AP Statistics Homework

Arnav Patri

April 27, 2022

Contents

Chapter 8

Estimating Proportions with Confidence

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.

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

b.

$$\begin{aligned} \hat{p} &= \frac{914}{4579} \approx 0.2 \\ s_{\hat{p}} &= \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = \sqrt{\frac{0.2\,(1-0.2)}{4569}} \approx 0.006 \\ \text{ME} &= z^* s_{\hat{p}} \approx (2.326)(0.006) \approx 0.014 \\ \text{confidence interval} &= \hat{p} \pm \text{ME} \approx 0.2 \pm 0.014 \approx (0.186, 0.213) \end{aligned}$$

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.

$$\begin{aligned} z^* &= -\,\text{invnorm}\left(\text{area} : \frac{1-C\%}{2} = \frac{1-0.9}{2}, \mu : 0, \sigma : 1, \text{Tail} : \text{LEFT}\right) \approx 1.645 \\ \hat{p} &= \frac{36}{50} = 0.72 \\ s_{\hat{p}} &= \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = \sqrt{\frac{0.72\,(1-0.72)}{50}} \approx 0.063 \\ \text{ME} &= z^* s_{\hat{p}} \approx (1.645)(0.063) = 0.104 \\ \text{confidence interval} &= \hat{p} \pm \text{ME} \approx 0.72 \pm 0.104 = (0.616, 0.824) \end{aligned}$$

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

$$z^* = -\operatorname{invnorm}\left(\text{area} : \frac{1-C\%}{2} = \frac{1-0.49}{2}, \mu : 0, \sigma : 1, \text{Tail} : \text{LEFT}\right) \approx 0.659$$
$$s_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = \sqrt{\frac{0.49\,(1-0.49)}{2001}} \approx 0.011$$
$$\text{ME} = z^* s_{\hat{p}} \approx (0.659)(0.011) \approx 0.007$$
$$\text{confidence interval} = \hat{p} \pm \text{ME} \approx 0.49 \pm 0.007 = (0.482, 0.497)$$

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 $0.49 = 0.67p_{18-29} + 0.29p_{65+}$ 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.

$$z^* = -\operatorname{invnorm}\left(\text{area} : \frac{1-C\%}{2} = \frac{1-0.99}{2}, \mu : 0, \sigma : 1, \text{Tail} : \text{LEFT}\right) \approx 2.576$$
$$s_{\hat{p}} = \sqrt{\frac{\hat{p}\,(1-\hat{p})}{n}} = \sqrt{\frac{0.55\,(1-0.55)}{1480}} \approx 0.03$$
$$\text{ME} = z^* s_{\hat{p}} \approx (2.576)(0.013) \approx 0.033$$
$$\text{confidence interval} = p^* \pm \text{ME} \approx 0.55 \pm 0.033 = (0.517, 0.583)$$

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

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

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.

$$
\begin{aligned}
z^* &= -\operatorname{invnorm}\left(\text{area}: \frac{1-C\%}{2} = \frac{1-0.99}{2}, \mu: 0, \sigma: 1, \text{Tail}: \text{LEFT}\right) \approx 2.576 \\
0.03 \geq \text{ME} = z^* s_{\hat{p}} &= 2.576\sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n}} \approx 2.576\sqrt{\frac{0.44\left(1-0.44\right)}{n}} \\
n &\geq \left\lceil \frac{2.576^2(0.44)(0.56)}{0.03^2} \right\rceil \approx \lceil 1816.487 \rceil = 1817
\end{aligned}
$$

b.

$$
\begin{aligned}
0.03 \geq \text{ME} = 2.576\sqrt{\frac{\hat{p}(1-\hat{p})}{n}} &= 2.576\sqrt{\frac{0.5(1-0.5)}{n}} \\
n &\geq \left\lceil \frac{2.576^2(0.5)(0.5)}{0.03^2} \right\rceil \approx \lceil 1843.027 \rceil = 1844 \\
1844 &> 1817
\end{aligned}
$$

53. Teens and their TV sets

a.

$$
\begin{aligned}
s_{\hat{p}} &= \sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n}} = \sqrt{\frac{0.64\left(1-0.64\right)}{1028}} \approx 0.15 \\
\text{ME} &= z^* s_{\hat{p}} \approx 0.015 z^* = 0.03 \\
z^* &\approx 2.004 \\
C\% &= 1 - 2\,\text{normalcdf}\left(\text{lower}: -\infty, \text{upper}: -z^* \approx -2.004, \mu: 0, \sigma: 1\right) \approx 95.492\%
\end{aligned}
$$

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.

$$
\begin{aligned}
z^* &= -\operatorname{invnorm}\left(\text{area}: \frac{1-C\%}{2} = \frac{1-0.95}{2}, \mu: 0, \sigma: 1, \text{Tail}: \text{LEFT}\right) \approx 1.96 \\
\hat{p} &= \frac{317}{400} \approx 0.793 \\
s_{\hat{p}} &= \sqrt{\frac{\hat{p}\left(1-\hat{p}\right)}{n}} \approx \sqrt{\frac{0.793\left(1-0.793\right)}{400}} \approx 0.02 \\
\text{ME} &= z^* s_{\hat{p}} \approx (0.793)(0.02) \approx 0.04
\end{aligned}
$$

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

58.

$$
\begin{aligned}
\hat{p} &= \frac{0.565 + 0.695}{2} = 0.63 \\
s_{\hat{p}} &= \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} = \sqrt{\frac{0.63(1-0.63)}{100}} \approx 0.048 \\
\text{ME} &= \frac{0.695 - 0.565}{2} = 0.065 \\
&= z^* s_{\hat{p}} \approx 0.048 z^* \\
z^* &\approx 1.346 \\
C\% &= 1 - 2\,\text{normalcdf}\left(\text{lower} : -\infty, \text{upper} : -z^* \approx 1.346, \mu : 0, \sigma : 1\right) \approx 0.823
\end{aligned}
$$

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.

$$
\begin{aligned}
z^* &= \text{invnorm}\left(\text{area} : \frac{1-C}{2} = \frac{1-0.99}{2}, \mu : 0, \sigma : 1, \text{Tail} : \text{LEFT}\right) \approx \\
\hat{p}_{M-W} &= \hat{p}_M - \hat{p}_W = \frac{986}{2253} - \frac{923}{2729} \approx \\
s_{\hat{p}_{M-W}} &= \sqrt{\frac{\hat{p}_M\left(1-\hat{p}_M\right)}{n_M} + \frac{\hat{p}_W\left(1-\hat{p}_W\right)}{n_W}} = \sqrt{\frac{\frac{289}{2253}\left(1-\frac{289}{2253}\right)}{2253} + \frac{\frac{923}{2729}\left(1-\frac{923}{2729}\right)}{2729}} \approx \\
\text{ME} &= z^* s_{\hat{p}_{M-W}} \approx \\
\text{confidence interval} &= \hat{p}_{M-W} \pm \text{ME} \approx
\end{aligned}
$$

Chapter 9

Testing Claims about Proportions

9.1 Significance Tests: The Basics

1. No homework? $H_0 : p = 0.75$ and $H_a : p < 0.75$ where p is the true proportion of students at Mr.Tabor's school that completed their math homework last night.

3. How about juice? $H_0 : \mu = 180$ ml and $H_a : \mu \neq 180$ ml where μ is the true mean amount of liquid in a bottle (dispensed by the machine) in milliliters.

4. Attitudes $H_0 : \mu = 115$ and $H_a : \mu > 115$ where μ is the true mean score on the SSHA for students at the teacher's college that are over the age of 30.

5. Cold cabin? $H_0 : \sigma = 3°\text{F}$ and $H_a : \sigma > 3°\text{F}$ where σ is the true standard deviation of the temperature allowed by the thermostat in degrees Fahrenheit.

7. Stating hypotheses

a. The null hypothesis must be a statement of equality while the alternative hypothesis must be an inequality; $H_0 : p = 0.37$; $H_a : p > 0.37$.

b. Hypotheses must always make predictions regarding a population parameter rather than a sample statistic; $H_0 : \mu = 3000$ grams; $H_a : \mu < 3000$ grams.

9. No homework?

a. If $H_0 : p = 0.75$ is true, then 75% of all students at Mr.Tabor's school completed their math homework last night.

b. Assuming that $H_0 : p = 0.75$ is true, the probability that $\hat{p} \leq 0.68$ for a random sample is 12.65%.

10. Attitudes

a. If $H_0 : \mu = 115$ is true, then the true mean score on the SSHA for students at the teacher's college that are over the age of 30 is 115.

b. Assuming that $H_0 : p = 115$ is true, then the probability that $\hat{p} \geq 125.7$ due to sheer random chance, as is the case in this sample, is 1.01%.

13. Interpreting a P-value The interpretation did not include the assumption that $H_0 : \mu = 100$ is true or the inequality $\mu > 100$.

15. No homework At a confidence level of $\alpha = 0.05$, there is not satisfactory evidence supporting $H_a : p < 0.75$, as the P-level of 0.1265 is greater than α, so the null hypothesis $H_0 : p = 0.75$ cannot be disregarded.

16. Attitudes At a confidence level of $\alpha = 0.05$, there is satisfactory evidence supporting the claim that the average SSHA score for students above the age of 30 is higher, as the P-level of 0.0101 is less than α, and the null hypothesis $H_0 : p = 115$ can be disregarded.

19. Making conclusions It was not specified that the P-value was greater than the α, simply that it was large. Additionally, a P-value greater than the significance level does not support H_0, instead not supporting H_a.

21. Heavy bread?

a. $\mu =$ true mean weight of a loaf of bread produced at the bakery (in pounds); $H_0 : \mu = 1$; $H_a : \mu < 1$.

b. The sample mean is less than that predicted by the H_0, which would support H_a.

c. Assuming that H_0 is true, there is an 8.06% chance of this sample's outcome occurring in a random sample.

d. Because the P-level of H_0 against H_a is greater than the $\alpha = 0.01$ significance level, the data does not provide convincing evidence for the hypothesis that the true mean weight of a loaf of bread produced at the bakery is less than 1 lbs, and H_0 cannot be rejected.

23. Opening a restaurant A Type I error would be finding convincing evidence for the true mean income of those living near the potential location being greater than \$85,000 when such is not the case. This would result in the restaurant being opened in a place where the people in the vicinity are unable to afford to eat there, meaning that the restaurant would have to either reduce its prices (by cutting either margins or costs) or close and relocate.
A Type II error would be failing to find convincing evidence of the true mean income of those living close to the potential location being at least \$85,000. This would result in the location being passed up despite being suitable.

