Chapter 8 Homework

Arnav Patri

April 3, 2022

8.1 Confidence Intervals: The Basics

1. Got shoes? The parameter is the average number of pairs of shoes that female teens own, which is a quantitative value, so the appropriate point estimate is $\bar{x}$.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{607}{20} = 30.35$$

3. Going to the prom The parameter is the proportion of seniors at Tonya's school planning to attend prom, making the appropriate point estimator $\hat{p}$.

$$\hat{p} = \frac{36}{50} = 0.72$$

5. Prayer in school

    a. It can be said with 95% confidence that the true proportion of U.S. adults that favor an amendment that would permit organized prayer in public schools is within the interval $(0.63, 0.69)$.

    b. The point estimate $\hat{p}$ is in the middle of the confidence interval, making it the average of the bounds.

$$\hat{p} = \frac{0.63 + 0.69}{2} = 0.66$$

    c. It is not accurate to say that two-thirds of U.S. adults favor this amendment based on this poll, as two-thirds is equal to $0.6\overline{7}$, and values lower than this appear in the confidence interval.

7. Bottling cola

    a. 12 is contained within the confidence interval, so it does not provide convincing evidence 12 is not the true mean.

    b. 12 is only one of the possible values of afforded by the interval, so there is not convincing evidence that it is the true mean.

9. Shoes

    a. There is a 95% chance that the difference between the averages number of pairs of shoes owned by girls and boys in the school is contained within the interval $(10.8, 26.5)$.

    b. Evidence that there is indeed a difference in the average number of pairs of shoes owned by girls and boys within the school, as 0 is not contained within $(10.8, 26.5)$.

11. More prayer in school Over many random samples of size 172, the true proportion of U.S. adults that favor an amendment that would allow organized prayer in public schools will be captured within the confidence interval 95% of the time.

15. How confident? Of the 25 confidence intervals, only 4 did not contain the mean, so the confidence level is likely $(25 - 4)/25$, making it 84%. It is therefore most likely that the confidence level used was the value closest to this, 80%.

23. A larger confidence interval means that there will be a wider range of results, so there will be a higher change of the true value being contained in the interval. The answer is therefore **b**.

24. Increasing the sample size reduces the the standard deviation of the sample, as it is inversely proportional to the square root of the sample size. This in turn reduces the standard error of the statistic which is proportional to the standard deviation of the sample, which results in the margin of error being reduced as well due to its proportionality to the standard error. The size of the confidence interval is determined by the margin of error, so it is narrowed. The confidence level is the same, though, so the changes of failing to capture the parameter remain constant, making the answer **e**.

25. The margin of error does not account for any sort of bias, so the answer is **e**.

26. A confidence level of 95% means that there is a 95% chance of the population parameter being captured in the confidence interval. It can therefore be said that over many samples, the confidence intervals will capture the population parameter 95% of the time, making the answer **c**.

8.2 Estimating a Population Proportion

29. Rating school food The sample is random, as the it was an SRS, so the randomness condition is met. 10% of the population is 17.5, which is far less than the sample size of 50, so the 10% condition is not met. There were 14 successes, and $175 - 14 = 161$, so there were 161 failures. As both of these figures are greater than ten, the Large Counts condition is met.

31. Salty chips The sample is stated to have been random, so the randomness condition is met. 25 is 10% of 250, which is less than one thousand. As the population is comprised of thousands of bags, the 10% condition is met. There were 3 successes, which is less than 10, so the Large Counts condition is not met.

33. The 10% condition

    a. The 10% condition checks for independence between trials when sampling without replacement. This is important because the formulas used are only valid when independence can be assumed.

    b. If the 10% condition is violated, then when sampling without replacement, independence between trials cannot be assumed, as after one trial, the population size will decrease by 1, changing the population proportion. When the 10% condition is not met, this effect is significant.

35. Selling online

a.

$$z^* = -\operatorname{invNorm}\left(\text{area}: \frac{1 - C\%}{2} = \frac{1 - 0.98}{2}, \mu : 0, \sigma : 1, \text{Tail}: \text{LEFT}\right) \approx 2.326$$

b.

$$\hat{p} = \frac{914}{4579} \approx 0.2$$

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{0.2(1 - 0.2)}{4569}} \approx 0.006$$

$$ME = z^* s_{\hat{p}} \approx (2.326)(0.006) \approx 0.014$$

$$\text{confidence interval} = \hat{p} \pm ME \approx 0.2 \pm 0.014 \approx (0.186, 0.213)$$

c. It can be said with 95% confidence that the interval $(0.186, 0.213)$ contains the true proportion of all American adults who would report having earned money by selling something online in the previous year.

