Frequently Asked Questions

1. How do you get the interval estimates for a population proportion and difference of two population proportions?

   Answer:

   A confidence interval for the population proportion P gives an estimated range of values which is likely to include an unknown population parameter, in this case P, the estimated range being calculated from a given set of sample data.

   100 (1- α) % C.I for the population proportion P is given by

   $[p - Z_{\alpha/2}\sqrt{\frac{PQ}{n}}\ ,\ p + Z_{\alpha/2}\sqrt{\frac{PQ}{n}}\ ]$

   Note: If P is not known then we use the estimate of P

   The estimate of P is p

   100 (1- α) % C.I for the population proportion P is given by

   $[p - Z_{\alpha/2}\sqrt{\frac{pq}{n}}\ ,\ p + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\ ]$

   Similarly when we have two populations

   100 (1- α) % C.I for the difference of population proportions is given by

   $[(p_1 - p_2)) - Z_{\alpha/2}\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}\ ,\ (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{P_1Q_1}{n_1} + \frac{P_2Q_2}{n_2}}\ ]$

   Note: If P_1 and P_2 are not known then we use the estimates of P_1 and P_2 as p_1 and p_2 .

   Then 100 (1- α) % C.I for the difference of two population proportions is given by

   $[(p_1 - p_2)) - Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}\ ,(p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}\ ]$

2. From a large consignment of mangoes 500 mangoes are selected and 135 of them are found to be rotten ones. Establish a 95% confidence interval for the percentage of rotten mangoes in the whole consignment

   Answer:

   Let xi denote the number of rotten mangoes = 135

N be the total number of mangoes =500

Estimate of the proportion of rotten mangoes

$$p = \frac{xi}{n} = \frac{135}{500} = 0.27$$

q=1-p=0.73

Since 100 (1- α) % =95%; α=0.05. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2} = 1.96$

Then 100 (1- α) % C.I for the population proportion P is given by

$$[p-Z_{\alpha/2}\sqrt{\frac{pq}{n}}, p + Z_{\alpha/2}\sqrt{\frac{pq}{n}}]$$

$$[0.27- 1.96 \sqrt{\frac{(0.27)(0.73)}{500}}, 0.27+ 1.96 \sqrt{\frac{(0.27)(0.73)}{500}}]$$

[0.2311, 0.3089]

Hence we can conclude that between 23 and 31 percent of mangoes in a large consignment are rotten ones

3. In a year there were 956 births in a town of which 520 were males. Establish a 95% confidence interval for the percentage of male births in a town

Answer:

Let xi denote the number of male births = 520

n be the total number of births =956

Estimate of the proportion of male births in a town

$$p = \frac{xi}{n} = \frac{520}{956} = 0.54$$

q=1-p=0.46

Since 100 (1- α) % =95%; α=0.05. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2} = 1.96$

Then 100 (1- α) % C.I for the population proportion P is given by

$$[p - Z_{\alpha/2}\sqrt{\frac{pq}{n}}\ ,\ p + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\ ]$$

$$[0.54 - 1.96\sqrt{\frac{(0.54)(0.46)}{956}}\ , 0.54 + 1.96\sqrt{\frac{(0.54)(0.46)}{956}}\ ]$$

[0.5084,0.5716]

Hence 95% CI for the percentage of male births is [0.5084,0.5716]

That is we conclude that the percentage of male births in a town is in between 51 and 57 percent .

4. In a sample of 600 high school students from a state 400 are found to use dot pens. In a sample of 900 students from a neighboring state 450 were found to use dot pens. Establish 99% confidence interval for the difference between the proportions of the users of dot pens among the students.

Answer:

Proportion of school students who use dot pens = p_1= 400/600 = 0.67

Proportion of school students of a neighbouring state who use dot pens = p_2= 450/600 = 0.75

q_1 =1- p_1= 1- 0.67 = 0.33

q_2 =1- p_2= 1- 0.75 = 0.25

Since 100 (1- α) % =99%; α=0.01. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2}$ = 2.57

Then 100 (1- α) % C.I for the difference of two proportions of the users of dot pens among the students is given by

$$[(p_1 - p_2)) - Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}\ , (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}\ ]$$

[(0.75-0.67) - 2.57 $\sqrt{\frac{(0.67)(0.33)}{600}+\frac{(0.75)(0.25)}{600}}$, (0.75-0.67) + 2.57 $\sqrt{\frac{(0.67)(0.33)}{600}+\frac{(0.75)(0.25)}{600}}$] = [0.0131, 0.1469]

Hence we can conclude that there is a difference of 1% to 15% in the proportion of users of dot pens among the students of two states.