25. Awful accidents

a. A Type I error would occur if convincing evidence of the true proportion of calls involving life-threatening injuries over this 6-month period for which emergency personnel took over 8 minutes to arrive being less than 0.22 was found despite this hypothesis being false.
A Type II error would occur if convincing evidence of the true proportion of calls involving life-threatening injuries over this timeframe for which it took an excess of 8 minutes for emergency personnel to arrive being less than 0.22 was not found despite this hypothesis being true.

b. In this case, a Type I error would be more harmful, as it would make it seem as though there was less room for improvement than there really actually is, which will likely result in a reduced drive to improve, potentially resulting in the proportion staying the same or even increasing, resulting in more deaths due to wasted time.

c. As the probability of a Type I error occurring is equal to α and that a Type I error would be more serious than a Type II one, the significance level should be lower than $\alpha = 0.05$.

27. More lefties?

a. $p = 0$ the true proportion of students at Simon's school that are left-handed; $H_0 : p = 0.1$; $H_a : p > 0.1$.

b. The P-value of H_0 against the result of the sample is 24/200, equal to 12%. This means that the probability of receiving the observed results due to sheer chance assuming, that H_0 is true.

c. The P-value is greater than the the assumed confidence level of $\alpha = 0.05$, so the data provided by the survey is not enough to warrant disregarding H_0 and convincingly support the conclusion that the true proportion of students at Simon's school that are left-handed is greater than 0.1.

9.2 Tests About a Population Proportion

35. Home computers The randomness condition is met, as there are Jason's school is large, so there are likely over 600 students at school making the sample size n of 60 less than a tenth of the population, so independence can be assumed, and the sample itself is random. The Large Counts condition is also met, as np_0 and $n(1 - p_0)$ are both greater than 10, at 48 and 12 respectively. A Normal distribution can therefore be used to approximate the sampling distribution of $\hat{p}$.

37. The chips project

a. There are 400 students in the population, so the sample size n of 50 is over 10% of the population, so independence cannot be assumed.

b. As np_0 and $n(1 - p_0)$ are both 25, which is greater than 10, the Large Counts condition is met, so the sampling distribution of $\hat{p}$ is approximately Normal.

39. Home computers

a. The sample proportion is 41/60, which is about 0.6833, which is less than 0.8.

b.

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{41}{60} - 0.8}{\sqrt{\frac{0.8(1-0.8)}{60}}} \approx -2.27$$

$$P\text{-value} = \text{normalcdf (lower} : -\infty, \text{upper} : z \approx -2.27, \mu : 0, \sigma : 1) \approx 0.0116$$

c. As the P-value is less than α, H_0 can be rejected, as there is convincing evidence that the true proportion of all students at Jason's school that own computers is less than 0.8.

41. Significance tests

a.

$$P\text{-value} = \text{normalcdf (lower} : z \approx 2.19, \text{upper} : \infty, \mu : 0, \sigma : 1) \approx 0.0143$$

Assuming that $H_0 : p = 0.5$ is true, there is about a 1.43% chance of having received this result from a random sample.

b. The P-level of 0.0143 is greater than α, so H_0 cannot be rejected, as there is not convincing evidence of the true proportion of being greater than 0.5.

c.

$$\hat{p} = p_0 + z\sqrt{\frac{p_0\,(1 - p_0)}{n}} = 0.5 + 2.19\sqrt{\frac{0.5\,(1 - 0.5)}{200}} \approx 0.5774$$

43. Bullies in middle school

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{\frac{445}{558} - 0.75}{\sqrt{\frac{0.75(1-0.75)}{558}}} \approx 2.591$$

$$P\text{-value} = \text{normalcdf}\,(\text{lower} : z \approx 2.591, \text{upper} : \infty, \mu : 0, \sigma : 1) \approx 0.005$$

The probability of evidence of the sample proportion of middle school students that engage in bullying being at least 445/558 is about 0.005 for a random sample assuming that the true population proportion p is equal to 0.75, which is less than the significance level $\alpha = 0.05$, so there this data provides convincing evidence that p is greater than 0.75.

59. Potato chips If the true proportion p of potatoes with blemishes in a shipment is 0.11, there is a 76.5% chance of convincing evidence being found for $H_a : p > 0.08$.

60. Upscale Restaurant If the true mean income of people living near the location

61. Powerful potatoes

a. Increasing the significance level also increases the test's power.

b. Decreasing the sample size also decreases the test's power.

c. Decreasing the difference between p_a and p, the effect size, decreases makes it less likely that a significant difference will be detected, decreasing the power.

62. Restaurant power

a. Decreasing the sample size decreases the standard deviation, decreasing the range within which values are significant, making it harder to reject H_0, increasing the chances of a Type II error occurring, reducing the power.

b. Reducing the effect size makes a difference in the parameter more difficult to detect, increasing the chances of a Type II error occurring, reducing the power.

c. Increasing the significance level makes it easier to reject H_0, decreasing the chances of a Type II error, increasing the power.

63. Potato power problems

a. Using a higher significance level increases the chance of the null hypothesis being rejected despite the alternative hypothesis being false.

b. Increasing the sample size makes the study more costly to perform.

64. Restaurant power problems

a. Increasing the significance level also increases the probability of a Type I error.

b. Increasing the sample size makes the study more expensive to carry out.

65. Better parking

a. At a significance level of $\alpha = 0.05$, there is a 75% chance of convincing evidence for more than 37% of students at the school approving of the provided parking being found by a random sample if the true population proportion is 0.45.

b.

$$P(\text{Type I Error}) = \alpha = 0.05 \qquad P(\text{Type II Error}) = 1 - \text{power} = 1 - 0.75 = 0.25$$

c. Power can be increased by increasing α or the sample size.

67. Error probabilities and power

a. The power of the test is one minus the probability of making a Type II, making it 0.86.

b. The probability of making a Type I error is equal to the significance level, so it is 0.05.

70.

$$z = \frac{\frac{64}{100} - 0.5}{\sqrt{\frac{\frac{64}{100}\left(1-\frac{64}{100}\right)}{100}}} = \frac{0.64 - 0.5}{\sqrt{\frac{0.64(1-0.64)}{0.36}}100}$$

The answer is therefore **a**.

71. A significance test requires randomness, Large Counts (for approximate Normality), and the 10% condition (for independence), so the answer is **e**.

72.

$$P\text{-value} = 1 - \text{normalcdf}\,(\text{lower} : -2.43, \text{upper} : 2.43, \mu : 0, \sigma : 1) \approx 0.015$$

As the P-value is between 0.01 and 0.05, the answer is **b**.

73. The only interval that doesn't contain 0.3 is $(0, 19, 0.27)$, so the answer is **a**.

74.

$$P(\text{Type II Error}) = 1 - \text{power} = 1 - 0.9 = 0.1$$

The probability of a Type II error is 0.1, so the answer is **b**.

9.3 Tests About a Difference in Proportions

85. Bag lunch?

a. The survey is stated to be an SRS.

$$
\begin{aligned}
n_1 \leq 0.1N_1, n_2 \leq 0.1N_2 \hat{p}_1 - \hat{p}_2 &= \frac{52}{80} - \frac{78}{104} = -0.1 \\
\hat{p}_C &= \frac{52+78}{80+104} \approx 0.707 \\
n_1\hat{p}_C, n_1(1-\hat{p}_C), n_2\hat{p}_C, n_2(1-\hat{p}_C) &\geq 10 \\
s_{\hat{p}_1-\hat{p}_2} = \sqrt{\hat{p}_C\left(1-\hat{p}_C\right)\left(\frac{1}{n_1}-\frac{1}{n_2}\right)} &\approx \sqrt{0.707\left(1-0.707\right)\left(\frac{1}{80}-\frac{1}{104}\right)} \approx 0.068 \\
z &= \frac{\hat{p}_1-\hat{p}_2-\mu_{\hat{p}_1-\hat{p}_2}}{s_{\hat{p}_1-\hat{p}_2}} \approx \frac{-0.1-0}{0.707} \approx -1.471 \\
P\text{-value} &= \text{normalcdf}\left(\text{lower} : -\infty, \text{upper} : z \approx -1.471, \mu : 0, \sigma : 1\right) \approx 0.071
\end{aligned}
$$

As 0.071 is greater than $\alpha = 0.05$, H_0 cannot be rejected, and the data does not provide convincing evidence for H_a.

b. If there is no difference between the proportions of sophomores and seniors at Phoebe's school that bring a bag lunch, there is a 7.1% chance of the sample difference being at most -0.1 in a random sample.

87. Preventing peanut allergies

a. The study is random and the 10% condition is met.

$$
\begin{aligned}
\hat{p}_1 - \hat{p}_2 &= \frac{10}{307} - \frac{55}{321} \approx -0.139 \\
\hat{p}_C &= \frac{10+55}{307+321} \approx 0.104 \\
s_{\hat{p}_1-\hat{p}_2} &= \sqrt{\hat{p}_C\left(1-\hat{p}_C\right)\left(\frac{1}{n_1}-\frac{1}{n_2}\right)} = \sqrt{0.104\left(1-0.104\right)\left(\frac{1}{307}-\frac{1}{321}\right)} \approx 0.024 \\
z &= \frac{\hat{p}_1-\hat{p}_2-\mu_{\hat{p}_1-\hat{p}_2}}{s_{\hat{p}_1-\hat{p}_2}} \approx \frac{-0.139-0}{0.024} \approx -4.256 \\
P\text{-value} &= \text{normalcdf}\left(\text{lower} : -\infty, \text{upper} : z \approx -4.256, \mu : 0, \sigma : 1\right) \approx 0
\end{aligned}
$$

As the P-value is less than the significance level $\alpha = 0.05$, H_0 can be rejected and the data convincingly supports H_a.

b. As H_0 was rejected, only a Type II error could have been made.

c. The study was an experiment, so the results are not generalizable.

d. This confidence interval states with 95% certainty that the difference is between -0.185 and -0.93 while the test simply stated that it was not 0.

89. Preventing peanut allergies

a. The sample size increasing decreases the standard error and therefore the standardized test statistic, making it more likely for the P-value to fall below α, making H_0 more likely to be rejected, decreasing the probability of a Type II error, increasing the power. This would, however, increase the cost of the study considerably.

b. Increasing the α increases the chances of the P-value falling below α, making H_0 more likely to be rejected, decreasing the probability of a Type II error, increasing the power. This would make a Type I error more likely, though.

c. This would eliminate a source of variability, making it easier to reject H_0, increasing the probability of a Type II error and therefore increasing the power. This would limit the scope of inference, though.

