

Chapter 23 – Inferences About Means**1. *t*-models, part I.**

- a) 1.74 b) 2.37 c) 0.0524 d) 0.0889

2. *t*-models, part II.

- a) 2.37 b) 2.63 c) 0.9829 d) 0.0381

3. *t*-models, part III.

As the number of degrees of freedom increases, the shape and center of *t*-models do not change. The spread of *t*-models decreases as the number of degrees of freedom increases, and the shape of the distribution becomes closer to Normal.

4. *t*-models, part IV (last one!).

As the number of degrees of freedom increases, the critical value of *t* for a 95% confidence interval gets smaller, approaching approximately 1.960, the critical value of *z* for a 95% confidence interval.

5. Cattle.

- a) Not correct. A confidence interval is for the mean weight gain of the population of all cows. It says nothing about individual cows. This interpretation also appears to imply that there is something special about the interval that was generated, when this interval is actually one of many that could have been generated, depending on the cows that were chosen for the sample.
- b) Not correct. A confidence interval is for the mean weight gain of the population of all cows, not individual cows.
- c) Not correct. We don't need a confidence interval about the average weight gain for cows in this study. We are certain that the mean weight gain of the cows in this study is 56 pounds. Confidence intervals are for the mean weight gain of the population of all cows.
- d) Not correct. This statement implies that the average weight gain varies. It doesn't. We just don't know what it is, and we are trying to find it. The average weight gain is either between 45 and 67 pounds, or it isn't.
- e) Not correct. This statement implies that there is something special about our interval, when this interval is actually one of many that could have been generated, depending on the cows that were chosen for the sample. The correct interpretation is that 95% of samples of this size will produce an interval that will contain the mean weight gain of the population of all cows.

6. Teachers.

- a) Not correct. Actually, 9 out of 10 samples will produce intervals that will contain the mean salary for Nevada teachers. Different samples are expected to produce different intervals.
- b) Correct! This is the one!
- c) Not correct. A confidence interval is about the mean salary of the population of Nevada teachers, not the salaries of individual teachers.

- d) Not correct. A confidence interval is about the mean salary of the population of Nevada teachers and doesn't tell us about the sample, nor does it tell us about individual salaries.
- e) Not correct. The population is teachers' salaries in Nevada, not the entire United States.

7. Meal plan.

- a) Not correct. The confidence interval is not about the individual students in the population.
- b) Not correct. The confidence interval is not about individual students in the sample. In fact, we know exactly what these students spent, so there is no need to estimate.
- c) Not correct. We know that the mean cost for students in this sample was \$1196.
- d) Not correct. A confidence interval is not about other sample means.
- e) This is the correct interpretation of a confidence interval. It estimates a population parameter.

8. Rain.

- a) Not correct. The confidence interval is not about the years in the sample.
- b) Not correct. The confidence interval does not predict what will happen in any one year.
- c) Not correct. The confidence interval is not based on a sample of days.
- d) This is the correct interpretation of a confidence interval. It estimates a population parameter.
- e) Not correct. We know exactly what the mean was in the sample. The mean snowfall was 23" per winter over the last century.

9. Pulse rates.

- a) We are 95% confident the interval 70.9 to 74.5 beats per minute contains the true mean heart rate.
- b) The width of the interval is about $74.5 - 70.9 = 3.6$ beats per minute. The margin of error is half of that, about 1.8 beats per minute.
- c) The margin of error would have been larger. More confidence requires a larger critical value of t , which increases the margin of error.

10. Crawling.

- a) We are 95% confident that the interval 29.2 to 31.8 weeks contains the true mean age at which babies begin to crawl.
- b) The width of the interval is about $31.8 - 29.2 = 2.6$ weeks. The margin of error is half of that, about 1.3 weeks.
- c) The margin of error would have been smaller. Less confidence requires a smaller critical value of t , which decreases the margin of error.

11. CEO compensation.

We should be hesitant to trust this confidence interval, since the conditions for inference are not met. The distribution is highly skewed and there is an outlier.

12. Credit card charges.

The analysts did not find the confidence interval useful because the conditions for inference were not met. There is one cardholder who spent over \$3,000,000 on his card. This made the standard deviation, and therefore the standard error, huge. The t -interval is too wide to be of any use.

13. Normal temperature.

a) **Randomization condition:** The adults were randomly selected.

10% condition: 52 adults are less than 10% of all adults.

Nearly Normal condition: The sample of 52 adults is large, and the histogram shows no serious skewness, outliers, or multiple modes.