5. In a sample of 800 citizens from a state 600 are found to be tea drinkers. In a sample of 1000 citizens from a neighbouring state 550 were found to be tea drinkers. Establish 99% confidence interval for the difference between the proportions of tea drinkers.

Answer:

Proportion of citizens who are found to be tea drinkers = p_1= 600/800 = 0.75

Proportion of citizens of a neighbouring state who are found to be tea drinkers = p_2= 550/1000 = 0.55

q_1 =1- p_1= 1- 0.75 = 0.25

q_2 =1- p_2= 1- 0.55 = 0.45

Since 100 (1- α) % =99%; α=0.01. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2}$ = 2.57

Then 100 (1- α) % C.I for the difference between the proportions of tea drinkers is given by

$$[(p_1 - p_2)) - Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}, (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}]$$

[(0.75-0.67) - 2.57 $\sqrt{\frac{(0.67)(0.33)}{600}+\frac{(0.75)(0.25)}{600}}$,(0.75-0.67)+ 2.57 $\sqrt{\frac{(0.67)(0.33)}{600}+\frac{(0.75)(0.25)}{600}}$] = [0.0131,0.1469]

Hence we can conclude that there is a difference of 1% to 15% in the proportion of users of dot pens among the students of two states

6. A simple random sample of size 100 is drawn from a population of students of size 300 who have two wheelers. Obtain 95% CI for the proportion of students in a college who possess two wheelers?

Answer:

Given n = 100 and N = 300

A fraction of students in a college who have two wheelers

Xi = 0 if she/ he hasn't possess two wheeler

Xi = 1 if she/ he has a two wheeler

Student : 1,2,3, ,4, ………………………….., 99,100

Xi : 0,1,1,………………………………..,1,1,1

∑Xi = 65

Estimate of the proportion of students in a college who have two wheelers

$$p = \frac{\sum_{i=1}^{n} x_i}{100} = \frac{65}{100} = 0.65$$

q=1-p=0.35

Since 100 (1- α) % =95%; α=0.05. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2} = 1.96$

Then 100 (1- α) % C.I for the population proportion P is given by

$$\left[p - Z_{\alpha/2}\sqrt{\frac{pq}{n}},\ p + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\right]$$

$$\left[0.65 - 1.96\sqrt{\frac{(0.65)(0.35)}{100}},\ 0.65 + 1.96\sqrt{\frac{(0.65)(0.35)}{100}}\right]$$

[0.55651, 0.74349]

Hence we can conclude that between 56 and 74 percent of students in a college possess two wheelers. In other words, with a margin of error of .09 , 65% of the students possess two wheelers.

7. In a sample of 850 men from district A, 350 are found to be smokers. In a sample of 1200 men from another district B, the number of smokers is 575. Establish

99% confidence interval for the difference between the proportions of the smokers in the two districts.

Answer:

Proportion of smokers in district A = p_1= 350/850 = 0.41

Proportion of smokers in district B = p2= 575/1200 = 0.48

q_1 =1- p_1= 1- 0.41 = 0.59

q_2 =1- p_2= 1- 0.48 = 0.52

Since 100 (1- α) % =99%; α=0.01. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2}$ = 2.57

Then 100 (1- α) % C.I for the difference between the proportions of the smokers in the two districts is given by

$$[(p_1 - p_2)) - Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}, (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1}+\frac{p_2q_2}{n_2}}]$$

$$[(0.48-0.41) - 2.57\sqrt{\frac{(0.41)(0.59)}{850}+\frac{(0.48)(0.52)}{1200}}, (0.48-0.41) + 2.57\sqrt{\frac{(0.41)(0.59)}{850}+\frac{(0.48)(0.52)}{1200}}] = [0.0126, 0.1274]$$

Hence we can conclude that there is a difference of 1% to 13% in the proportion of smokers in the two districts.

8. In a hospital 480 female and 520 male babies were born in a week. Establish a 99% confidence interval for the proportion of male babies' birth as a whole.

Answer:

Let xi denote the number of male births = 520

n be the total number of births =480+520=1000

Estimate of the proportion of male babies births in a hospital

$$p = \frac{xi}{n} = \frac{520}{1000} = 0.52$$

q=1-p=0.48

Since 100 (1- α) % =99%; α=0.01. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2} = 2.57$

Then 100 (1- α) % C.I for the population proportion P is given by

$$[\, p - Z_{\alpha/2}\sqrt{\frac{pq}{n}}\, ,\ p + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\,]$$

$$[0.52 - 2.57\sqrt{\frac{(0.52)(0.48)}{1000}}\, ,\ 0.52 + 2.57\sqrt{\frac{(0.52)(0.48)}{1000}}\,]$$

[0.4794, 0.5606]

Hence 99% CI for the proportion of male babies' births is [0.4794, 0.5606]

That is we conclude that the percentage of male babies births in a hospital is in between 48 and 56 percent.