93. Texting and driving

a. p_A = proportion of people that received version A that answered "Yes"; p_B = proportion of people that received version B that answered "Yes"; $H_0 : p_A - p_B = 0$; $H_a : p_A - p_B > 0$.

b.

$$\hat{p}_C = \frac{x_A + x_B}{n_A + n_B} = \frac{18 + 14}{25 + 25} = 0.64$$
$$n_A(1 - \hat{p}_C) = 9 < 10$$

The Large Counts condition is not met, so standard error cannot be calculated, so the P-value cannot be calculated.

c.

$$\hat{p}_A - \hat{p}_b = \frac{18}{25} - \frac{14}{25} = 0.16$$
$$P\text{-value} = P(\hat{p}_A - \hat{p}_B \geq 0.16) = \frac{14}{100} = 0.14$$

d. The probability of getting a difference between the proportions of people that received each version that answered "Yes" of at least 0.16 in a random sample is about 0.14. As this value is greater than the significance level $\alpha = 0.05$, H_0 cannot be rejected and the data does not provide convincing evidence for H_a.

95.

$$H_0 : p_M - p_F = 0 \qquad\qquad H_a : p_M - p_F \neq 0$$

The answer is therefore **a**.

96. The P-value was greater than $\alpha = 0.1$, so H_0 cannot be rejected, and there is not convincing evidence supporting H_a, so the answer is **b**.

97.

$$s_{\hat{p}_1 - \hat{p}_2} = \sqrt{\hat{p}_C\,(1 - \hat{p}_C)\left(\frac{1}{n_1} - \frac{1}{n_2}\right)} = \sqrt{\frac{500 + 410}{550 + 484}\left(1 - \frac{500 + 410}{550 + 484}\right)\left(\frac{1}{550} - \frac{1}{484}\right)}$$
$$\approx \sqrt{0.88\,(0.12)\left(\frac{1}{550} - \frac{1}{484}\right)}$$

The answer is therefore **c**.

98. The total number of rats in each group is 10, so it is not possible for the Large Counts condition to be met, making the answer **e**.

Chapter 10

Estimating Means with Confidence

10.1 Estimating a Population Mean

1. Critical values

a.

$$\begin{aligned} \text{df} &= n - 1 = 9 \\ t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.95}{2} = 0.025, \text{df} : 9\right)\right| \approx 2.262 \end{aligned}$$

b.

$$\begin{aligned} \text{df} &= 20 - 1 = 19 \\ t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.99}{2} = 0.005, \text{df} : 19\right)\right| \approx 2.861 \end{aligned}$$

c.

$$\begin{aligned} \text{df} &= 77 - 1 = 76 \\ t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.9}{2} = 0.05, \text{df} : 76\right)\right| \approx 1.665 \end{aligned}$$

3. Weeds among the corn 28 is less than 30 and the data contains outliers, so the Central Limit theorem does not take effect. Normality can therefore not be assumed.

5. Check them all

a. The Randomness condition is not met, as the sample contains all individuals from the same class.
The 10% condition is met, as my school contains more than 32% people, so the sample size of 32 is less than 10% of the population.
The Normality condition met, as the sample size of 32 is greater than 30, so the Central Limit theorem is applicable.

b. The Randomness condition is met, as each individual is randomly selected.
The 10% condition is met, as the city is large, so there were likely over 1000 home sales in the previous 6 months, so the sample size of 100 is less than 10% of the population size. Independence is therefore justified.

7. Blood pressure

$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}} = \frac{9.3}{\sqrt{27}} \approx 1.79$$

The means of random samples of size 27, the mean seated systolic blood pressure is expected to vary by about 1.79 from the true mean.

9. Bone loss by nursing mothers

a.

$$\mu = \text{mean \% change in BMC of breast-feeding mothers over 3 months of breast-feeding}$$

As a single parameter is concerned, a 1-sample t interval should be constructed.
The Randomness condition is met by the fact that this was a random sample.
The 10% condition is met, as there are far more than 470 breast-feeding mothers, so the sample size of 47 is less than 10% of the population size.
The Normality condition is met by the Central Limit theorem, which is applicable due to the sample size being 47, which is greater than 30.

$$\begin{aligned}
\text{df} &= n - 1 = 47 - 1 = 46 \\
t^* &= \left| \text{invT}\left(\text{area} : \frac{1 - 0.99}{2} = 0.005, \text{df} : 46\right) \right| \approx 2.687 \\
s_{\bar{x}} &= \frac{s_x}{\sqrt{n}} = \frac{2.506}{\sqrt{47}} \approx 0.366 \\
ME &= t^* s_{\bar{x}} \approx 2.687 \times 0.366 \approx 0.982 \\
\text{confidence interval} &= \bar{x} \pm \text{ME} \approx -3.587 \pm 0.982 \approx (-4.569, -2.605)
\end{aligned}$$

It can be said with 99% confidence that the true mean value μ of the percent change in BMC over 3 months of breast-feeding is contained within the interval $(-4.569, -2.605)$.

b. The 99% confidence interval provides convincing evidence of breast-feeding mothers on average losing bone mineral, as every value contained within the interval was negative.

11. America's favorite cookie

a.

$$\mu = \text{mean weight of an Oreo cookie in grams}$$

There is a single numerical value being tested, so a 1-sample t interval should be constructed.
The Randomness condition is met, as the sample is stated to be random.
The 10% condition is met, as more than 360 Oreo cookies exist, so the sample size of 36 is less than the population size.
The Normality condition is met, as the sample size of 36 is greater than 30, so the Central Limit

theorem is applicable.

$$\begin{aligned}
\text{df} &= n-1 = 36-1 = 35 \\
t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.9}{2} = 0.05, \text{df} : 35\right)\right| \approx 1.6896 \\
s_{\bar{x}} &= \frac{s_x}{\sqrt{n}} = \frac{0.0817}{\sqrt{36}} \approx 0.0136 \\
ME &= t^* s_{\bar{x}} \approx 1.6896 \times 0.0136 \approx 0.023 \\
\text{confidence interval} &= \bar{x} \pm \text{ME} \approx 11.3921 \pm 0.023 \approx (11.3691, 11.4151)
\end{aligned}$$

It can be said with 99% confidence that the true mean value μ of the weight of an Oreo cookie is contained within the interval $(1.3691, 11.4151)$.

b. Over many random samples of size 36 taken from the same population of all Oreo cookies, the true mean will be contained within 90% of the 90% confidence interval constructed about each sample mean.

13. Pepperoni pizza

$\mu =$ mean number of pepperonis on a large pepperoni pizza from their favorite pizza restaurant

There is a single numerical value, so a 1-sample t interval should be constructed.
The Randomness condition is met, as each time is stated to be random.
The 10% condition is met, as the restaurant likely made over 100 pizzas over the week, so the sample size of 10 is less than 10% of the population size. The Normality condition is met, as the data lacks strong skew or outliers.

$$\begin{aligned}
\text{df} &= n-1 = 10-1 = 9 \\
t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.95}{2} = 0.025, \text{df} : 9\right)\right| \approx 2.262 \\
\bar{x} &= \frac{\sum x_i}{n} = \frac{374}{10} = 37.4 \\
s_x &= \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} \approx 7.662 \\
s_{\bar{x}} &= \frac{s_x}{\sqrt{n}} \approx \frac{7.662}{\sqrt{10}} \approx 2.423 \\
\text{ME} &= t^* s_{\bar{x}} \approx 2.262 \times 2.423 \approx 5.481 \\
\text{confidence interval} &= \bar{x} \pm \text{ME} \approx 37.4 \pm 1.733 \approx (32.919, 42.881)
\end{aligned}$$

It can be said with 95% confidence that the true mean number of pepperoni slices on a large pepperoni pizza from Melissa and Madeline's favorite pizza restaurant is contained within the interval $(32.919, 42.881)$.

15. A plethora of pepperoni

a. The sample size of 10 is less than 30, so the Central Limit theorem could not be used to justify Normality.

b. The interval contains 40, so there is not convincing evidence of the average number of pepperonis being less than 40.

c. The margin of error could be reduced either by increasing the sample size, reducing s_x and $s_{\bar{x}}$, or decreasing the confidence level, reducing t^*. The former would be more expensive while the latter would make the interval less likely to contain the true mean number of pepperonis.

17. Estimating BMI

$$\begin{aligned}
z^* &= \left|\text{invnorm}\left(\text{area} : \frac{1-0.99}{2} = 0.005, \mu : 0, \sigma : 1, \text{Tail} : \text{LEFT}\right)\right| \approx 2.576 \\
\text{ME} &\geq z^* \sigma \bar{x} = z^* \frac{\sigma}{\sqrt{n}} \\
n &\geq \left(\frac{z^* \sigma}{\text{ME}}\right)^2 \approx \left(\frac{2.576 \times 7.5}{1}\right) \approx 373.213 \\
n &= \lceil 373.213 \rceil = 374
\end{aligned}$$

19. Willows in Yellowstone

a.

$$\begin{aligned}
s_{\bar{x}} &= \frac{s_x}{\sqrt{n}} \\
s_x &= s_{\bar{x}} \sqrt{n} = 19.03\sqrt{23} \approx 91.265
\end{aligned}$$

b.

$$C = \text{tcdf}\left(\text{lower} : -1, \text{upper} : 1, \text{df} : 23 - 1 = 22\right) \approx 0.672$$

21.

z can only be used when the population standard deviation is already known, so the answer is **b**.

22.

$$\begin{aligned}
\text{df} &= n - 1 = 23 - 1 = 22 \\
t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.98}{2} = 0.01, \text{df} : 22\right)\right| \approx 2.508
\end{aligned}$$

The answer is therefore **e**.

23.

The margin of error is minimized when $s_{\bar{x}}$ and t^* are minimized. These are in turn minimized when n is maximized and C is minimized so the answer is **b**.

24.

The most significant detriment to the validity of a t interval listed is the data containing a clear outlier, as that would mean that the data fails to satisfy the Normality condition. (The Central Limit theorem is not applicable, as the sample size of 24 is less than the 30 required for it to take effect.) The answer is therefore **a**.