The people in the sample had a mean temperature of 98.2846° and a standard deviation in temperature of 0.682379° . Since the conditions are satisfied, the sampling distribution of the mean can be modeled by a Student's t model, with $52 - 1 = 51$ degrees of freedom. We will use a one-sample t -interval with 98% confidence for the mean body temperature. (By hand, use $t_{50}^* \approx 2.403$ from the table.)

$$b) \bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 98.2846 \pm t_{51}^* \left(\frac{0.682379}{\sqrt{52}} \right) \approx (98.06, 98.51)$$

- c) We are 98% confident that the interval 98.06°F to 98.51°F contains the true mean body temperature for adults. (If you calculated the interval by hand, using $t_{50}^* \approx 2.403$ from the table, your interval may be slightly different than intervals calculated using technology. With the rounding used here, they are identical. Even if they aren't, it's not a big deal.)
- d) 98% of all random samples of size 52 will produce intervals that contain the true mean body temperature of adults.
- e) Since the interval is completely below the body temperature of 98.6°F , there is strong evidence that the true mean body temperature of adults is lower than 98.6°F .

14. Parking.

a) **Randomization condition:** The weekdays were not randomly selected. We will assume that the weekdays in our sample are representative of all weekdays.

10% condition: 44 weekdays are less than 10% of all weekdays.

Nearly Normal condition: We don't have the actual data, but since the sample of 44 weekdays is fairly large it is okay to proceed.

The weekdays in the sample had a mean revenue of \$126 and a standard deviation in revenue of \$15. The sampling distribution of the mean can be modeled by a Student's t model, with $44 - 1 = 43$ degrees of freedom. We will use a one-sample t -interval with 90% confidence for the mean daily income of the parking garage. (By hand, use $t_{40}^* \approx 1.684$)

- b) $\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 126 \pm t_{43}^* \left(\frac{15}{\sqrt{44}} \right) \approx (122.2, 129.8)$
- c) We are 90% confident that the interval \$122.20 to \$129.80 contains the true mean daily income of the parking garage. (If you calculated the interval by hand, using $t_{40}^* \approx 1.684$ from the table, your interval will be (122.19, 129.81), ever so slightly wider from the interval calculated using technology. This is not a big deal.)
- d) 90% of all random samples of size 44 will produce intervals that contain the true mean daily income of the parking garage.
- e) Since the interval is completely below the \$130 predicted by the consultant, there is evidence that the average daily parking revenue is lower than \$130.

15. Normal temperatures, part II.

- a) The 90% confidence interval would be narrower than the 98% confidence interval. We can be more precise with our interval when we are less confident.
- b) The 98% confidence interval has a greater chance of containing the true mean body temperature of adults than the 90% confidence interval, but the 98% confidence interval is less precise (wider) than the 90% confidence interval.
- c) The 98% confidence interval would be narrower if the sample size were increased from 52 people to 500 people. The smaller standard error would result in a smaller margin of error.
- d) Our sample of 52 people gave us a 98% confidence interval with a margin of error of $(98.51 - 98.05)/2 = 0.225^\circ\text{F}$. In order to get a margin of error of 0.1, less than half of that, we need a sample over 4 times as large. It should be safe to use $t_{100}^* \approx 2.364$ from the table, since the sample will need to be larger than 101. Or we could use $z^* \approx 2.326$, since we expect the sample to be large. We need a sample of about 252 people in order to estimate the mean body temperature of adults to within 0.1°F .

$$ME = t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right)$$

$$0.1 = 2.326 \left(\frac{0.682379}{\sqrt{n}} \right)$$

$$n = \frac{(2.326)^2 (0.682379)^2}{(0.1)^2}$$

$$n \approx 252$$

16. Parking II.

- a) The 95% confidence interval would be wider than the 90% confidence interval. We can be more confident that our interval contains the mean parking revenue when we are less precise. This would be better for the city because the 95% confidence interval is more likely to contain the true mean parking revenue.
- b) The 95% confidence interval is wider than the 90% confidence interval, and therefore less precise. It would be difficult for budget planners to use this wider interval, since they need precise figures for the budget.
- c) By collecting a larger sample of parking revenue on weekdays, they could create a more precise interval without sacrificing confidence.

- d) The confidence interval that was calculated in Exercise 10 won't help us to estimate the sample size. That interval was for 90% confidence. Now we want 95% confidence. A quick estimate with a critical value of $z^* = 2$ (from the 68-95-99.7 rule) gives us a sample size of 100, which will probably work fine. Let's be a bit more precise, just for fun! Conservatively, let's choose t^* with fewer degrees of freedom, which will give us a wider interval. From the table, the next available number of degrees of freedom is $t_{80}^* \approx 1.990$, not much different than the estimate of 2 that was used before. If we substitute 1.990 for t^* , we can estimate a sample size of about 99. Why not play it a bit safe? Use $n = 100$.

$$ME = t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right)$$

$$3 = 2 \left(\frac{15}{\sqrt{n}} \right)$$

$$n = \frac{(2)^2 (15)^2}{(3)^2}$$

$$n = 100$$

17. Speed of Light.

- a) $\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 756.22 \pm t_{22}^* \left(\frac{107.12}{\sqrt{23}} \right) \approx (709.9, 802.5)$
- b) We are 95% confident that the interval 299,709.9 to 299,802.5 km/sec contains the speed of light.
- c) We have assumed that the measurements are independent of each other and that the distribution of the population of all possible measurements is Normal. The assumption of independence seems reasonable, but it might be a good idea to look at a display of the measurements made by Michelson to verify that the Nearly Normal Condition is satisfied.

