D. The sample size

Ans: 225

E. The sampling design

Ans: Voluntary Voting

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

Ans: Only people who are extremely happy or people who are extremely unhappy may be included in the sample which makes the output unreasonable.

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.

With 95% Confidence Interval, the purchase range is from 50-100, and \$100 lies within this range.

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.

Because we cannot get into such conclusion with the provided data.

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

Ans: False.

It can be applied for any kind of Confidence Interval.

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

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

Ans: Option D

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:

```
import numpy as np
from scipy import stats
from scipy.stats import norm
# Apply One-Sample One-Tail Z-Test
z_scores = (0.046-0.05)/(np.sqrt((0.05*(1-0.05))/2000))
z_scores
-0.820782681668124
p_value = 1 - stats.norm.cdf(abs(z_scores))
p_value
0.20588503245107104
```

Hence Null hypothesis is rejected, since $0.205 > 0.05$.

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?

Ans: We are given that WebSideStory claims that its sample includes all the daily Internet users.

This means that the 4.6% is the population percentage. Comparing it with Microsoft's claim that Mozilla has a less than 5% share of the whole market is True.

Hence, we can conclude that Mozilla has a less than 5% share of the market.

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: No, because interval is only 95% confident.

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

Ans: Yes.

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

Ans: Yes.

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: No, because the interval is not for the mean of another sample.

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

Ans: No, because the interval is not corresponding to 95% confidence interval.

7. Which is shorter: a 95% z -interval or a 95% t -interval for μ 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

Ans: 95% Z-interval is shorter because t-critical is greater than Z-critical value.

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

Ans: Consider $p \sim 0.5$

95% confidence interval z-score = 1.960

$$N = ((1.960/0.4)^2 * 0.5 * (1-0.5)) * 100 = 600.25 \sim 600$$

Option A

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

Ans: Consider $p \sim 0.5$

z-score for 98% interval = 2.326

$$N = 100((2.326/0.4)^2 * (1-0.5)) = 848.36 \sim 848$$

Option C