10.2 Estimating a Difference in Means

31. Is red wine better than white wine

a. Both distributions are slightly skewed left without any outliers. The center of the white distribution appear lower than that of the red, while its minimum and maximum are both lower than those of the red distribution.

b.

$$\begin{aligned}
\mu_R &= \text{mean percent change in polyphenol levels for healthy men that drink red wine} \\
\mu_W &= \text{mean percent change in polyphenol levels for healthy men that drink white wine} \\
\mu_{\text{diff}} &= \mu_{R-W}
\end{aligned}$$

As there are two quantitative values, a two-sample t interval should be constructed.
The Randomness condition is met, as the assignment was random.
The 10% condition is met, as there are over 900 healthy men that drink red wine, and the same is true of white wine, which means that the sample size of 9 is less than 10% of each population.
The Normality condition is met, as both distributions have little skew and lack outliers.

$$\begin{aligned}
s_{\bar{x}_W-\bar{x}_R} &= \sqrt{\frac{s_W^2}{n_W}+\frac{s_R^2}{n_R}} \approx \sqrt{\frac{2.517^2}{9}+\frac{3.292^2}{9}} \approx 1.381 \\
\text{df} &= \frac{s^4_{\bar{x}_R-\bar{x}_W}}{\frac{s_R^4/n_R^2}{n_R-1}+\frac{s_W^4/n_W^2}{n_W-1}} \approx \frac{1.381^4}{\frac{2.517^4/9^2}{9-1}+\frac{3.292^4/9^2}{9-1}} \approx 14.971 \\
t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.9}{2} = 0.05, \text{df} :\approx 14.971\right)\right| \approx 1.753 \\
\text{ME} &= t^* s_{\bar{x}_W-\bar{x}_R} \approx 1.753 \times 1.381 \approx 2.422 \\
\bar{x}_{\text{diff}} &= \bar{x}_W - \bar{x}_R \approx 5.5 - 0.233 = 5.267 \\
\text{confidence interval} &= \bar{x}_{\text{diff}} \pm \text{ME} \approx 5.267 \pm 2.422 \approx (2.845, 7.689)
\end{aligned}$$

It can be said with 90% confidence that the true mean difference between the percent change in polyphenol levels of healthy men drinking red and white wine μ_{R-W} is contained within the interval $(2.845, 7.689)$.

33. Paying for college

a. Earnings cannot be negative, yet the standard deviations of each group is almost as large as their means. Regardless, as the sample sizes of both groups is greater than 30, the Central Limit theorem applies, so Normality can be justified.

b.

$$\begin{aligned}
\mu_M &= \text{mean earnings of male university students with summer jobs in \$} \\
\mu_W &= \text{mean earnings of females university students with summer jobs in \$} \\
\mu_{\text{diff}} &= \mu_{M-W}
\end{aligned}$$

The Randomness condition is met, as the sample was random and independent.
The 10% condition is met, as there are more than $6,750$ university students with summer jobs that are male/female, so the sample sizes of 675 and 621 are both less than 10% of their respective populations.
The Normality condition is met, as stated in part a.

$$
\begin{aligned}
s_{\bar{x}_M-\bar{x}_W} &= \sqrt{\frac{s_M^2}{n_M}+\frac{s_W^2}{n_W}} = \sqrt{\frac{1368.37^2}{675}+\frac{1037.46^2}{621}} \approx 67.136 \\
\text{df} &= \frac{s_{M-W}^4}{\frac{s_M^4/n_M^2}{n_M-1}+\frac{s_W^4/n_W^2}{n_W-1}} \approx \frac{67.136^4}{\frac{1368.37^4/675^2}{675-1}+\frac{1037.46^4/621^2}{621-1}} \approx 1249.213 \\
t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.9}{2} = 0.05, \text{df} :\approx 1249.213\right)\right| \approx 1.646 \\
\text{ME} &= t^* s_{M-W} \approx 1.646 \times 67.136 \approx 110.51 \\
\bar{x}_{\text{diff}} &= \bar{x}_M - \bar{x}_W = 1884.52 - 1360.39 \approx 524.13 \\
\text{confidence interval} &= \bar{x}_{\text{diff}} \pm \text{ME} \approx 523.13 \pm 110.51 \approx (413.62, 634.64)
\end{aligned}
$$

It can be said with 90% confidence that the true mean dollar difference between the earnings of male and female university students with summer jobs μ_{M-W} is contained within the interval $(413.62, 634.64)$.

35. Reaction times

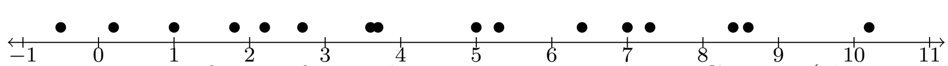
a. The confidence interval does not provide convincing evidence of a difference in the true mean reaction times of athletes and non-athletes, as its contains 0, meaning that there is a chance of the difference being 0.

b. It does not provide convincing evidence of a difference not existing, as not only are nonzero values contained, but the interval is centered above zero.

c. To decrease the width of a confidence interval, the margin of error must be decreased, so n must be increased, reducing the standard error, or the confidence level must be decreased, reducing t^*. The former is makes data collection more time-consuming and/or expensive while the latter makes the true mean less likely to be contained within the interval.

37. Groovy tires

a.



Difference in estimates of wear for each tire using Weight vs Groove (thousands of miles)

b. The distribution appears to be centered around 7 rather than 0, implying that the two methods give different estimates of tire wear on average.

c.

$$\bar{x}_{\text{diff}} \approx 4.556 \qquad\qquad s_{\text{diff}} \approx 3.226$$

Using the weight method to estimate wear yields an estimate on average 4.556 thousand miles greater than that using the groove method.

41. Groovy tires

$$\begin{aligned}
\mu_W &= \text{mean estimate of wear using the weight technique (thousands of miles)}\\
\mu_G &= \text{mean estimate of wear using the groove technique (thousands of miles)}\\
\mu_{\text{diff}} &= \mu_{W-G}
\end{aligned}$$

The values for the differences can be individually analyzed, so a 1-sample t test can be used.
The Randomness condition is met, as the samples are random.
The 10% condition is met, as each technique has been used over 160 times, so the sample sizes of 16 is less 10% of each population size.
The Normality condition is met, as the distribution of the difference is symmetrical without outliers.

$$\begin{aligned}
\text{df} &= n-1 = 16-1 = 15\\
t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.95}{2} = 0.025, \text{df} : 15\right)\right| \approx 2.131\\
s_{\bar{x}_W-\bar{x}_G} &= \frac{s_{\text{diff}}}{\sqrt{n}} \approx \frac{3.226}{\sqrt{16}} \approx 0.807\\
\text{ME} &= t^* s_{\bar{x}_W-\bar{x}_G} \approx 2.131 \times 0.807 \approx 1.719\\
\text{confidence interval} &= \bar{x}_{\text{diff}} \pm \text{ME} \approx 4.556 \pm 1.719 \approx (2.837, 6.275)
\end{aligned}$$

It can be said with 95% confidence that the true mean difference between the wears predicted by the weight and groove methods (in miles) contained within the interval $(2.837, 6.275)$.

43. Does playing piano make you smarter?

a.

$$\begin{aligned}
\mu_B &= \text{mean reasoning score of preschool children before 6 months of piano lessons}\\
\mu_A &= \text{mean reasoning score of preschool children after 6 months of piano lessons.}\\
\mu_{\text{diff}} &= \mu_{A-B}
\end{aligned}$$

As the difference is being dealt with as a single value rather than the difference of two means, a 1-sample t test should be performed.
The Randomness condition is met, as sample was taken randomly.
The 10% condition is met, as there are more than 340 preschoolers, so the sample size of 34 is less than 10% of the population size.
The Normality condition is met, as 34 is greater than 30, so the Central Limit theorem can be applied.

$$\begin{aligned}
\text{df} &= n-1 = 34-1 = 33\\
t^* &= \left|\text{invT}\left(\text{area} : \frac{1-0.9}{2} = 0.05, \text{df} : 33\right)\right| \approx 1.692\\
s_{\bar{x}_A-\bar{x}_B} &= \frac{s_{\text{diff}}}{\sqrt{n}} = \frac{3.055}{\sqrt{33}} \approx 0.532\\
\text{ME} &= t^* s_{\bar{x}_A-\bar{x}_B} \approx 1.692 \times 0.532 \approx 0.9\\
\text{confidence interval} &= \bar{x}_{\text{diff}} \pm \text{ME} \approx 3.618 \pm 0.9 \approx (2.718, 4.518)
\end{aligned}$$

It can be said with 90% confidence that the true mean difference change in reasoning scores of preschool children after 6 months of piano lessons is contained within the interval $(2.718, 4.518)$.

b. The 90% confidence interval provides convincing evidence that 6 months of piano lessons may result in an increase in the average reasoning scores of preschoolers, as it only contains positive values.

45. Chewing gum

a. The reason that the is paired is that every individual had data collected for both treatments.

b. The Randomness condition is met, as the volunteers were randomly assigned.
The 10% condition is met, as over far more than 300 people exist, so the sample sizes of 30 are far less than 10% of the population size.
The Normality condition is met, as a sample size of 30 is exactly large enough for the Central Limit theorem to take effect.

c. It can be said with 95% confidence that the true mean difference in the number of words remembered when using or not using gum is contained within the interval $(-0.67, 1.54)$.

d. The interval does not provide convincing evidence of chewing gum helping with short-term memory, as it contains 0, implying that there may be no difference.

49.
A 1-sample t interval is better for paired data, so the answer is **d**.

50.

$$\begin{aligned}
\text{df} &= n - 1 = 10 - 1 = 9 \\
t^* &= \left| \text{invT}\left(\text{area} : \frac{1 - 0.95}{2} = 0.0025, \text{df} : 9\right) \right| \approx 2.262 \\
s_{\bar{x}_A - \bar{x}_B} &= \frac{s_{A-B}}{\sqrt{n}} = \frac{0.83}{\sqrt{10}} \\
\text{ME} &= t^* s_{\bar{x}_A - \bar{x}_B} \approx 2.262\left(\frac{0.83}{\sqrt{10}}\right) \\
\text{confidence interval} &= \bar{x}_{\text{diff}} \pm \text{ME}
\end{aligned}$$

The answer is therefore **e**.

51.
0 being contained within the interval means that there is a chance that the true difference is 0 and there being no difference, so the answer is **b**.

52.

$$\begin{aligned}
s_{\bar{x}_A - \bar{x}_J} &= \sqrt{\frac{s_A^2}{n_A} + \frac{s_J^2}{n_J}} \approx \sqrt{\frac{2.4^2}{30} + \frac{2.26^2}{30}} \approx 0.602 \\
\text{df} &= \frac{s^4_{\bar{x}_A - \bar{x}_J}}{\frac{s_A^4/n_A^2}{n_A - 1} + \frac{s_J^4/n_J^2}{n_J - 1}} \approx \frac{0.602^4}{\frac{2.4^4/30^2}{30-1} + \frac{2.26^4/30^2}{30-1}} \approx 57.792 \\
t^* &= \left| \text{invT}\left(\text{area} : \frac{1 - 0.95}{2} = 0.025, \text{df} :\approx 57.792\right) \right| \approx 2.002 \\
\text{ME} &= t^* s_{\bar{x}_A - \bar{x}_J} \approx 2.002\sqrt{\frac{2.4^2}{30} + \frac{2.26^2}{30}} \\
\text{confidence interval} &= \bar{x}_{\text{diff}} \pm \text{ME}
\end{aligned}$$

The answer is closest to **d**.

Chapter 11

Testing Claims About Means

11.1 Tests About a Population Mean

1. Attitudes
The Randomness condition is met, as an SRS is being performed.
The 10% condition is met, as the sample size of 45 is less than 100, making it less than 10% of the population size.
The Normality condition is met, as the sample size of 45 is greater than 30, meaning that Central Limit Theorem is applicable to justify Normality.