18. Better light.

- a) $SE(\bar{y}) = \left(\frac{s}{\sqrt{n}} \right) = \left(\frac{79.0}{\sqrt{100}} \right) = 7.9$ km/sec.
- b) The interval should be narrower. There are three reasons for this: the larger sample size results in a smaller standard error (reducing the margin of error), the larger sample size results in a greater number of degrees of freedom (decreasing the value of t^* , reducing the margin of error), and the smaller standard deviation in measurements results in a smaller standard error (reducing the margin of error). Additionally, the interval will have a different center, since the sample mean is different.
- c) We must assume that the measurements are independent of one another. Since the sample size is large, the Nearly Normal Condition is overridden, but it would still be nice to look at a graphical display of the measurements. A one-sample t -interval for the speed of light can be constructed, with $100 - 1 = 99$ degrees of freedom, at 95% confidence.

$$\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 852.4 \pm t_{99}^* \left(\frac{79.0}{\sqrt{100}} \right) \approx (836.72, 868.08)$$

We are 95% confident that the interval 299,836.72 to 299,868.08 km/sec contains the speed of light.

Since the interval for the new method does not contain the true speed of light as reported by Stigler, 299,710.5 km/sec., there is no evidence to support the accuracy of Michelson's new methods.

The interval for Michelson's old method (from Exercise 14) does contain the true speed of light as reported by Stigler. There is some evidence that Michelson's previous measurement technique was a good one, if not very precise.

19. Departures.

- a) **Randomization condition:** Since there is no time trend, the monthly on-time departure rates should be independent. This is not a random sample, but should be representative.
10% condition: These months represent fewer than 10% of all months.
Nearly Normal condition: The histogram looks unimodal, and slightly skewed to the left. Since the sample size is 144, this should not be of concern.
- b) The on-time departure rates in the sample had a mean of 81.1838%, and a standard deviation in of 4.47094%. Since the conditions have been satisfied, construct a one-sample t -interval, with $144 - 1 = 143$ degrees of freedom, at 90% confidence.

$$\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 81.18 \pm t_{143}^* \left(\frac{4.47}{\sqrt{144}} \right) \approx (80.57, 81.80)$$

- c) We are 90% confident that the interval from 80.57% to 81.80% contains the true mean monthly percentage of on-time flight departures.

20. Late arrivals.

- a) **Randomization condition:** Since there is no time trend, the monthly on-time arrival rates should be independent. This is not a random sample, but should be representative.
10% condition: These months represent fewer than 10% of all months.
Nearly Normal condition: The histogram looks unimodal and symmetric.
- b) The on-time arrival rates in the sample had a mean of 20.08%, and a standard deviation in of 4.09%. Since the conditions have been satisfied, construct a one-sample t -interval, with $144 - 1 = 143$ degrees of freedom, at 99% confidence.

$$\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 20.08 \pm t_{143}^* \left(\frac{4.09}{\sqrt{144}} \right) \approx (19.19, 21.0)$$

- c) We are 99% confident that the interval from 19.19% to 21.0% contains the true mean monthly percentage of on-time flight arrivals.

21. For example, 2nd look.

The 95% confidence interval lies entirely above the 0.08 ppm limit. This is evidence that mirex contamination is too high and consistent with rejecting the null hypothesis. We used an upper-tail test, so the P-value should be smaller than $\frac{1}{2}(1 - 0.95) = 0.025$, and it was.

22. Hot dogs.

The 90% confidence interval contains the 325 mg limit. They can't assert that the mean sodium content is less than 325 mg, consistent with not rejecting the null hypothesis. They used an upper-tail test, so the P-value should be more than $\frac{1}{2}(1 - 0.90) = 0.05$, and it was.

23. Pizza.

If in fact the mean cholesterol of pizza eaters does not indicate a health risk, then only 7 out of every 100 samples would be expected to have mean cholesterol as high or higher than the mean cholesterol observed in the sample.

24. Golf balls.

If in fact this ball meets the velocity standard, then 34% of samples tested would be expected to have mean speeds at least as high as the mean speed recorded in the sample.

25. TV Safety.

- a) The inspectors are performing an upper-tail test. They need to prove that the stands will support 500 pounds (or more) easily.
- b) The inspectors commit a Type I error if they certify the stands as safe, when they are not.
- c) The inspectors commit a Type II error if they decide the stands are not safe, when they are.

26. Catheters.

- a) Quality control personnel are conducting a two-sided test. If the catheters are too big, they won't fit through the vein. If they are too small, the examination apparatus may not fit through the catheter.
- b) The quality control personnel commit a Type I error if catheters are rejected, when in fact the diameters are fine. The manufacturing process is stopped needlessly.
- c) The quality control personnel commit a Type II error if catheters are being produced that do not meet the specifications, and this goes unnoticed. Defective catheters are being produced and sold.