37. More online sales

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{\frac{914}{4579}\left(1 - \frac{914}{4579}\right)}{4579}}$$

Over many random samples of size 4579 taken from this population, $\hat{p}$ will differ from p by an average of about about 0.006.

39. Going to the prom

a. The population is the seniors of Tonya's school while the parameter of interest is the proportion of those students that are planning to go to prom.

b. The random condition is met, as it is stated that the 50 students were selected in an SRS. The 10% condition is met, 10% of 750 is 75, which is greater than the sample size of 50. The Large Counts condition is met, as there were 36 successes and the results were binary, so the number of successes is equal to the sample size minus the number of successes, and $50 - 36 = 14$, so both the number of successes and failures are greater than 10.

c.

$$z^* = -\operatorname{invNorm}\left(\text{area}: \frac{1 - C\%}{2} = \frac{1 - 0.9}{2}, \mu : 0, \sigma : 1, \text{Tail}: \text{LEFT}\right) \approx 1.645$$

$$\hat{p} = \frac{36}{50} = 0.72$$

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = \sqrt{\frac{0.72(1 - 0.72)}{50}} \approx 0.063$$

$$ME = z^* s_{\hat{p}} \approx (1.645)(0.063) = 0.104$$

$$\text{confidence interval} = \hat{p} \pm ME \approx 0.72 \pm 0.104 = (0.616, 0.824)$$

d. It can be said with 90% confidence that the true proportion of seniors at Tonya's school planning to attend prom is contained within the interval $(0.616, 0.824)$.

41. Video games

$$z^* = -\text{invNorm}\left(\text{area}: \frac{1 - C\%}{2} = \frac{1 - 0.49}{2}, \mu: 0, \sigma: 1, \text{Tail}: \text{LEFT}\right) \approx 0.659$$

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1 - \hat{p})}{n}} = \sqrt{\frac{0.49\,(1 - 0.49)}{2001}} \approx 0.011$$

$$ME = z^* s_{\hat{p}} \approx (0.659)(0.011) \approx 0.007$$

$$\text{confidence interval} = \hat{p} \pm ME \approx 0.49 \pm 0.007 = (0.482, 0.497)$$

43. Age and video games

a. It is not made clear whether the Large Counts condition was met for each individual population, is it is not specified how many from each population are represented in the sample. Had these been the only age groups in the sample, the equations $0.49 = 0.67 p_{18-29} + 0.29 p_{65+}$ and $1 = p_{18-29} + p_{65+}$ could have been used, but there is evidence to indicate that this is the case.

b. The number of adults ages 18-29 that participated in the sample must be less than the sample size, as there is at least one other age group that participated, so the margin of error would be greater than that calculated for all participants in the study, as it is inversely proportional to the square root of the sample size.

45. Food fight

a.

$$z^* = -\text{invNorm}\left(\text{area}: \frac{1 - C\%}{2} = \frac{1 - 0.99}{2}, \mu: 0, \sigma: 1, \text{Tail}: \text{LEFT}\right) \approx 2.576$$

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1 - \hat{p})}{n}} = \sqrt{\frac{0.55\,(1 - 0.55)}{1480}} \approx 0.03$$

$$ME = z^* s_{\hat{p}} \approx (2.576)(0.013) \approx 0.033$$

$$\text{confidence interval} = p^* \pm ME \approx 0.55 \pm 0.033 = (0.517, 0.583)$$

It can be said with 99% confidence that that true proportion of U.S. adults that agree with the statement that "organic produce is better for health than conventionally grown produce" falls within the interval $(0.517, 0.583)$.

b. This interval provides convincing evidence that the majority of U.S. believe that organic produce has health benefits, as the low bound of the confidence interval was 51.7%, which is a majority.

47. Prom totals

$$\text{confidence interval} = \text{nint}\,((\hat{p} \pm ME)N) \approx \text{nint}\,((0.616, 0.824)(750)) = (618, 462)$$

It can be said with 90% confidence that the number of seniors at Tonya's school planning to attend prom is contained within the interval $(618, 462)$.

