

Chapter 19, Day 2

Confidence Intervals for Proportions

Page 456, #12 and/or 16

12. Contaminated chicken, second course.

- a) **Independence assumption:** It is very important that the researchers kept the chicken samples separated. Otherwise, *salmonella* could be transmitted from sample to sample.
Randomization condition: It's not clear how the sample was chosen. We will assume the sample from 23 states is at least representative of all broiler chicken sold.
10% condition: 525 is far less than 10% of all broiler chickens.
Success/Failure condition: $n\hat{p} = 79$ and $n\hat{q} = 446$ are both greater than 10, so the sample is large enough.

- b) Since the conditions are met, we can use a one-proportion z-interval to estimate the percentage of broiler chicken packages that are infected with *salmonella*.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} = (0.15) \pm 1.960 \sqrt{\frac{(0.15)(0.85)}{525}} = (0.12, 0.18)$$

- c) We are 95% confident that between 12% and 18% of all broiler chicken sold in U.S. food stores is infected with *salmonella*.

16. Take the offer.

- a) **Randomization condition:** Offers were sent to a random sample of 50,000 cardholders.
10% condition: We assume that there are more than 500,000 cardholders.
Success/Failure condition: $n\hat{p} = 1,184$ and $n\hat{q} = 48,816$ are both much greater than 10, so the sample is large enough.

$$\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} = \left(\frac{1,184}{50,000} \right) \pm 1.960 \sqrt{\frac{\left(\frac{1,184}{50,000} \right) \left(\frac{48,816}{50,000} \right)}{50,000}} = (0.0223, 0.0250)$$

We are 95% confident that the between 2.23% and 2.5% of all cardholders would register for double miles.

- b) The confidence interval gives the set of plausible values with 95% confidence. Since 2% is outside the interval, there is evidence that the true proportion is above 2%. The campaign should be worth the expense.

Be careful how you interpret a confidence interval.

Page 455, #6

6. More conclusions.

- a) Not correct. This statement implies certainty. There is no level of confidence in the statement.
- b) Not correct. We know that 56% of the spins in this experiment landed heads. There is no need to make an interval for the sample proportion.
- c) Correct.
- d) Not correct. The interval should be about the proportion of heads, not the spins.
- e) Not correct. The interval should be about the proportion of heads, not the percentage of euros.

Interpretation of a CI: “We are $x\%$ confident that the true parameter (in context) is between _____ and _____.”

Interpretation of a confidence level: “ $x\%$ of samples of size n contain the true population parameter (in context).”

When collecting data, plan for inference.

Decide on the desired confidence interval.

Choose your sample size based on this.

Choosing Your Sample Size

- The question of how large a sample to take is an important step in planning any study.
- Choose a Margin or Error (ME) and a Confidence Interval Level.
- The formula requires \hat{p} , which we don't have yet because we have not taken the sample. A good estimate for \hat{p} , which will yield the largest value for $\hat{p}\hat{q}$ (and therefore for n) is 0.50.
- Solve the formula for n (*rounding up*).

$$ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

Page 458, #30

30. Hiring.

a)

$$ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$$
$$0.05 = 2.326 \sqrt{\frac{(0.5)(0.5)}{n}}$$
$$n = \frac{(2.326)^2(0.5)(0.5)}{(0.05)^2}$$
$$n = 542 \text{ businesses}$$

In order to estimate the percentage of businesses planning to hire additional employees within the next 60 days to within 5% with 98% confidence, we would need a sample of at least 542 businesses. All decimals in the final answer must be rounded up, to the next business.

b)

$$ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.03 = 2.326 \sqrt{\frac{(0.5)(0.5)}{n}}$$

$$n = \frac{(2.326)^2(0.5)(0.5)}{(0.03)^2}$$

$$n = 1503 \text{ businesses}$$

In order to estimate the percentage of businesses planning to hire additional employees within the next 60 days to within 3% with 98% confidence, we would need a sample of at least 1503 businesses. All decimals in the final answer must be rounded up, to the next business.

(Alternatively, the margin of error is being decreased to 3/5 of its original size, so the sample size must increase by a factor of 25/9. $542(25/9) \approx 1506$ businesses. A bit off, because 542 was rounded, but close enough!

c)

$$ME = z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$0.01 = 2.326 \sqrt{\frac{(0.5)(0.5)}{n}}$$

$$n = \frac{(2.326)^2(0.5)(0.5)}{(0.01)^2}$$

$$n = 13,526 \text{ businesses}$$

In order to estimate the percentage of businesses planning to hire additional employees within the next 60 days to within 1% with 98% confidence, we would need a sample of at least 13,526 businesses.

(Alternatively, the margin of error has been decreased to 1/5 of its original size, so a sample 25 times as large would be needed. $25(542) = 13,550$. Close enough!

It would probably be very expensive and time consuming to sample that many businesses.

Page 456, #18

Page 458, #34