5. Two-sided test

a.

$$\mu = \text{mean battery life of tablet computer when playing videos (hrs)}$$

$$H_0 : \mu = 11.5 \qquad\qquad H_a^{:}\mu < 11.5$$

b. The randomness condition is met, as the tablets are selected randomly.
The 10% condition is met, as over 200 tablets are likely produced per day, making the sample size of 20 less than 10ôf the population size.
The Normality condition is not met, as the sample distribution is heavily skewed right and the Central Limit Theorem does not take effect counteract this, as the sample size of 20 is less than 30.

7. Attitudes

a.

$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}} = \frac{29.8}{\sqrt{45}} \approx 4.442$$
$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} \approx \frac{125.7 - 115}{4.442} \approx 2.409$$

b.

$$P\text{-value} = P(T > t) = \text{tcdf}\left(\text{lower} : t \approx 2.409, \text{upper} : \infty, \text{df} : n - 1 = 45 - 1 = 44\right) \approx 0.01$$

As the P-value of 0.01 is less than the significance level $\alpha = 0.05$, the data provides convincing evidence that students above the age of 30 at the teacher's school have higher than average SSHA scores, and the hypothesis that they have average scores can be rejected.

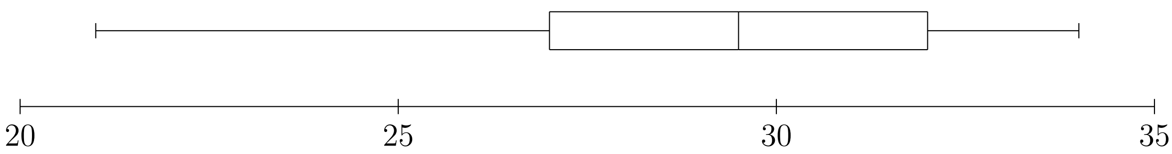
9. Construction zones

a.

$$\mu = \text{average speed of drivers in 25 mph construction zone (mph)}$$

$$H_0 : \mu = 25 \qquad\qquad H_a : \mu > 25$$

It is being tested whether a mean's true value is greater than its null value, a 1-sided 1-sample t test should be performed.
The randomness condition is met, as the sample was random.
The 10% condition is met, as more than 100 drivers likely passed through the 25 mph construction zone, so the sample size of 10 is less than 10% of the population size.



The modified box plot of the distribution is only slightly skewed and lacks any outliers, indicating approximate Normality.

$$\bar{x} = \frac{\sum x_i}{n} = \frac{278}{10} = 27.8$$
$$s_x = \sqrt{\frac{\sum(x_i - \bar{x})}{n-1}} \approx 3.938$$
$$s_{\bar{x}} = \frac{s_x}{\sqrt{n}} \approx \frac{3.938}{\sqrt{9}} \approx 1.245$$
$$t = \frac{\bar{x} - \mu_0}{s_{\bar{x}}} \approx \frac{27.8 - 25}{1.245} \approx 3.051$$
$$P\text{-value} = P(T > t) = \text{tcdf}\left(\text{lower} : t \approx 3.051, \text{upper} : \infty, \text{df} : n - 1 = 10 - 1 = 9\right) \approx 0.007$$

As the P-value of about 0.007 is less than the significance level $\alpha = 0.01$, H_0 can be rejected. The data provides convincing evidence that the mean speed of drivers in the 25 mph construction zone is greater than 25 mph.

b. A low significance level that the power is lower, so the probability of a Type II error is higher. Making this error here would mean concluding that the data does not provide convincing evidence of the true mean speed of drivers in the 25 mph construction zone is greater than 25 mph, failing to reject H_0.

11. Reading level

$$\mu = \text{mean reading level of pages in the novel}$$

$$H_0 : \mu = 5 \qquad\qquad H_a : \mu < 5$$

It is being tested whether the true mean is less than the null mean, so a 1-sided 1-sample t test is appropriate.
The randomness condition is met, as the sample is random.
The 10% condition is met, as it can be assumed that the novel is at least 400 pages long, so the sample size of 40 is less than 10% of the population size.

The Normality condition is met, as the sample size of 40 is greater than 30 and therefore large enough to fulfill the Central Limit Theorem to ensure Normality.

$$
\begin{aligned}
s_{\bar{x}} &= \frac{s_x}{\sqrt{n}} = \frac{0.8}{\sqrt{40}} \approx 0.126 \\
t &= \frac{\bar{x} - \mu_0}{s_{\bar{x}}} \approx \frac{4.8 - 5}{0.126} \approx -1.581 \\
P\text{-value} &= \text{tcdf}\left(\text{lower} : -\infty, \text{upper} : t \approx -1.581, \text{df} : n - 1 = 40 - 1 = 39\right) \approx 0.061
\end{aligned}
$$

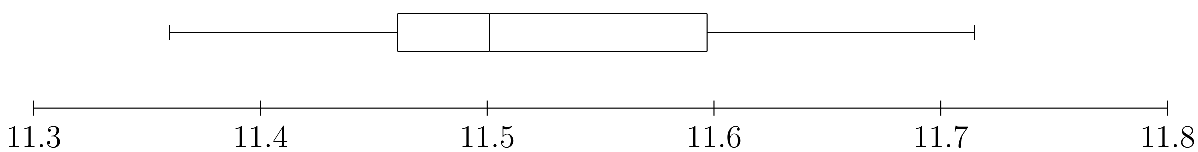
As the P-value of about 0.061 is greater than the significance level $\alpha = 0.05$, H_0 cannot be rejected. The data does not provide convincing evidence that the true mean reading level of the novel's pages is less than 5.

13. Pressing pills

$$\mu = \text{mean tablet hardness}$$

$$H_0 : \mu = 11.5 \qquad\qquad H_a : \mu \neq 11.5$$

As it is being determined whether or not the true mean varies from the null value, a 2-sided 1-sample t test should be carried out.
The Randomness condition is met, as the sample is random.
The 10% condition is met, as over 200 pills are likely produced per large batch, so the sample size of 20 is likely less than 10% of the population size.



The modified box plot of the distribution is roughly symmetrical and lacks outliers, so approximate Normality is justified.

$$
\begin{aligned}
\bar{x} &= \frac{\sum x_i}{n} \approx 11.513 \\
s_x &= \sqrt{\frac{\sum(x_i - \bar{x})}{n - 1}} \approx 0.094 \\
s_{\bar{x}} &= \frac{s_x}{\sqrt{n}} \approx \frac{0.094}{\sqrt{20}} \approx 0.021 \\
t &= \frac{\bar{x} - \mu_0}{s_{\bar{x}}} \approx \frac{11.513 - 11.5}{0.021} \approx 0.608 \\
P\text{-value} &= 1 - \text{tcdf}\left(\text{lower} : -t \approx -0.608, \text{upper} : t \approx 0.608, \text{df} : n - 1 = 20 - 1 = 19\right) \approx 0.45
\end{aligned}
$$

As the P-value of 0.45 is greater than the significance level $\alpha = 0.05$, H_0 cannot be rejected. The data does not provide convincing evidence of H_a.

15. Pressing pills

a. A 95% confidence interval provides a range of plausible values rather than simply stating whether or not it is plausible for the true mean to be a particular value as well. (Whether or not it is plausible for μ to be equal to be 11.55 can be determined by whether or not it is contained within the interval, which it in this case is.)

b. If the true mean is equal to 11.55, there is a 61% chance of convincing evidence being found for the alternative hypothesis.

c.

$$P(\text{Type II error}) = 1 - \text{power} = 1 - 0.61 = 0.39$$

d. The probability of a Type II error occurring can be decreased by either increasing the sample size, causing the standard errors of both the sample and the statistic to decrease, or increasing the effect size, in this case making μ larger than 11.55, making the observed differences in results larger and therefore easier to detect.

19. Tests and confidence intervals

a. As the two-sided test about μ has a P-value of 0.06, which is greater than 0.05, H_0 cannot be rejected at a significance level of $\alpha = 0.05$, so a confidence interval with a confidence level of $1 - \alpha = 0.95$ includes the null value of 10.

b. As the two-sided test about μ has a P-value of 0.06, which is less than 0.1, H_0 can be rejected at a significance level of $\alpha = 0.1$, so a confidence interval with a confidence level of $1 - \alpha = 0.9$ does not include the null value of 10.

21. Do you have ESP?

a. The P-value of their results is at most 0.01, meaning that the probability of them occurring due to random chance assuming the null hypothesis — that they are guessing randomly — is 0.01. As such, the expected number of people to receive results of that probability in a sample of size n is the product of the P-value and the sample size, in this case 1% of 500, which is 5. Therefore, 4 people having these results is not indicative of them possessing ESP.

b. To test whether or not these 4 subjects truly have ESP, the experiment should be repeated, as the chances of a person getting results with a P-value of at most 0.01 twice due only to random chance is the equal to the product of the sample size and the square of the P-value, making it .05 in this case, which is less than 1, meaning that none are expected to pass.

23. Improving SAT scores
The P-value of 0.0148 is less than the significance level $\alpha = 0.05$, making it statistically significant. The difference between the sample mean of those that used the app and the population mean of those within the program was only two points, though, which is an insignificant amount.

25. Sampling shoppers
The shoppers were consecutive, making the sample a convenience sample, which makes it subject to bias, shifting the sample mean in an unknown way that will not be accounted for in the test.

26. Ages of presidents
The population data is all that is available, so the parameter is known, making a significance test useless.

27.
The standard deviation of a mean is dependent on that of a quantitative value, which cannot be known without the population data. The answer is therefore **b**.

28.

$$P\text{-value} = \text{tcdf}\,(\text{lower} : -\infty, \text{upper} : -2.25, \text{df} : n-1 = 20-1 = 19) \approx 0.018$$

The answer is therefore **a**.

29.

$$|t| > \left|\text{invT}\left(\text{area} : \frac{1-0.005}{2} = 0.0025, \text{df} : n-1 = 15-1 = 14\right)]\right| \approx 3.326$$

The answer is therefore **d**.

30.
The test is not statistically significant at a significance level $\alpha = 0.01$, so the answer is **c**.

31.
Accuracy is vital for inference, so the answer is **a**.

32.
Increasing the sample size and significance level makes it more likely for a test to be statistically significant, so the answer is **a**.