27. TV safety revisited.

- a) The value of α should be decreased. This means a smaller chance of declaring the stands safe under the null hypothesis that they are not safe.
- b) The power of the test is the probability of correctly detecting that the stands can safely hold over 500 pounds.
- c)
 - 1) The company could redesign the stands so that their strength is more consistent, as measured by the standard deviation. Redesigning the manufacturing process is likely to be quite costly.
 - 2) The company could increase the number of stands tested. This costs them both time to perform the test and money to pay the quality control personnel.
 - 3) The company could increase α , effectively lowering their standards for what is required to certify the stands "safe". This is a big risk, since there is a greater chance of Type I error, namely allowing unsafe stands to be sold.
 - 4) The company could make the stands stronger, increasing the mean amount of weight that the stands can safely hold. This type of redesign is expensive.

28. Catheters again.

- a) If the level of significance is lowered to $\alpha = 0.01$, the probability of Type II error will increase. Lowering α will lower the probability of incorrectly rejecting the null hypothesis when it's true, which will increase the probability of incorrectly *failing* to reject the null hypothesis when it is *false*.
- b) The power of this test is the probability of correctly detecting deviations from 2 mm in diameter.
- c) The power of the test will increase as the actual mean diameter gets farther and farther away from 2 mm. Larger deviations from what is expected are easier to detect than small deviations.
- d) In order to increase the power of the test, the company could increase the sample size, reducing the standard error of the sampling distribution model. They could also increase the value of α , requiring a lower standard of proof to identify a faulty manufacturing process.

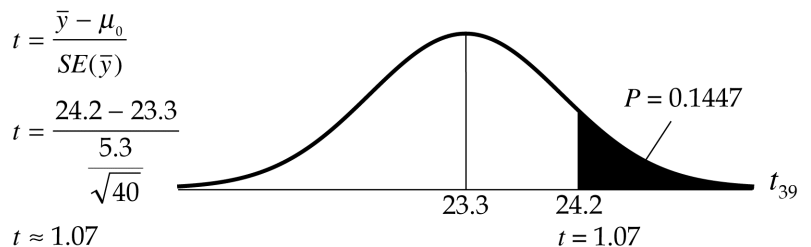
29. Marriage.

- a) H_0 : The mean age at which American men first marry is 23.3 years. ($\mu = 23.3$)
 H_A : The mean age at which American men first marry is greater than 23.3 years. ($\mu > 23.3$)
- b) **Randomization condition:** The 40 men were selected randomly.
10% condition: 40 men are less than 10% of all recently married men.
Nearly Normal condition: The population of ages of men at first marriage is likely to be skewed to the right. It is much more likely that there are men who marry for the first time at an older age than at an age that is very young. We should examine the distribution of the sample to check for serious skewness and outliers, but with a large sample of 40 men, it should be safe to proceed.
- c) Since the conditions for inference are satisfied, we can model the sampling distribution of the mean age of men at first marriage with $N\left(23.3, \frac{\sigma}{\sqrt{n}}\right)$. Since we do not know σ , the standard deviation of the population, $\sigma(\bar{y})$ will be estimated by $SE(\bar{y}) = \frac{s}{\sqrt{n}}$, and we will use a Student's t model, with $40 - 1 = 39$ degrees of freedom, $t_{39}\left(23.3, \frac{s}{\sqrt{40}}\right)$.

- d) The mean age at first marriage in the sample was 24.2 years, with a standard deviation in age of 5.3 years. Use a one-sample t -test, modeling the sampling distribution of \bar{y} with

$$t_{39}\left(23.3, \frac{5.3}{\sqrt{40}}\right).$$

The P -value is 0.1447.



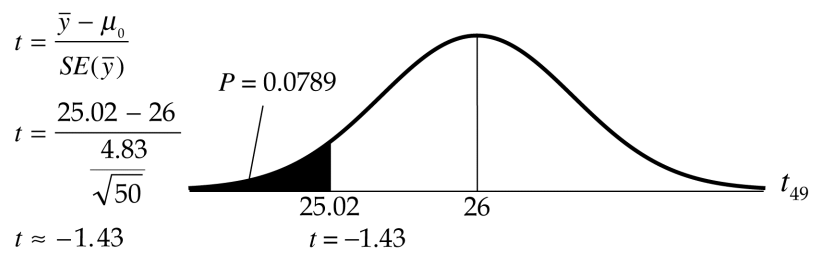
394 **Part VI Learning About the World**

- e) If the mean age at first marriage is still 23.3 years, there is a 14.5% chance of getting a sample mean of 24.2 years or older simply from natural sampling variation.
- f) Since the P -value = 0.1447 is high, we fail to reject the null hypothesis. There is no evidence to suggest that the mean age of men at first marriage has changed from 23.3 years, the mean in 1960.

