L15. Confidence Interval on the Population Proportion

Example 1

The fraction of defective integrated circuits produced in a photolithography process is being studied. A random sample of 300 circuits is tested, revealing 15 defectives. Calculate a 90% two-sided confidence interval on the fraction of defective circuits produced in this particular process.

Solution

Let x = the number of integrated circuits that are defective

$$n = 300 \quad ; \quad x = 15 \quad \Rightarrow \quad \hat{p} = \frac{x}{n} = \frac{15}{300} = 0.05$$

$n\hat{p} = 300(0.05) = 15 > 5$
$n(1-\hat{p}) = 300(1-0.05) = 285 > 5$

90% confidence:

$$\hat{p} \pm Z_{\alpha/z}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \Rightarrow \quad 0.05 \pm 1.645\sqrt{\frac{0.05(1-0.05)}{300}}$$

$$0.05 \pm 0.0207$$

$$0.0293 < p < 0.0707$$

Interpretation:

With repeated sampling, we a 90% confident that the actual percentage of defective circuits is between 2.93% and 7.07%

Example 2

A mail order company promises its customers that the products ordered will be mailed within 72 hours after an order has been placed. The quality control department checks from time to time to see if this promise is fulfilled. Recently the quality control department took a sample of 50 orders and found that 35 of them were mailed 72 hours after the order was placed.

a. Construct a 98% confidence interval for the percentage of all orders that were mailed within 72 hours of their placement.

b. Construct a 98% lower-confidence bound for the percentage of all orders that were mailed within 72 hours of their placement.

c. Construct a 96% upper-confidence bound for the proportion of all orders that were mailed within 72 hours of the order being placed.

Solution

Let x = the number of orders which are shipped out within 72 hours

$$n = 50 \quad ; \quad x = 35 \quad \Rightarrow \quad \hat{p} = \frac{x}{n} = \frac{35}{50} = 0.70$$

$n\hat{p} = 50(0.70) = 35 > 5$
$n(1-\hat{p}) = 50(1-0.70) = 15 > 5$

a. 98% confidence:

$$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \Rightarrow \quad 0.70 \pm 2.33\sqrt{\frac{0.70(1-0.70)}{50}}$$

$$0.70 \pm 0.1510$$

$$0.549 < p < 0.851$$

Interpretation:

With repeated sampling, we a 98% confident that the actual percentage of packages which are sent out within 72 hours is between 54.90% and 85.10%

b. 98% LCB:

$$\hat{p} - Z_{\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \quad \Rightarrow \quad 0.70 \pm 2.05\sqrt{\frac{0.70(1-0.70)}{50}} \leq p$$

$$0.70 - 0.1328 \leq p$$

$$0.5672 \leq p$$

Interpretation:

With repeated sampling, we a 98% confident that at least 56.72% of all orders are mailed out within 72 hours of placement.

c. 96% UCB:

$$p \leq \hat{p} + Z_{\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \Rightarrow \quad p \leq 0.70 + 1.75\sqrt{\frac{0.70(1-0.70)}{50}}$$

$$p \leq 0.70 + 0.1134$$

$$p \leq 0.8134$$

Interpretation:

With repeated sampling, we a 96% confident that at most 81.34% of all orders are mailed out within 72 hours of placement.

Example 3

In 2011, scientists in the US used $660000 in federal research money to examine whether distant prayer could heal AIDS, $406000 to see if squirting brewed coffee into someone's intestines would help with pancreatic cancer, and $1.25 million to examine whether massages made people with advanced cancer feel better. The distant prayers and coffee enemas did not work, but the massages did. In a group of 200 cancer patients, all of who received regular massages, 165 of them reported feeling better after their treatment.

a. Construct a 85% confidence interval for the true proportion of patients who felt better after receiving a massage.

b. Suppose that a preliminary study has found that 84% of patients felt better after receiving a massage. What should the sample size be so that the 85% confidence interval for the population proportion has a margin of error 0.05?

c. In the absence of a preliminary estimate, for $\hat{p}$, what sample size should be used so that the 85% confidence interval for the population proportion has a margin of error 0.05?