11.2 Tests About a Difference in Means

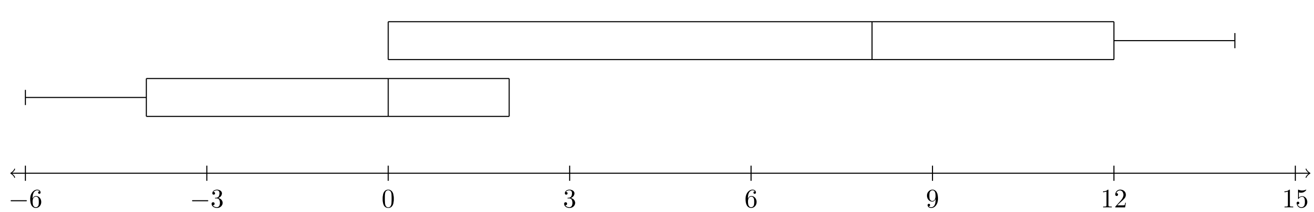
41. Fish oil

a.

μ_F = mean decrease in diastolic blood pressure in males with high blood pressure after 4 weeks with fish oil
μ_R = mean decrease in diastolic blood pressure in males with high blood pressure after 4 weeks with regular oil

$$H_0 : \mu_F = \mu_R \qquad H_a : \mu_F > \mu_R$$

As there are two groups that were randomly assigned treatments, a 2-sample t test should be performed.
The assignment of treatments was random, so the randomness condition is met.



The modified box plots of each distribution lack skew and outliers, so approximate Normality can be assumed.

$$\begin{aligned}
\bar{x}_F &= \frac{\sum x_{F,i}}{n_F} \approx 6.571 \\
\bar{x}_R &= \frac{\sum x_{R,i}}{n_R} \approx -1.143 \\
\bar{x}_F - \bar{x}_R &\approx 6.571 + 1.143 \approx 7.714 \\
s_F &= \sqrt{\frac{\sum(\bar{x}_F - x_{F,i})}{n_F - 1}} \approx 5.855 \\
s_R &= \sqrt{\frac{\sum(\bar{x}_R - x_{R,i})}{n_R - 1}} \approx 3.185 \\
s_{\bar{x}_F - \bar{x}_R} &= \sqrt{\frac{s_F^2}{n_f} + \frac{s_R^2}{n_R}} \approx \sqrt{\frac{5.855^2}{7} + \frac{3.185^2}{7}} \approx 2.519 \\
t &= \frac{\bar{x}_F - \bar{x}_R}{s_{\bar{x}_F - \bar{x}_R}} \approx \frac{7.714}{2.519} \approx 3.062 \\
\text{df} &= \frac{s^4_{\bar{x}_F - \bar{x}_R}}{\frac{s_F^4/n_F^2}{n_F - 1} + \frac{s_R^4/n_R^2}{n_R - 1}} \approx \frac{2.519^4}{\frac{5.855^4/7^2}{7-1} + \frac{3.185^4/7^2}{7-1}} \approx 9.264 \\
P\text{-value} &= \text{tcdf}\left(\text{lower} : t \approx 3.062, \text{upper} : \infty, \text{df} :\approx 9.274\right) \approx 0.007
\end{aligned}$$

As the P-value of 0.007 is less than the significance level of $\alpha = 0.05$, the null hypothesis can be rejected. The data does provide convincing evidence the fish oil helps reduce blood pressure more, on average, than regular oil for men like those in the study.

b. Assuming that fish oil has no affect on blood pressure compared to regular oil, the probability of getting a difference between the results of two studies at least as large as that observed here is about 0.7%.

43. Who talks more — men or women?

$$\begin{aligned}
\mu_F &= \text{mean number of words spoken by female students at the university per day} \\
\mu_M &= \text{mean number of words spoken by male students at the university per day}
\end{aligned}$$

$$H_0 : \mu_F = \mu_M \qquad\qquad H_a : \mu_F \neq \mu_M$$

As the data is not paired, and there are two means, a 2-sample t test should be performed.
The randomness condition is met, as the samples of the people are random.
The 10% condition is met, as the university is large, so there are likely more than 560 male students and 560 female students.
The Normality condition is met by the Central Limit theorem, as the sample sizes of 56 are greater than 30, making them large enough for the theorem to justify Normality.

$$\begin{aligned}
\bar{x}_F - \bar{x}_M &= 16177 - 16569 = -392 \\
s_{\bar{x}_F - \bar{x}_M} &= \sqrt{\frac{s_F^2}{n_F} + \frac{s_M^2}{n_M}} = \sqrt{\frac{7520^2}{56} + \frac{9108^2}{56}} \approx 1578.347 \\
t &= \frac{x_F - x_M}{s_{\bar{x}_F - \bar{x}_M}} \approx \frac{-392}{1578.347} \approx -0.248 \\
\text{df} &= \frac{s^4_{\bar{x}_F - \bar{x}_M}}{\frac{s_F^4/n_F^2}{n_F - 1} + \frac{s_M^4/n_M^2}{n_M - 1}} \approx \frac{1578.347^4}{\frac{7520^4/56^2}{56-1} + \frac{9108^4/56^2}{56-1}} \approx 106.195 \\
P\text{-value} &= 2\,\text{tcdf}\left(\text{lower} : -\infty, \text{upper} : t \approx -0.248, \text{df} :\approx 106.195\right) \approx 0.804
\end{aligned}$$

As the P-value of 0.804 is greater than the significance level $\alpha = 0.05$, H_0 cannot be rejected. The data does not provide convincing evidence that there is a difference between the mean number of words spoken per day by a given male and female student of this school.

45. Who talks more — men or women?

a.

$$t^* = \left|\text{invT}\left(\text{area} : \frac{1-C}{2} = \frac{1-0.95}{2} = 0.025, \text{df} :\approx 106.195\right)\right| \approx 1.983$$
$$\text{confidence interval} = \bar{x}_F - \bar{x}_M \pm t^* s_{\bar{x}_F - \bar{x}_M} \approx -392 \pm (1.983)(1578.347) \approx (-3521.862, 2737.862)$$

It can be said with 95% confidence that the true mean difference between the number of words spoken on a given day by a given male and female student of this university is contained within the interval $(-3521.862, 2737.862)$.

b. In addition to showing that it is plausible for no difference to exist, as 0 is contained within the interval, it shows a range of plausible values for the mean.

47. Teaching reading

a. Those that did the activities had a higher average score than those that did not, having a higher mean and median. Their scores were also less variable, as the distribution has a lower range and interquartile range.

b. The P-value of the test was 0.013, which is less than the significance level $\alpha = 0.05$, meaning that the null hypothesis can be rejected. The evidence convincingly supports the hypothesis that third-graders in the program that do the activities have higher a higher mean DRP score.

c. It can be concluded that the new reading activities resulted in an increase in the mean DRP score, as the data convincingly supports the alternative hypothesis and since this is an experiment, it can be used to justify causality.

d. Since the null hypothesis was rejected, a Type I error may have occurred.

49. A better drug?

a.

$$s_{\bar{x}_{\text{new}} - \bar{x}_{\text{cur}}} = \sqrt{\frac{s_{\text{new}}^2}{n_{\text{new}}} + \frac{s_{\text{cur}}^2}{n_{\text{cur}}}} = \sqrt{\frac{13.3^2}{15} + \frac{11.93^2}{14}} \approx 4.686$$
$$t = \frac{\bar{x}_{\text{new}} - \bar{x}_{\text{cur}} - \mu_{(\text{new} - \bar{x}_{\text{cur}}),0}}{s_{\bar{x}_{\text{new}} - \bar{x}_{\text{cur}}}} \approx \frac{68.7 - 54.1 - 10}{4.686} \approx 0.982$$

b.

$$\text{df} = \frac{s_{\bar{x}_{\text{new}} - \bar{x}_{\text{new}}}^4}{\frac{s_{\text{new}}^4 / n_{\text{new}}^2}{n_{\text{new}} - 1} + \frac{s_{\text{cur}}^4 / n_{\text{cur}}^2}{n_{\text{cur}} - 1}} \approx \frac{4.686^4}{\frac{13.3^4/15^2}{15-1} + \frac{11.93^4/14^2}{14-1}} \approx 26.963$$
$$P\text{-value} = \text{tcdf}\left(\text{lower} : t \approx 0.982, \text{upper} : \infty, \text{df} :\approx 26.963\right) \approx 0.168$$

As the P-value of about 0.168 is greater than the significance level $\alpha = 0.05$, H_0 cannot be rejected. The data does not provide convincing evidence that there exists a difference between the mean cholesterol reduction caused by the current and new drugs.

c. A test's power can be increased by increasing the sample size, reducing the standard error of the difference and increasing the standardized test statistic and degrees of freedom, resulting in a lower P-value, as the tails are reduced. It can also be increased by increasing the significance level, making it easier for the P-value to fall beneath it.

51. Rewards and creativity

a. Random assignment is the only way to justify causality in an experiment.

b. A Type I error would occur if it was concluded that the data provides convincing evidence that intrinsic rewards promote creativity better than extrinsic ones, the null hypothesis being rejected, despite that not being the case in reality.
A Type II error would occur if it was concluded that the data does not provide convincing evidence that intrinsic rewards promote creativity better than intrinsic ones, the null hypothesis not being rejected, despite it being the case in reality.

c.

$$P\text{-value} = P(\bar{x}_{\text{intrinsic}} - \bar{x}_{\text{extrinsic}} > 19.833 - 15.739) \approx \frac{1}{47} \approx 0.021$$

Assuming that there is no difference between the true values of the means, the probability of data producing a difference in sample means of at least 4.144, that observed, is about 2.1%.

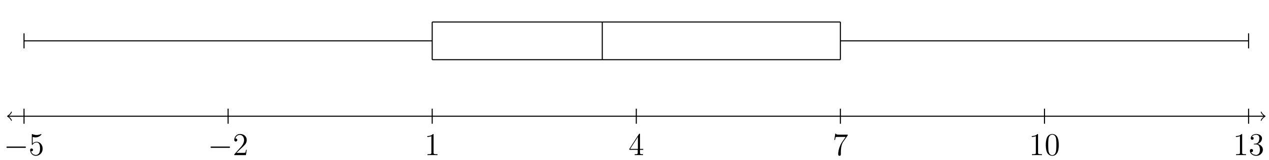
d. Because the P-value of about 0.021 is less than the significance level of $\alpha = 0.05$, the null hypothesis can be rejected. The data provides convincing evidence that there is a nonzero difference between the true mean creativity (based on a scored poem) depending on whether the motivation is intrinsic or extrinsic.

53. Drive-thru or go inside?

$$\mu_I = \text{time taken to place order inside (s)}$$
$$\mu_D = \text{time take to place order through drive-thru (s)}$$
$$\mu_{\text{diff}} = \mu_{I-D} = \mu_I - \mu_D$$

$$H_0 : \mu_I = \mu_D \qquad\qquad H_a : \mu_I > \mu_D$$

The data is paired, as each ordered at the same time, so a paired t test should be performed.
The randomness condition is met, as the times were randomly generated.
The independence condition is met via the 10% rule, as there were likely over 100 orders of each type placed over the two-week period at the establishment, making the sample size of 10 less than 10% of each population.