30. Fuel economy.

- a) H_0 : The mean mileage of the cars in the fleet is 26 mpg. ($\mu = 26$)
 H_A : The mean mileage of the cars in the fleet is less than 26 mpg. ($\mu < 26$)
- b) **Randomization condition:** The 50 trips were selected randomly.
10% condition: 50 trips are less than 10% of all trips.
Nearly Normal condition: We don't have the actual data, so we cannot look at the distribution of the data, but the sample is large, so we can proceed.
- c) Since the conditions for inference are satisfied, we can model the sampling distribution of the mean mileage of cars in the fleet with $N\left(26, \frac{\sigma}{\sqrt{n}}\right)$. Since we do not know σ , the standard deviation of the population, $\sigma(\bar{y})$ will be estimated by $SE(\bar{y}) = \frac{s}{\sqrt{n}}$, and we will use a Student's t model, with $50 - 1 = 49$ degrees of freedom, $t_{49}\left(26, \frac{s}{\sqrt{50}}\right)$.

- d) The trips in the sample had a mean mileage of 25.02 mpg, with a standard deviation of 4.83 mpg. Use a one-sample t -test, modeling the sampling distribution of \bar{y} with $t_{49}\left(26, \frac{4.83}{\sqrt{50}}\right)$.

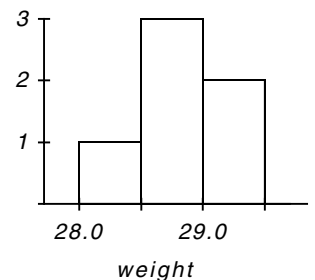


The P -value is 0.0789.

- e) If the mean mileage of cars in the fleet is 26 mpg, the chance that a sample mean of a sample of size 50 is 25.02 mpg or less simply due to sampling error is 7.9%.
- f) Since the P -value = 0.0789 is fairly high, we fail to reject the null hypothesis. There is little evidence to suggest that the mean mileage of cars in the fleet is less than 26 mpg.

31. Ruffles.

- a) **Randomization condition:** The 6 bags were not selected at random, but it is reasonable to think that these bags are representative of all bags of chips.
10% condition: 6 bags are less than 10% of all bags of chips.
Nearly Normal condition: The histogram of the weights of chips in the sample is nearly normal.



- b) $\bar{y} \approx 28.78$ grams, $s \approx 0.40$ grams

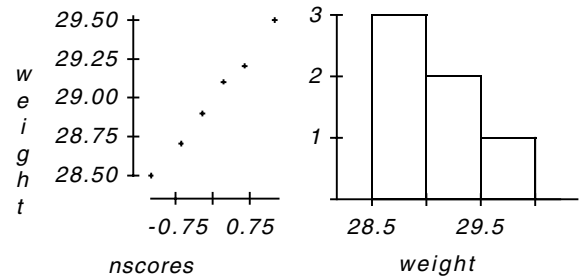
- c) Since the conditions for inference have been satisfied, use a one-sample t -interval, with $6 - 1 = 5$ degrees of freedom, at 95% confidence.

$$\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 28.78 \pm t_5^* \left(\frac{0.40}{\sqrt{6}} \right) \approx (28.36, 29.21)$$

- d) We are 95% confident that the mean weight of the contents of Ruffles bags is between 28.36 and 29.21 grams.
- e) Since the interval is above the stated weight of 28.3 grams, there is evidence that the company is filling the bags to more than the stated weight, on average.

32. Doritos.

- a) **Randomization condition:** The 6 bags were not selected at random, but it is reasonable to think that these bags are representative of all bags.
10% condition: 6 is less than 10% of all bags.
Nearly Normal condition: The Normal probability plot is reasonably straight. Although the histogram of the weights of chips in the sample is not symmetric, any apparent “skewness” is the result of a single bag of chips. It is safe to proceed.



- b) $\bar{y} \approx 28.98$ grams, $s \approx 0.36$ grams
- c) Since the conditions for inference have been satisfied, use a one-sample t -interval, with $6 - 1 = 5$ degrees of freedom, at 95% confidence.

$$\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 28.98 \pm t_5^* \left(\frac{0.36}{\sqrt{6}} \right) \approx (28.61, 29.36)$$

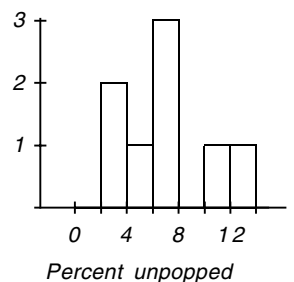
- d) We are 95% confident that the interval 28.61 to 29.36 grams contains the true mean weight of the contents of Doritos bags.
- e) Since the interval is above the stated weight of 28.3 grams, there is evidence that the company is filling the bags to more than the stated weight, on average.

33. Popcorn.

- a) Hopps made a Type I error. He mistakenly rejected the null hypothesis that the proportion of unpopped kernels was 10% (or higher).
- b) H_0 : The mean proportion of unpopped kernels is 10%. ($\mu = 10$)

H_A : The mean proportion of unpopped kernels is lower than 10%. ($\mu < 10$)

Randomization condition: The 8 bags were randomly selected.
10% condition: 8 bags are less than 10% of all bags.
Nearly Normal condition: The histogram of the percentage of unpopped kernels is unimodal and roughly symmetric.