Solution

Let x = the number of patient who felt better after receiving a massage.

$$n = 200 \quad ; \quad x = 165 \quad \Rightarrow \quad \hat{p} = \frac{x}{n} = \frac{165}{200} = 0.825$$

$n\hat{p} = 200(0.825) = 165 > 5$
$n(1-\hat{p}) = 200(1-0.825) = 35 > 5$

a. 15% confidence:

$$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \Rightarrow \quad 0.825 \pm 1.44\sqrt{\frac{0.825(1-0.825)}{200}}$$

$$0.825 \pm 0.0387$$

$$0.7863 < p < 0.8637$$

Interpretation:

With repeated sampling, we a 85% confident that the actual percentage of patients who felt better after receiving a massage is between 78.63% and 86.37%

b. $\hat{p} = 0.84$

$$n = \hat{p}(1-\hat{p})\left(\frac{Z_{\alpha/2}}{E}\right)^2 = 0.84(1-0.84)\left(\frac{1.44}{0.05}\right)^2 = 111.4767$$

$\Rightarrow$ 111 patients need to be included in the sample.

c. No preliminary estimate $\therefore \hat{p} = 0.50$

$$n = \hat{p}(1-\hat{p})\left(\frac{Z_{\alpha/2}}{E}\right)^2 = 0.50(1-0.50)\left(\frac{1.44}{0.05}\right)^2 = 207.36$$

$\Rightarrow$ 207 patients need to be included in the sample.

Example 4

A random sample of 53 suspension helmets used by motorcyclists and race-car drivers was subjected to an impact test, and some damage was observed on 18 of these helmets.

a. Find a 97% two-sided confidence interval on the true proportion of helmets that would show damage from this test.

b. In a previous study, it was found that 37% of the helmets showed some damage after being subjected to the impact test. How many helmets must be tested in order to be 97% confident that the error in estimating p is at most 0.10?

c. How large must the sample be if we wish to be at least 97% confident that the error in estimating p is less than 0.02 regardless of the true value of p?

d. Find a 97% lower-confidence bound for the true proportion of helmets that showed some damage from the impact test.

Solution

Let x = the number of patient who felt better after receiving a massage.

$$n = 53 \quad ; \quad x = 18 \quad \Rightarrow \quad \hat{p} = \frac{x}{n} = \frac{18}{53} = 0.3396$$

$n\hat{p} = 53(0.3396) = 18 > 5$
$n(1-\hat{p}) = 53(1-0.3396) = 35 > 5$

a. 15% confidence:

$$\hat{p} \pm Z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \quad \Rightarrow \quad 0.3396 \pm 2.17\sqrt{\frac{0.3396(1-0.3396)}{53}}$$

$$0.3396 \pm 0.1412$$

$$0.1984 < p < 0.4808$$

Interpretation:

With repeated sampling, we a 97% confident that the true proportion of helmets that show damage from this test is between 19.84% and 48.08%

b. $\hat{p} = 0.37$

$$n = \hat{p}(1-\hat{p})\left(\frac{Z_{\alpha/2}}{E}\right)^2 = 0.37(1-0.37)\left(\frac{2.17}{0.15}\right)^2 = 109.7644$$

$\Rightarrow$ 110 helmets are needed in the sample.

c. No preliminary estimate $\therefore \hat{p} = 0.50$

$$n = \hat{p}(1-\hat{p})\left(\frac{Z_{\alpha/2}}{E}\right)^2 = 0.50(1-0.50)\left(\frac{2.17}{0.10}\right)^2 = 117.7225$$

$\Rightarrow$ 118 helmets need to be included in the sample.

d. 97% LCB

$$\hat{p} - Z_\alpha\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \leq p \quad \Rightarrow \quad 0.3396 \pm 1.88\sqrt{\frac{0.3396(1-0.3396)}{53}} \leq p$$

$$0.3396 - 0.1222 \leq p$$

$$0.2174 \leq p$$

Interpretation:

With repeated sampling, we a 97% confident that at least 21.74% of all helmets are damaged in the impact test.