The distribution of the differences is symmetrical and lacks outliers, so Normality is justified.

$$\begin{aligned}
\bar{x}_{\text{diff}} &= \frac{\sum x_{\text{diff},i}}{n} = 3.9 \\
s_{\text{diff}} &= \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} \approx 5.646 \\
s_{\bar{x}_{\text{diff}}} &= \frac{s_{\text{diff}}}{\sqrt{n}} \approx \frac{5.646}{\sqrt{10}} \approx 1.786 \\
t &= \frac{\bar{x}_{\text{diff}} - \mu_0}{s_{\bar{x}_{\text{diff}}}} \approx \frac{3.9 - 0}{1.786} \approx 2.184 \\
\text{df} &= n - 1 = 10 - 1 = 9 \\
P\text{-value} &= \text{tcdf}\left(\text{lower} : 2.184, \text{upper} : \infty, \text{df} : 9\right) \approx 0.028
\end{aligned}$$

As the P-value is less than 0.05, the null hypothesis can be rejected. The data provides convincing evidence that the mean time taken inside is greater than that taken going through the drive-thru.

59. Two samples or paired data?

a. The cars are of the same make and model and the assignment is random, so the data is paired.

b. The subjects are exposed to both treatments, so the data is paired.

c. The subjects are not the same, so the data is not paired.

61. Have a ball!

1. Each student threw both balls, so the data is paired.

2. The median is greater that 0, indicating that the center of the distribution of the differences is 0, so a difference may exist.

3.

$$\begin{aligned}
\mu_B &= \text{mean baseball throw distance (yd)} \\
\mu_S &= \text{mean softball throw distance (yd)} \\
\mu_{\text{diff}} &= \mu_{B-S} = \mu_B - \mu_S
\end{aligned}$$

$$H_0 : \mu_{\text{diff}} = 0 \qquad\qquad H_a : \mu_{\text{diff}} > 0$$

4. The distribution contains an outlier, meaning that Normality cannot be justified, the sample size of 24 being less than 30 and therefore too small for central limit theorem to justify Normality.

5.

$$\begin{aligned}
\bar{x}_{\text{diff}} &= \frac{\sum x_i}{n} \approx 6.542 \\
P\text{-value} &= P(\bar{x}_{\text{diff}} > \bar{x}_{\text{diff}}) \approx P(\bar{x}_{\text{diff}} > 6.542) \approx 0
\end{aligned}$$

As the P-value of 0 is less than the significance level $\alpha = 0.05$, the null hypothesis can be rejected. The data provides convincing evidence that the mean distance that one can throw a baseball is greater than that of a softball.

63. Coaching and SAT scores

a. The change in an individual is being examined, so the data is paired. A paired t test using gain should therefore be used.

b.

$$\mu_G = \text{mean gain in SAT score from Try 1 to 2}$$

$$H_0 : \mu_G = 0 \qquad\qquad H_a : \mu_G > 0$$

A single difference in means is being tests, so a 2 sample t test should be performed.
The randomness condition is met, as the sample was randomly taken.
The independence condition is met, as there are likely over 4270 coached students, making the sample size of 427 less than 10% of the population size of 4270.
The Normality condition is met, as the sample size of 427 is greater than 30, so the central limit theorem can be applied to justify Normality.

$$s_{\bar{x}_G} = \frac{s_G}{\sqrt{n_C}} = \frac{59}{\sqrt{427}} \approx 2.855$$

$$t = \frac{\bar{x}_G}{s_{\bar{x}_G}} \approx \frac{29}{2.855} \approx 10.157$$

$$\text{df} = n_C - 1 = 427 - 1 = 426$$

$$P\text{-value} = \text{tcdf}\left(\text{lower} : t \approx 10.157, \text{upper} : \infty, \text{df} : 426\right) \approx 3.774 \times 10^{-22}$$

As the P-value of about 3.774×10^{-22} is less than the significance level $\alpha = 0.05$, the null hypothesis can be rejected. The data provides convincing evidence that the coaching does increase SAT scores on average.

64. Coaching and SAT scores

a.

μ_C = mean score gain for those that were coachedμ_U = mean score gain for those that were not coached

$$H_0 : \mu_C = \mu_U \qquad\qquad H_a : \mu_C > \mu_U$$

There are two unpaired means, so a 2 sample t test should be performed. The samples were randomly taken, so the Randomness condition is met.
The independence condition is met, as there are well over 4,270 and 27330 coached and uncoached students respectively, making the sample sized of 247 and 2733 less than 10% of their respective population sizes.
The Normality condition is met, as sample sizes of 247 and 2733 are both greater than 30, so the central limit theorem is applicable to justify Normality.

$$\bar{x}_C - \bar{x}_U = 29 - 21 = 8$$

$$s_{\bar{x}_C - \bar{x}_U} = \sqrt{\frac{s_C^2}{n_C} + \frac{s_U^2}{n_U}} = \sqrt{\frac{59^2}{427} + \frac{52^2}{2733}} \approx 3.024$$

$$t = \frac{\bar{x}_C - \bar{x}_U}{s_{\bar{x}_C - \bar{x}_U}} \approx \frac{8}{3.024} \approx 2.646$$

$$\text{df} = \frac{s^4_{\bar{x}_C - \bar{x}_U}}{\frac{s_C^4/n_C^2}{n_C - 1} + \frac{s_U^4/n_U^2}{n_U - 1}} \approx \frac{3.024^4}{\frac{59^4/427^2}{427-1} + \frac{52^4/2733^2}{2733-1}} \approx 534.45$$

$$P\text{-value} = \text{tcdf}\left(\text{lower} : t \approx 2.646, \text{upper} : \infty, \text{df} :\approx 534.45\right) \approx 0.004$$

As the P-value of about 0.004 is less than the significance level $\alpha = 0.05$, the null hypothesis can be rejected. The data provides convincing evidence that the mean gain in SAT score for those that were coached is greater than that of those that were not.

b. The interval exclusively contains values that are greater than 0, telling us that the true mean gain is not 0. It also provides a range of plausible values for the true mean, though.

c. I believe that coaching courses are worth paying for, as the increase is significant both statistically and practically.

65.
A Type II error occurs when the null hypothesis is not rejected despite the alternate hypothesis being true. The answer is therefore **d**.

66.
The P-value is the probability of receiving results at least as significant as those supported assuming that the null hypothesis is true. The answer is therefore **b**.

67.
A two-sided t test with significance level α produces the same results as a confidence interval with confidence level $C = 1 - 2\alpha$. The answer is therefore **d**.

68.
For a significance test to be performed about a difference in paired means, the distribution of the differences must be checked for Normality. The answer is therefore **c**.

69.
To test a difference between paired measured concentrations, a paired t test should be performed. The answer is therefore **a**.

Chapter 12

Inference for Distributions and Relationships

12.1 Chi-Square Tests for Goodness of Fit

1. Aw, nuts!

a.

$$\begin{aligned}
&H_0 : \text{The true distribution of mixed nuts is just as the company claims.}\\
&H_a : \text{The true distribution of mixed nuts differs from what the company claims.}
\end{aligned}$$

b.

$$\begin{aligned}
\text{Cashew} &= 0.52(150) = 78 = 78 & \text{Almond} &= 0.27(150) = 40.5\\
\text{Macadamia} &= 0.13(150) = 19.5 & \text{Brazil} &= 0.08(150) = 12
\end{aligned}$$

$$\chi^2 = \sum \left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}}\right] \approx 6.599$$

3. P-values

1.

$$\begin{aligned}
&0.05 < P\text{-value} < 0.1\\
&\qquad P\text{-value} = \chi^2\text{cdf}\,(\text{lower} : 19.03, \text{upper} : \infty, \text{df} : 11) \approx 0.061
\end{aligned}$$

2.

$$\begin{aligned}
&P\text{-value} < 0.0005\\
&P\text{-value} = \chi^2\text{cdf}\,(\text{lower} : 19.03, \text{upper} : \infty, \text{df} : 3) \approx 2.695 \times 10^{-4}
\end{aligned}$$

5. Aw, nuts!

a. There are over 5 counts expected in each category, so the Large Counts condition is met.

$$\text{df} = \text{categories} - 1 = 4 - 1 = 3$$

b.

$$0.05 < P\text{-value} < 0.1$$
$$P\text{-value} = \chi^2\text{cdf}\left(\text{lower} : \chi^2 \approx 6.599, \text{upper} : \infty, \text{df} : n-1 = 4-1 = 3\right) \approx 0.086$$

c. Over many trials, there is an 8.6% probability of receiving a difference at at least as great as that observed from the claimed proportions.

d. As the P-value of about 8.6% is greater than the significance level $\alpha = 0.05$, the null hypothesis cannot be rejected. The data does not provide convincing evidence that the true proportions of deluxe mixed nuts significantly differ from those claimed.

9. Munching Froot Loops

$$H_0 : \text{The proportions of Froot Loop colors are those claimed by Kellogg's.}$$
$$H_a : \text{The proportions of Froot Loop colors are not those claimed by Kellogg's.}$$

The sample was random, so the Randomness condition is met.
There are over 1200 Froot Loops, so the sample size of 120 is less than 10% of the population size.
As the probabilities are equal, the expected counts of each color is equal, being $120/6 = 20$.
As the expected counts of all categories is 20, which is greater than 5, the Large Counts condition is met.

$$\chi^2 = \sum\left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}}\right] = 7.9$$
$$\text{df} = \text{categories} - 1 = 6 - 1 = 5$$
$$P\text{-value} = \chi^2\text{cdf}\,(\text{lower} : 7.9, \text{upper} : \infty, \text{df} : 5) \approx 0.162$$

As the P-value of about 0.162 is greater than the significance level of $\alpha = 0.05$, the null hypothesis cannot be rejected. The data does not provide convincing evidence that the true proportions of Froot Loops colors differs from those claimed by Kellogg's.

13. Birds in the tress

a.

$$H_0 : \text{The red-breasted nuthatches have no preference regarding trees.}$$
$$H_a : \text{The red-breasted nuthatches have some preference regarding trees.}$$

The sample was random, so the Randomness condition is met.
The sample size of 156 is likely less than 10% of the number of red-breasted nuthatches in the forest, so the 10% condition is met, justifying independence.

$$\text{Douglas firs} = (0.54)(312) = 84.24 \qquad \text{ponderosa pines} = (0.4)(156) = 62.4$$
$$\text{other} = (0.06)(156) = 9.36$$

The expected counts of each category are greater than 5, so the Large Counts condition is met.

$$\chi^2 = \sum\left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}}\right] \approx 7.418$$
$$\text{df} = \text{categories} - 1 = 3 - 1 = 2$$
$$P\text{-value} = \chi^2\text{cdf}\left(\text{lower} : \chi^2 \approx 7.418, \text{upper} : \infty, \text{df} : 2\right) \approx 0.024$$

As the P-value of about 0.024 is less than the significance level $\alpha = 0.05$, the null hypothesis can be rejected. The data provides convincing evidence that red-breasted nuthatches have some preference regarding trees.

b.

tree	observed	expected	observed − expected
Douglas firs	70	84.24	−14.24
ponderosa pines	79	62.4	16.6
other	7	9.36	−2.36

The greatest (and only) positive difference between the observed and expected counts indicates that red-breasted nuthatches prefer ponderosa pines. The greatest negative difference indicates that they are most averse to Douglas firs.