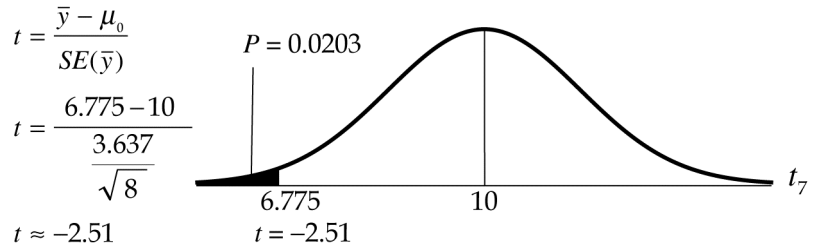


The bags in the sample had a mean percentage of unpopped kernels of 6.775 percent and a standard deviation in percentage of unpopped kernels of 3.637 percent. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean percentage of unpopped kernels with a Student's t model, with $8 - 1 = 7$ degrees of

freedom, $t_7\left(6.775, \frac{3.637}{\sqrt{8}}\right)$.

We will perform a one-sample t -test.

Since the P -value = 0.0203 is low, we reject the null hypothesis. There is evidence to suggest the mean percentage of unpopped kernels is less than 10% at this setting.



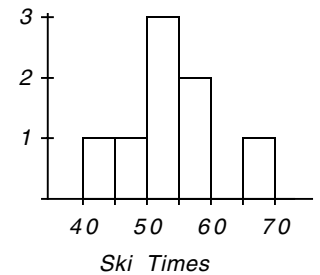
34. Ski wax.

- a) He would have made a Type II error. The null hypothesis, that his average time would be 55 seconds (or higher), was false, and he failed to realize this.
- b) H_0 : The mean time was 55 seconds. ($\mu = 55$)

H_A : The mean time was less than 55 seconds. ($\mu < 55$)

Independence assumption: Since the times are not randomly selected, we will assume that the times are independent, and representative of all times.

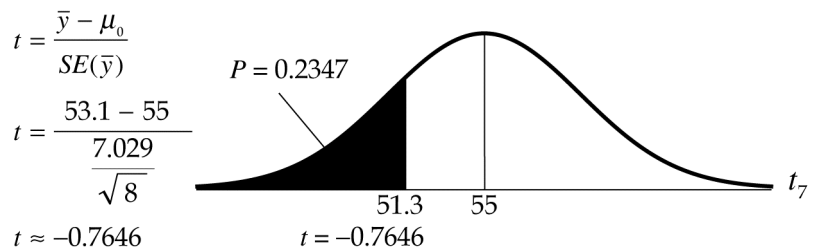
Nearly Normal condition: The histogram of the times is unimodal and roughly symmetric.



The times in the sample had a mean of 53.1 seconds and a standard deviation of 7.029 seconds. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean time with a Student's t model, with $8 - 1 = 7$ degrees of freedom,

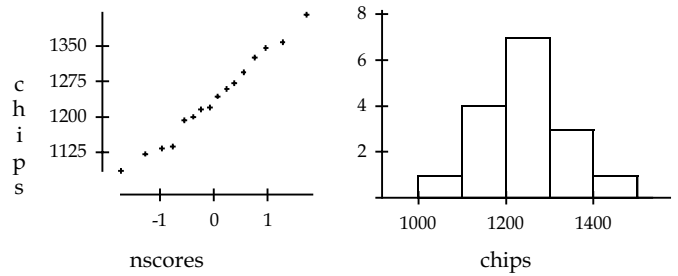
$t_7\left(53.1, \frac{7.029}{\sqrt{8}}\right)$. We will perform a one-sample t -test.

Since the P -value = 0.2347 is high, we fail to reject the null hypothesis. There is no evidence to suggest the mean time is less than 55 seconds. He should not buy the new ski wax.



35. Chips ahoy.

- a) **Randomization condition:** The bags of cookies were randomly selected.
10% condition: 16 bags are less than 10% of all bags
Nearly Normal condition: The Normal probability plot is reasonably straight, and the histogram of the number of chips per bag is unimodal and symmetric.



- b) The bags in the sample had with a mean number of chips or 1238.19, and a standard deviation of 94.282 chips. Since the conditions for inference have been satisfied, use a one-sample t -interval, with $16 - 1 = 15$ degrees of freedom, at 95% confidence.

$$\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 1238.19 \pm t_{15}^* \left(\frac{94.282}{\sqrt{16}} \right) \approx (1187.9, 1288.4)$$

We are 95% confident that the mean number of chips in an 18-ounce bag of Chips Ahoy cookies is between 1187.9 and 1288.4.