9. One thousand randomly selected Americans were asked if they believed the minimum wage should be raised. 600 said yes. Construct a 95% confidence interval for the proportion of Americans who believe that the minimum wage should be raised.

Answer:

We have p = 600/1000 =0 .6; q=1-p=0.4

Since 100 (1- α) % =95%; α=0.05. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2} = 1.96$ and $n = 1000$

Then 100 (1- α) % C.I for the population proportion P is given by

$$[\, p - Z_{\alpha/2}\sqrt{\frac{pq}{n}}\, ,\ p + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\,]$$

$$[0.6 - 1.96\sqrt{\frac{(0.6)(0.4)}{1000}}\, ,\ 0.6 + 1.96\sqrt{\frac{(0.6)(0.4)}{1000}}\,]$$

= [0.57, 0.63]

Hence we can conclude that between 57 and 63 percent of all Americans agree with the proposal. In other words, with a margin of error of .03 , 60% agree.

10. 150 heads and 250 tails resulted in 400 tosses of a coin. Find 90% CI for the probability of a head

Answer:

Let xi denote the number of heads = 150

n be the total number of tosses =150+250=400

Estimate of the probability of a head

$$p = \frac{xi}{n} = \frac{150}{400} = 0.38$$

q=1-p=0.62

Since 100 (1- α) % =90%; α=0.10. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2}$ = 1.64

Then 100 (1- α) % C.I for the population proportion P is given by

$$[\, p - Z_{\alpha/2}\sqrt{\frac{pq}{n}},\ p + Z_{\alpha/2}\sqrt{\frac{pq}{n}}\,]$$

$$[0.38 - 1.64\sqrt{\frac{(0.38)(0.62)}{400}},\ 0.38 + 1.64\sqrt{\frac{(0.38)(0.62)}{400}}] =$$

[0.3402, 0.4198]

Hence 90% CI for the probability of head is [0.3402, 0.4198]

11. In a poll of 256 voters chosen at random from all registered voters in s constituency 56% are in favour of candidate A. Find the 95% CI for the population of the registered voters in the constituency who are in favor of candidate A

Answer:

Let xi denote the number of voters who are in favor of candidate A = 56

n be the total number of voters = 100

Estimate of the proportion of voters who are in favor of candidate A

$$p = \frac{xi}{n} = \frac{56}{100} = 0.56$$

q=1-p=0.44

Since 100 (1- α) % =95%; α=0.05. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2} = 1.96$

Then 100 (1- α) % C.I for the population proportion P is given by

$$[\, p - Z_{\alpha/2} \sqrt{\frac{pq}{n}}\, ,\ p + Z_{\alpha/2} \sqrt{\frac{pq}{n}}\,]$$

$$[0.56 - 1.96 \sqrt{\frac{(0.56)(0.44)}{100}}\, ,\ 0.56 + 1.96 \sqrt{\frac{(0.56)(0.44)}{100}}\,]$$

Hence 95% CI for the true proportion of defectives in the batch is [0.4627, 0.6573]

That is we conclude that the percentage of voters in the constituency who are in favor of candidate A is in between 46 and 66 percent.

12. In a random sample of 100 articles taken from a large batch of articles, 10 are found to be defective. Obtain 95% CI for the true proportion of defectives in the batch

Answer:

Let

xi denote the number of defective articles = 10

n be the total number of articles =100

Estimate of the proportion of defective articles

$$p = \frac{xi}{n} = \frac{10}{100} = 0.10$$

q=1-p=0.90

Since 100 (1- α) % =95%; α=0.05. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2} = 1.96$

Then 100 (1- α) % C.I for the population proportion P is given by

$$[\, p - Z_{\alpha/2} \sqrt{\frac{pq}{n}}\, ,\ p + Z_{\alpha/2} \sqrt{\frac{pq}{n}}\,]$$

$$[0.10 - 1.96 \sqrt{\frac{(0.10)(0.90)}{100}}\, ,\ 0.10 + 1.96 \sqrt{\frac{(0.10)(0.90)}{100}}\,]$$

[0.0412, 0.1588]

Hence 95% CI for the true population of defectives in the batch is [0.0412, 0.1588]

That is we conclude that the percentage of defectives in the batch is between 4 and 16 percent.