15. Mendel and the peas

a.

H_0 : The actual proportions of smooth and wrinkled peas are 0.75 and 0.25 respectively.
H_a : The actual proportions of smooth and wrinkled peas are not 0.75 and 0.25 respectively

	observed	expected
smooth	423	$0.75(423+133) = 417$
wrinkled	133	$0.25(423+133) = 139$

$$\chi^2 = \sum \left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}}\right] \approx 0.345$$
$$\text{df} = \text{categories} - 1 = 2 - 1 = 1$$
$$P\text{-value} = \text{tcdf}\left(\text{lower} : \chi^2 \approx 0.345, \text{upper} : \infty, \text{df} : 1\right) \approx 0.557$$

As the P-value of about 0.557 is greater than the significance level $\alpha = 0.05$, the null hypothesis cannot be rejected. The data does not provide convincing evidence that the true proportions of smooth and wrinkled peas are not 0.75 and 0.25 respectively.

b.

$$\hat{p} = \frac{423}{423+133} \approx 0.761$$
$$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \approx \sqrt{\frac{0.761(1-0.761)}{423+133}} \approx 0.018$$
$$z = \frac{\hat{p} - p_0}{s_{\hat{p}}} \approx \frac{0.761 - 0.75}{0.018} \approx 0.596$$
$$P\text{-value} = 1 - \text{normalcdf}\left(\text{lower} : -z \approx -0.596, \text{upper} : z \approx 0.596, \mu : 0, \sigma : 1\right) \approx 0.551$$

As the P-value of about 0.551 is greater than the significance level $\alpha = 0.05$, the null hypothesis cannot be rejected. The data does not provide convincing evidence that the true proportion smooth peas is not equal to 0.75.

17. Is your random number generator working?

a.

$$H_0 : \text{Each digit's proportion is } 0.1. \qquad H_a : \text{Not every digit's proportion is } 0.1$$

The randomness condition is met, as the numbers were randomly generated.
The independence condition is met, as each digit being generated was independent of the last.

digit	0	1	2	3	3	5	6	7	8	9
observed	18	20	25	25	21	15	18	19	13	26
expected	$0.1(200) = 20$	20	20	20	20	20	20	20	20	20

The Large Counts condition is met, as each digit is expected to be generated 20 times, which is greater than 5.

b.

$$\chi^2 = \sum \left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}}\right] = 6.7$$

$$\text{df} = \text{categories} - 1 = 10 - 1 = 9$$

$$P\text{-value} = \chi^2\text{cdf}\left(\text{lower} : \chi^2 = 6.7, \text{upper} : \infty, \text{df} : 9\right) \approx 0.668$$

As the P-value of 0.668 is greater than the significance level $\alpha = 0.05$, the null hypothesis cannot be rejected. The data does not provide convincing evidence of the true proportions of randomly generated digits being nonuniform.

c. The probability of a Type I error occurring is equal to the significance level $\alpha = 0.05$.

d. The probability of a Type I error occurring is the same for every trial.
The number of trials is fixed at 25.
A Type I error can either occur or not occur, making the outcome binary.
The number of Type I errors is therefore a binomial variable.

$$P(X \geq 1) = 1 - P(X = 0) = 1 - \binom{25}{0}(0.05)^0(1 - 0.05)^{25-0} \approx 0.277$$

19.
The null hypothesis should describe population proportions, making the answer **d**.

20.

$$\chi^2 = \sum \left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}}\right]$$

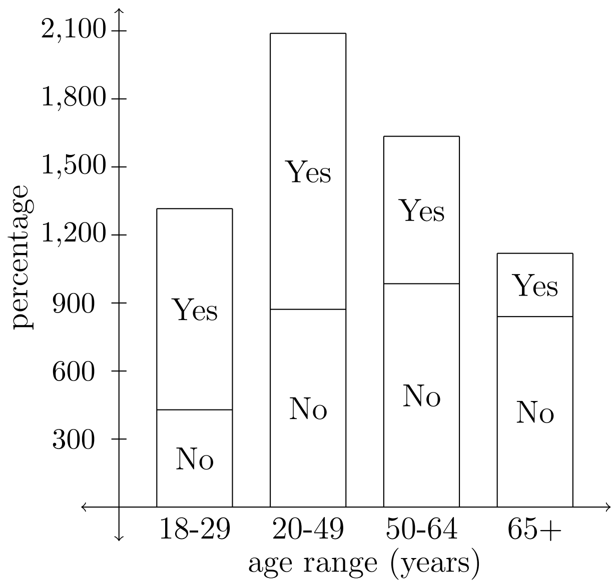
The answer is therefore **a**.

21.
A P-value less than the significance level results in H_0 being rejected due to the data providing convincing evidence for H_a. The answer is therefore **c**.

22.
The degrees of freedom for a chi-square test is determined by the number of categories, not the sample size. The answer is therefore **c**.

23. Video games

a.



b. The lower the age, the higher the proportion of people that play video games.

12.2 Inference for Two-Way Tables

27. The color of candy
Those that took a colored survey showed a strong preference for the candy of the same color while those in the control group slightly preferred the blue candy.

29. More candy

a.

$$H_0 : \text{The distributions of candy colors remains the same regardless of the survey.}$$
$$H_a : \text{The distributions of candy colors is affected by the survey.}$$

b.

	Red	Blue
Red	$\frac{20(26)}{60} = 8.\overline{3}$	$\frac{20(26)}{60} = 8.\overline{3}$
Blue	$\frac{20(34)}{60} = 11.\overline{3}$	$\frac{20(34)}{60} = 11.\overline{3}$

c.

$$\chi^2 = \sum \left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}} \right] \approx 6.652$$

31. Last candy

a. The Randomness condition is met, as the assignment was random.
The independence condition is met, as this is an experiment.
The Large Counts condition is met, as all expected values are greater than 5.

b.

$$0.005 < P\text{-value} < 0.01$$
$$P\text{-value} = \text{tcdf}\left(\text{lower} : \chi^2 \approx 6.652, \text{upper} : \infty, \text{df} : (2-1)(3-1) = 2\right) \approx 0.036$$

c. If the survey has no affect on preferences, the probability of receiving data that differ from the expected values at least as much as those observed is about 3.6%.

d. As the P-value of about 0.036 is greater than the significance level $\alpha = 0.01$, the null hypothesis cannot be rejected. The data does not provide convincing evidence that the survey affects the preferences for candy color.

33. Sorry, no chi-square
The counts of the travelers from each category is not known, just the relative frequencies of each category.

34. Going nuts
The data describes means rather than counts (proportions).

35. Gummy bears

$$H_0 : \text{The distributions of colors is the same for name and store-brand gummy bears}$$
$$H_a : \text{The distributions of colors differs for name and store-brand gummy bears}$$

The randomness condition is met, as the samples were random.
The independence condition is met, as there are over 6 bags is far less than 10% of the number of bags of name or store-brand gummy bears.

	Name	Store
Red	130.83	218.17
Green	58.86	98.15
Yellow	50.61	84.39
Orange	77.97	130.03
White	54.742	91.268

Every expected count exceeds 5, so the Large Counts condition is met.

$$\chi^2 = \sum \left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}}\right] \approx 1.815$$
$$\text{df} = (2-1)(5-1) = 4$$
$$P\text{-value} = \chi^2\text{cdf}\left(\text{lower} : \chi^2 \approx 1.815, \text{upper} : \infty, \text{df} : 4\right) \approx 0.77$$

As the P-value of about 0.77 is greater than the significance level $\alpha = 0.05$, the null hypothesis cannot be rejected. The data does not provide convincing evidence of a difference existing between the true distributions of gummy bear colors for name and store brands.

37. How to quit smoking

a.

	Nicotine patch	Drug	Patch plus drug	Placebo	Total
Successes	40	74	87	25	226
Failures	204	170	158	135	667
Total	244	244	245	160	893

b.

$$H_0 : \text{The proportions of successes and failures is the same across all treatments.}$$
$$H_a : \text{The proportions of successes and failures is not the same across all treatments.}$$

The randomness condition is met, as the assignment was random.
The independence condition is met, as this is an experiment.

	Nicotine patch	Drug	Patch plus drug	Placebo
Successes	61.75	61.75	62	40.49
Failures	182.25	182.25	183	119.51

Every expected count is greater than 5, so the Large Counts condition is met.

$$\chi^2 = \sum \left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}}\right] \approx 34.937$$
$$\text{df} = (4-1)(2-1) = 3$$
$$P\text{-value} = \chi^2\text{cdf}\left(\text{lower} : \chi^2 \approx 34.937, \text{upper} : \infty, \text{df} : 3\right) \approx 1.256 \times 10^{-7}$$

As the P-value of about 1.256×10^{-7} is less than the significance level $\alpha = 0.05$, the null hypothesis can be rejected. The data provides convincing evidence that different treatments result in a different distribution of successes and failures.

39. How to quit smoking

The treatment with the largest positive difference between its actual and expected values is the combination of the patch and drug, suggesting that is is the most effective. The one with the largest negative difference, on the other hand, is the nicotine patch, indicating that it is the least effective treatment.

41. Relaxing in the sauna

a.

$$H_0 : \text{The distribution of suffering from SCD is independent of weekly sauna frequency.}$$
$$H_a : \text{Weekly sauna frequency affects the distribution of suffering from SCD.}$$

b.

	¡1	2-3	¿4
Yes	49.33	124.18	17.5
No	551.67	1388.8	184.5

c.

$$\chi^2 = \sum \left[\frac{(\text{observed} - \text{expected})^2}{\text{expected}}\right] \approx 6.032$$
$$\text{df} = (3-1)(2-1) = 2$$
$$P\text{-value} = \chi^2\text{cdf}\left(\text{lower} : \chi^2 \approx 6.032, \text{upper} : \infty, \text{df} : 2\right) \approx 0.049$$

d. As the P-value of about 0.049 is less than the significance level $\alpha = 0.05$, the null hypothesis can be rejected. The data provides convincing evidence that going to the sauna affects the distribution of SCD suffering.

43. Finger length

$$H_0 :$$

47.

53.

65.

66.

67.

68.

12.3 Inference for Slopes

69.

71.

73.

25.

77.

79.

83.

85.

87.

88.

89.

91.