- c) H_0 : The mean number of chips per bag is 1000. ($\mu = 1000$)

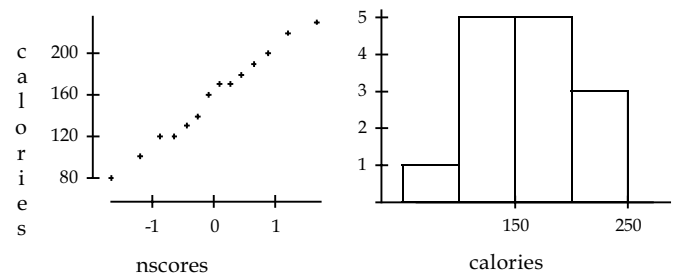
H_A : The mean number of chips per bag is greater than 1000. ($\mu > 1000$)

Since the confidence interval is well above 1000, there is strong evidence that the mean number of chips per bag is well above 1000.

However, since the "1000 Chip Challenge" is about individual bags, not means, the claim made by Nabisco may not be true. If the mean was around 1188 chips, the low end of our confidence interval, and the standard deviation of the population was about 94 chips, our best estimate obtained from our sample, a bag containing 1000 chips would be about 2 standard deviations below the mean. This is not likely to happen, but not an outrageous occurrence. These data do not provide evidence that the "1000 Chip Challenge" is true.

36. Yogurt.

- a) **Randomization condition:** The brands of vanilla yogurt may not be a random sample, but they are probably representative of all brands of yogurt.
10% condition: 14 brands of vanilla yogurt may not be less than 10% of all yogurt brands. Are there 140 brands of vanilla yogurt available?



Independence assumption: The Randomization Condition and the 10% Condition are designed to check the reasonableness of the assumption of independence. We had some trouble verifying these conditions. But is the calorie content per serving of one brand of yogurt likely to be associated with that of another brand? Probably not. We're okay.

Nearly Normal condition: The Normal probability plot is reasonably straight, and the histogram of the number of calories per serving is unimodal and symmetric.

- b) The brands in the sample had a mean calorie content of 157.857 calories, and a standard deviation of 44.7521 calories. Since the conditions for inference have been satisfied, use a one-sample t -interval, with $14 - 1 = 13$ degrees of freedom, at 95% confidence.

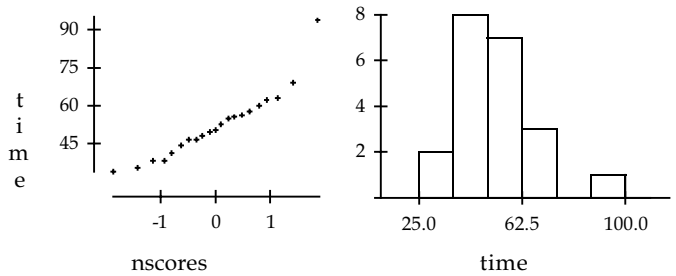
$$\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 157.857 \pm t_{13}^* \left(\frac{44.7521}{\sqrt{14}} \right) \approx (132.0, 183.7)$$

- c) We are 95% confident that the mean calorie content in a serving of vanilla yogurt is between 132.0 and 183.7 calories. There is evidence that the estimate of 120 calories made in the diet guide is too low. The 95% confidence interval is well above 120 calories.

37. Maze.

- a) **Independence assumption:** It is reasonable to think that the rats' times will be independent, as long as the times are for different rats.

Nearly Normal condition: There is an outlier in both the Normal probability plot and the histogram that should probably be eliminated before continuing the test. One rat took a long time to complete the maze.



- b) H_0 : The mean time for rats to complete this maze is 60 seconds. ($\mu = 60$)

H_A : The mean time for rats to complete this maze is not 60 seconds. ($\mu \neq 60$)

The rats in the sample finished the maze with a mean time of 52.21 seconds and a standard deviation in times of 13.5646 seconds. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean time in which rats complete the maze with a Student's t model, with

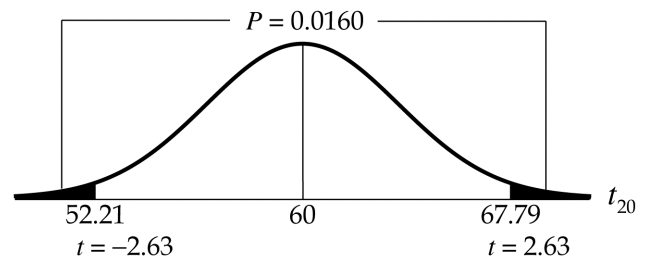
$21 - 1 = 20$ degrees of freedom, $t_{20} \left(60, \frac{13.5646}{\sqrt{21}} \right)$. We will perform a one-sample t -test.

Since the P -value = 0.0160 is low, we reject the null hypothesis. There is evidence that the mean time required for rats to finish the maze is not 60 seconds. Our evidence suggests that the mean time is actually less than 60 seconds.

$$t = \frac{\bar{y} - \mu_0}{SE(\bar{y})}$$

$$t = \frac{52.21 - 60}{\frac{13.56}{\sqrt{21}}}$$

$$t \approx -2.63$$



- c) Without the outlier, the rats in the sample finished the maze with a mean time of 50.13 seconds and standard deviation in times of 9.90 seconds. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean time in which rats complete the maze with a Student's t model, with $20 - 1 = 19$ degrees of freedom,

$t_{19} \left(60, \frac{9.90407}{\sqrt{20}} \right)$. We will perform a one-sample t -test.

