

## 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.

**Ans.** False: A Sample size of 30 is considered large enough, but that may not be adequate.

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 refers to a list of an item which responds to the question and not the ones which do not respond to the questions.

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:** All the readers of pc magazine.

B. The parameter of interest

**Ans:** Digital camera

C. The sampling frame

**Ans:** 9000

D. The sample size

**Ans:** 225

E. The sampling design

**Ans:** Kodak compact

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

**Ans.** Selection of the readers, selection of the issue which will contain the survey

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.

Reason: The 95% confidence interval for the average purchase of customers at a department store is \$50 to \$110. Which means that there is a 95% chance that the population mean will fall between \$50 and \$110. Hence, as \$100 falls between \$50 and \$110, it is a plausible value for the population mean at this confidence level.

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:** True

Reason: 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%. Hence, we can infer that fewer than half of all the moviegoers purchase concessions.

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

**Ans:** True

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

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

Ans.  $\frac{1}{2}$

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:

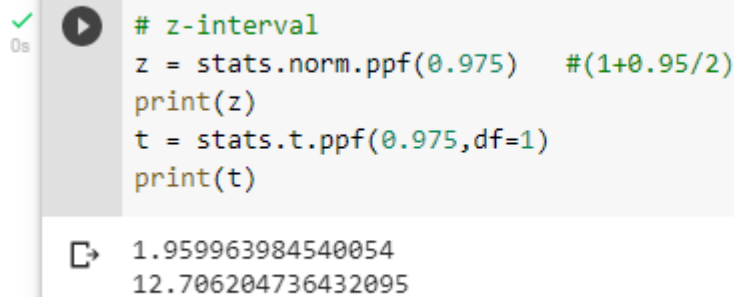
```
#Assume null hypothesis: H0: p>5%
#Alternate hypothesis: Ha: p<5%
z_score = (0.046-0.05)/(np.sqrt((0.05*(1-0.05))/2000))
print('z_score : ',z_score)
p_value = 1-stats.norm.cdf(abs(z_score))
print('p_value : ',p_value)
# Accept Null Hypothesis i.e. Mozilla market share > 5% Thus, Microsoft cannot conclude that Mozilla has a less than 5% share of the market

z_score : -0.820782681668124
p_value : 0.20588503245107104
```

- 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:** Yes, In this case, we have data on entire population and sample value accurately reflects the population. Thus we can conclude that Mozilla has share less than 5%.
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.  
**Ans:** Incorrect , All means 100% but the interval is given for 95%.
- B. 95% of shipments are between 205 and 295 books.  
**Ans:** Incorrect, the intervals doesn't describe individual shipments.
- C. The procedure that produced this interval generates ranges that hold the population mean for 95% of samples.  
**Ans:** 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.  
**Ans:** Incorrect, this interval doesn't describe the mean for another sample.
- E. We can be 95% confident that the range 160 to 340 holds the population mean.  
**Ans:** Correct
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

**Ans:** A



```

# z-interval
z = stats.norm.ppf(0.975)  #(1+0.95/2)
print(z)
t = stats.t.ppf(0.975,df=1)
print(t)

```

1.959963984540054  
12.706204736432095

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:** margin of error =  $z \cdot (\sqrt{p(1-p)/N})$

$$N = [(z/m)^2] \cdot p(1-p)$$

$$m = 0.04, p = 0.5, z(95\%) = 1.96$$

$$N = (1.96/0.04)^2 \cdot 0.25$$

$$N = 592.13 \text{ which is near to } 600.$$

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:** Here  $z = 2.33$

$$N = (2.33/0.04)^2 \cdot 0.25$$

$$N = 848.26.$$