49. School vouchers

a.

$$z^* = -\operatorname{invNorm}\left(\text{area} : \frac{1 - C\%}{2} = \frac{1 - 0.99}{2}, \mu : 0, \sigma : 1, \text{Tail} : \text{LEFT}\right) \approx 2.576$$

$$0.03 \geq ME = z^* s_{\hat{p}} = 2.576\sqrt{\frac{\hat{p}\,(1 - \hat{p})}{n}} \approx 2.576\sqrt{\frac{0.44\,(1 - 0.44)}{n}}$$

$$n \geq \left\lceil \frac{2.576^2(0.44)(0.56)}{0.03^2} \right\rceil \approx \lceil 1816.487 \rceil = 1817$$

b.

$$0.03 \geq ME = 2.576\sqrt{\frac{\hat{p}(1 - \hat{p})}{n}} = 2.576\sqrt{\frac{0.5(1 - 0.5)}{n}}$$

$$n \geq \left\lceil \frac{2.576^2(0.5)(0.5)}{0.03^2} \right\rceil \approx \lceil 1843.027 \rceil = 1844$$

$$1844 > 1817$$

53. Teens and their TV sets

a.

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1 - \hat{p})}{n}} = \sqrt{\frac{0.64\,(1 - 0.64)}{1028}} \approx 0.15$$

$$ME = z^* s_{\hat{p}} \approx 0.015 z^* = 0.03$$

$$z^* \approx 2.004$$

$$C\% = 1 - 2\,\operatorname{normalcdf}\left(\text{lower} : -\infty, \text{upper} : -z^* \approx -2.004, \mu : 0, \sigma : 1\right) \approx 95.492\%$$

b. Bias may have been introduced that only households already on the Gallup Poll Panel of households were used.

55. It can be said with 95% confidence that the true proportion of American adults that anticipate inheriting money or valuable possessions from a relative is with 0.28 ± 0.03, or $(0.25, 0.31)$

56. The margin of error is dependent on the standard error of the statistic, which cannot be calculated without knowing how increasing the sample size will impact that sample proportion, making the answer **e**.

57.

$$z^* = -\operatorname{invNorm}\left(\text{area} : \frac{1 - C\%}{2} = \frac{1 - 0.95}{2}, \mu : 0, \sigma : 1, \text{Tail} : \text{LEFT}\right) \approx 1.96$$

$$\hat{p} = \frac{317}{400} \approx 0.793$$

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1 - \hat{p})}{n}} \approx \sqrt{\frac{0.793\,(1 - 0.793)}{400}} \approx 0.02$$

$$ME = z^* s_{\hat{p}} \approx (0.793)(0.02) \approx 0.04$$

The margin of error is approximately 0.04, making the answer **d**.

58.

$$\hat{p} = \frac{0.565 + 0.695}{2} = 0.63$$

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.63(1-0.63)}{100}} \approx 0.048$$

$$ME = \frac{0.695 - 0.565}{2} = 0.065$$

$$= z^* s_{\hat{p}} \approx 0.048 z^*$$

$$z^* \approx 1.346$$

$$C\% = 1 - 2 \, \text{normalcdf} \left(\text{lower} : -\infty, \text{upper} : -z^* \approx 1.346, \mu : 0, \sigma : 1 \right) \approx 0.823$$

C% is most about 0.823, so the confidence level is about 82, so the answer is **a**.

8.3   Estimating a Difference in Proportions

61. Don't drink the water!   The randomness condition is not met, as the the populations in their entirety are used. The 10% condition does not apply, as no sampling took place. The Large Counts condition is not met, as there were only 3 successes from the West side.

65.

a.

$$z^* = \text{invNorm} \left(\text{area} : \frac{1-C}{2} = \frac{1-0.99}{2}, \mu : 0, \sigma : 1, \text{Tail} : \text{LEFT} \right) \approx$$

$$\hat{p}_{M-W} = \hat{p}_M - \hat{p}_W = \frac{986}{2253} - \frac{923}{2729} \approx$$

$$s_{\hat{p}_{M-W}} = \sqrt{\frac{\hat{p}_M\left(1-\hat{p}_M\right)}{n_M} + \frac{\hat{p}_W\left(1-\hat{p}_W\right)}{n_W}} = \sqrt{\frac{\frac{289}{2253}\left(1-\frac{289}{2253}\right)}{2253} + \frac{\frac{923}{2729}\left(1-\frac{923}{2729}\right)}{2729}} \approx$$

$$ME = z^* s_{\hat{p}_{M-W}} \approx$$

$$\text{confidence interval} = \hat{p}_{M-W} \pm ME \approx$$