This test results in a value of $t = -4.46$, and a two-sided P -value = 0.0003. Since the P -value is low, we reject the null hypothesis. There is evidence that the mean time required for rats to finish the maze is not 60 seconds. Our evidence suggests that the mean time is actually less than 60 seconds.

- d) According to both tests, there is evidence that the mean time required for rats to complete the maze is different than 60 seconds. The maze does not meet the “one-minute average” requirement. It should be noted that the test without the outlier is the appropriate test. The one slow rat made the mean time required seem much higher than it probably was.

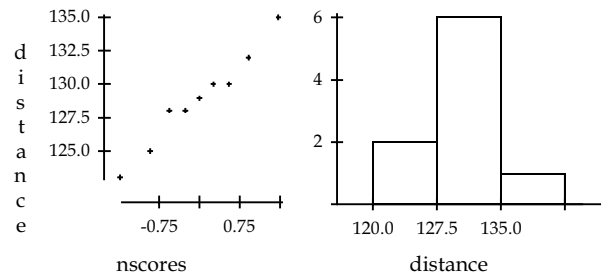
38. Braking.

H_0 : The mean braking distance is 125 feet. The tread pattern works adequately. ($\mu = 125$)

H_A : The mean braking distance is greater than 125 feet, and the new tread pattern should not be used. ($\mu > 125$)

Independence assumption: It is reasonable to think that the braking distances on the test track are independent of each other.

Nearly Normal condition: The braking distance of 102 feet is an outlier. After it is removed, the Normal probability plot is reasonably straight, and the histogram of braking distances unimodal and symmetric.



The braking distances in the sample had a mean of 128.889 feet, and a standard deviation of 3.55121 feet. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean braking distance with a Student’s t model, with $9 - 1 = 8$ degrees of freedom,

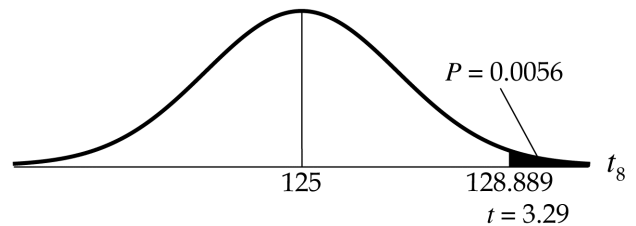
$t_9\left(125, \frac{3.55121}{\sqrt{9}}\right)$. We will perform a one-sample t -test.

Since the P -value = 0.0056 is low, we reject the null hypothesis. There is strong evidence that the mean braking distance of cars with these tires is greater than 125 feet. The new tread pattern should not be adopted.

$$t = \frac{\bar{y} - \mu_0}{SE(\bar{y})}$$

$$t = \frac{128.889 - 125}{\frac{3.55121}{\sqrt{9}}}$$

$$t \approx 3.29$$



39. Driving distance.

a) $\bar{y} \pm t_{n-1}^* \left(\frac{s}{\sqrt{n}} \right) = 288.6 \pm t_{201}^* \left(\frac{9.31}{\sqrt{202}} \right) \approx (287.3, 289.9)$

- b) These data are not a random sample of golfers. The top professionals are not representative of all golfers and were not selected at random. We might consider the 2006 data to represent the population of all professional golfers, past, present, and future.

400 **Part VI Learning About the World**

- c) The data are means for each golfer, so they are less variable than if we looked at separate drives.

40. **Wind power.**

- a) H_0 : The mean wind speed is 8 mph. It's not windy enough for a wind turbine. ($\mu = 8$)

H_A : The mean wind speed is greater than 8 mph. It's windy enough. ($\mu > 8$)

Independence assumption: The timeplot shows no pattern, so it seems reasonable that the measurements are independent.

Randomization condition: This is not a random sample, but an entire year is measured. These wind speeds should be representative of all wind speeds at this location.

10% condition: These wind speeds certainly represent fewer than 10% of all wind speeds.

Nearly Normal condition: The Normal probability plot is reasonably straight, and the histogram of the wind speeds is unimodal and reasonably symmetric.

The wind speeds in the sample had a mean of 8.091 mph, and a standard deviation of 3.813 mph. Since the conditions for inference are satisfied, we can model the sampling distribution of the mean wind speed with a Student's t model, with $1114 - 1 = 1113$ degrees

of freedom, $t_{1113} \left(8, \frac{3.813}{\sqrt{1114}} \right)$. We will perform a one-sample t -test.

Since the P -value = 0.43 is high, we fail to reject the null hypothesis. There is no evidence that the mean wind speed at this site is higher than 8 mph. Even though the mean wind speed for these 1114 measurements is 8.019 mph, I wouldn't recommend building a wind turbine at this site.

$$t = \frac{\bar{y} - \mu_0}{SE(\bar{y})}$$

$$t = \frac{8.019 - 8}{\frac{3.813}{\sqrt{1114}}}$$

$$t \approx 0.1663$$