13. An investigation of the performance of two machines in a factory manufacturing large number of fans gives the following results. Find the 95% confidence Interval for the difference in proportions of defectives.

	No of Fans Examined	No of Defectives Found
Machine 1	275	17
Machine 2	350	22

Answer:

Proportion of defective fans found produced by machine 1= p_1= 17/275 = 0.061

Proportion of defective fans found produced by machine 2 = p_2= 22/350 = 0.063

q_1 =1- p_1= 1- 0.061 = 0.939

q_2 =1- p_2= 1- 0.063 = 0.937

Since 100 (1- α) % =95%; α=0.05. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2}$ =1.96

Then 100 (1- α) % C.I for the for the difference in proportions of defectives is given by

$$\left[(p_1 - p_2) - Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}},\ (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}\right]$$

$$\left[(0.063-0.061) - 196\sqrt{\frac{(0.061)(0.939)}{275} + \frac{(0.063)(0.937)}{350}},(0.063-0.061) + 196\sqrt{\frac{(0.061)(0.939)}{275} + \frac{(0.063)(0.937)}{350}}\right]=[-0.03612, 0.04012]$$

Notice that this confidence interval includes the value of zero which implies there may be negligible difference in the difference of proportion of defective fans produced by two machines.

14. Here is the data related to proportion of persons who would vote for a guilty verdict in a particular sexual harassment case from a study by Egbert, Moore, Wuensch, and Castellow. Of 160 mock jurors of both sexes, 105 voted guilty and 55 voted not guilty. Among the 80 male jurors 47 voted guilty. Among the 80 female jurors 58 voted guilty. Construct a 95% confidence interval for the difference between the two proportions based on sex.

Answer:

Proportion of female jurors who voted guilty = p_1= 58/80 = 0.725

Proportion of male jurors who voted guilty = p_2= 47/80 = 0.588

q_1 =1- p_1= 1- 0.725 = 0.275

q_2 =1- p_2= 1- 0.588 = 0.412

Since 100 (1- α) % =95%; α=0.05. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2}$ = 1.96

Then 100 (1- α) % C.I for the difference of two proportions in males and females is given by

$$[((p_1 - p_2)) - Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}, (p_1 - p_2) + Z_{\alpha/2}\sqrt{\frac{p_1q_1}{n_1} + \frac{p_2q_2}{n_2}}]$$

$$[(0.725\text{-}0.588) - 1.96\sqrt{\frac{(0.725)(0.275)}{80} + \frac{(0.588)(0.412)}{80}}, (0.725\text{-}0.588) + 1.96\sqrt{\frac{(0.725)(0.275)}{80} + \frac{(0.588)90.412)}{80}}]$$

= [-.009, 0.283]

Notice that this confidence interval includes the value of zero which implies there may be negligible difference in the difference of two sexes who voted guilty.

15. Ten Life insurance policies in a sample of 200 taken out of 50,000 were found to be insured for less than Rs.5000. How many policies can be reasonably expected to be insured for less than Rs. 5000in the whole lot at 95% confidence level.

Answer:

Let

xi denote the number of policies who are insured for less than Rs.5000 = 10

n be the total number of police = 200

Estimate of the proportion of of policies who are insured for less than Rs.5000

$$p = \frac{xi}{n} = \frac{10}{200} = 0.05$$

q=1-p=0.95

Since 100 (1- α) % =95%; α=0.05. Then from the table of probabilities of Normal distribution

$Z_{\alpha/2} = 1.96$

Then 100 (1- α) % C.I for the population proportion P is given by

$$[p - Z_{\alpha/2} \sqrt{\frac{pq}{n}}, \; p + Z_{\alpha/2} \sqrt{\frac{pq}{n}}]$$

$$[0.05 - 1.96 \sqrt{\frac{(0.05)(0.95)}{200}}, \; 0.05 + 1.96 \sqrt{\frac{(0.05)(0.95)}{200}}] = [0.0198, 0.0802]$$

Hence 95% confidence interval for the proportion of policies who are insured for less than Rs. 5000 is [0.0198, 0.0802]

Therefore number of policies cab be reasonably expected to be insured for less than Rs. 5000 in the whole lot at 95% CI is [0.0198, 0.0802]

Hence number policies can be reasonably expected to be insured for less than Rs. 5000 in the whole lot at 95% confidence level is in between 990 and 4010