(b) In the above sample of 60 students, 20% of them lived within 2·5 km of the school. Find the 95% confidence interval for the proportion of students from that school who live within 2·5 km of the school.

Solution:

Confidence interval for a proportion:

$$\hat{p} - 1.96\sqrt{\frac{p(1-p)}{n}} \leq p \leq \hat{p} + 1.96\sqrt{\frac{p(1-p)}{n}}$$

$$0.2 - 1.96\sqrt{\frac{0.2(1-0.2)}{60}} \leq p \leq 0.2 + 1.96\sqrt{\frac{0.2(1-0.2)}{60}}$$

$$0.2 - 0.10 \leq p \leq 0.2 + 0.10$$

$$0.1 \leq p \leq 0.3$$

(c) Data from 10 years ago shows that, at that time, 26% of the student population lived within 2·5 km of the school. Based on your answer to part (b) is it possible to conclude, at the 5% level of significance, that the proportion of students living within 2·5 km of the school has changed since that time? Explain your answer.

Solution:

No. As 0·26 (26%) lies inside the confidence interval above. Thus, we would not reject at the 5% level of significance.

(d) A statistician wishes to estimate, with 95% confidence, the proportion of students who live within a certain distance of the school. She wishes to be accurate to within 10 percentage points of the true proportion. What is the minimum sample size necessary for the statistician to carry out this analysis?

Solution:

She needs her confidence interval to extend at most 0·1, (10%), in either direction.

$$\therefore \quad 1.96\sqrt{\frac{p(1-p)}{n}} \leq 0.1$$

$$1.96\sqrt{\frac{0.2(1-0.2)}{n}} = 0.1$$

$$1.96\sqrt{\frac{0.16}{n}} = 0.1$$

$$n = 61$$

or As there is only one sample use $1.96\sqrt{\frac{1}{4n}}$ or $\frac{1}{\sqrt{n}}$

$$\therefore \quad \frac{1}{\sqrt{n}} = 0.1$$

$$\frac{1}{n} = 0.01$$

$$n = \frac{1}{0.01} = 100$$

using $1.96\sqrt{\frac{1}{4n}}$ gives $n = 97$

ie. 3 different methods