Chapter 17 – Confidence Intervals for Means

Section 17.1

1. Salmon.

a) The shipment of 4 salmon has $SD(\bar{y}) = \frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{4}} = 1$ pound.

The shipment of 16 salmon has $SD(\bar{y}) = \frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{16}} = 0.5$ pounds.

The shipment of 100 salmon has $SD(\bar{y}) = \frac{\sigma}{\sqrt{n}} = \frac{2}{\sqrt{100}} = 0.2$ pounds.

b) The Normal model would better characterize the shipping weight of the pallets than the shipping weight of the boxes, since the pallets contain a large number of salmon. The Central Limit Theorem tells us that the distribution of means (and therefore totals) approaches the Normal model, regardless of the underlying distribution. As samples get larger, the approximation gets better.

2. LSAT.

a) Since the distribution of LSAT scores for all test takers is unimodal and symmetric, the distribution of scores for the test takers at these test preparation organizations are probably at least free of outliers and skewness. Therefore, the distribution of the mean score of classes of size 9 and 25 should each be roughly Normal, and each would have a mean of 151 points. The standard deviation of the distribution of the means would be $\frac{\sigma}{\sqrt{n}} = \frac{9}{\sqrt{9}} = 3$ points for the class of size 9 and $\frac{\sigma}{\sqrt{n}} = \frac{9}{\sqrt{25}} = 1.8$ points for the class of size 25.

b) The organization with the smaller class has a larger standard deviation of the mean. A class mean score of 160 is 3 standard deviations above the population mean, which is rare, but could happen. For the larger organization, a class mean of 160 is 5 standard deviations above the population mean, which would be highly unlikely

c) The smaller organization is at a greater risk of having to pay for LSAT retakes. They are more likely to have a low class mean for the same reason they are more likely to have a high class mean. The variability in the class mean score is greater when the class size is small.

3. Tips.

a) Since the distribution of tips is skewed to the right, we can't use the Normal model to determine the probability that a given party will tip at least \$20.

b) No. A sample of 4 parties is probably not a large enough sample for the CLT to allow us to use the Normal model to estimate the distribution of averages.

c) A sample of 10 parties may not be large enough to allow the use of a Normal model to describe the distribution of averages. It would be risky to attempt to estimate the probability that his next 10 parties tip an average of \$15. However, since the distribution of tips has $\mu = \$9.60$, with standard deviation $\sigma = \$5.40$, we still know that the mean of the sampling distribution model is $\mu_{\bar{y}} = \$9.60$ with standard deviation $SD(\bar{y}) = \frac{5.40}{\sqrt{10}} \approx \$1.71$.

We don't know the exact shape of the distribution, but we can still assess the likelihood of specific means. A mean tip of \$15 is over 3 standard deviations above the expected mean tip for 10 parties. That's not very likely to happen.

4. Groceries.

a) Since the distribution of Sunday purchases is skewed, we can't use the Normal model to determine the probability that a given purchase is at least \$40.

b) A sample of 10 customers may not be large enough for the CLT to allow the use of a Normal model for the sampling distribution model. If the distribution of Sunday purchases is only slightly skewed, the sample may be large enough, but if the distribution is heavily skewed, it would be very risky to attempt to determine the probability.

c) **Randomization condition:** Assume that the 50 Sunday purchases can be considered a representative sample of all purchases.

Independence assumption: It is reasonable to think that the Sunday purchases are mutually independent, unless there is a sale or other incentive to purchase more.

10% condition: The 50 purchases certainly represent less than 10% of all purchases.

Large Enough Sample condition: The sample of 50 purchases is large enough.

The mean Sunday purchase is $\mu = \$32$, with standard deviation $\sigma = \$20$. Since the conditions are met, the CLT allows us to model the sampling distribution of $\bar{y}$ with a Normal model, with $\mu_{\bar{y}} = \$32$ and standard deviation $SD(\bar{y}) = \frac{20}{\sqrt{50}} \approx \$2.83$.

$z = \frac{y-\mu}{\sigma/\sqrt{n}} = \frac{40-32}{20/\sqrt{50}} \approx 2.828$; According to the Normal model, the probability the mean Sunday purchase of 50 customers is at least \$40 is about 0.0023.

5. More tips.

a) **Randomization condition:** Assume that the tips from 40 parties can be considered a representative sample of all tips.

Independence assumption: It is reasonable to think that the tips are mutually independent, unless the service is particularly good or bad during this weekend.

10% condition: The tips of 40 parties certainly represent less than 10% of all tips.

Large Enough Sample condition: The sample of 40 parties is large enough.

The mean tip is $\mu = \$9.60$, with standard deviation $\sigma = \$5.40$. Since the conditions are satisfied, the CLT allows us to model the sampling distribution of $\bar{y}$ with a Normal model, with $\mu_{\bar{y}} = \$9.60$ and standard deviation $SD(\bar{y}) = \frac{5.40}{\sqrt{40}} \approx \$0.8538$.

In order to earn at least \$500, the waiter would have to average $\frac{500}{40} = \$12.50$ per party.

$z = \frac{y-\mu}{\sigma/\sqrt{n}} = \frac{12.50-9.60}{5.40/\sqrt{40}} \approx 3.397$; According to the Normal model, the probability that the waiter earns at least \$500 in tips in a weekend is approximately 0.0003.

5. (continued)

b) According to the Normal model, the waiter can expect to have a mean tip of about \$10.6942, which corresponds to about \$427.77 for 40 parties, in the best 10% of such weekends.

$$z = \frac{\bar{y} - \mu_{\bar{y}}}{SD(\bar{y})}$$
$$1.2816 = \frac{\bar{y} - 9.60}{\frac{5.40}{\sqrt{40}}}$$
$$\bar{y} \approx 10.6942$$

6. More groceries.

a) Assumptions and conditions for the use of the CLT were verified in Exercise 4.

The mean purchase is $\mu = \$32$, with standard deviation $\sigma = \$20$. Since the sample is large, the CLT allows us to model the sampling distribution of $\bar{y}$ with a Normal model, with $\mu_{\bar{y}} = \$32$ and standard deviation $SD(\bar{y}) = \frac{20}{\sqrt{312}} \approx \$1.1323$.

In order to have revenues of at least \$10,000, the mean Sunday purchase must be at least $\frac{10,000}{312} = \$32.0513$.

$z = \frac{y - \mu}{\sigma/\sqrt{n}} = \frac{32.0513 - 32}{20/\sqrt{312}} \approx 0.0453$; According to the Normal model, the probability of having a mean Sunday purchase at least that high (and therefore at total revenue of at least \$10,000) is 0.482.

b) According to the Normal model, the mean Sunday purchase on the worst 10% of such days is approximately \$30.548928, so 312 customers are expected to spend about \$9531.27.

$$z = \frac{\bar{y} - \mu_{\bar{y}}}{\sigma(\bar{y})}$$
$$-1.281552 = \frac{\bar{y} - 32}{\frac{20}{\sqrt{312}}}$$
$$\bar{y} \approx 30.548928$$

Section 17.2

7. *t*-models, part I.

a) 1.74

b) 2.37

8. *t*-models, part II.

a) 2.36

b) 2.62

9. *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.

10. *t*-models, part IV.

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.

11. Home sales.

a) The estimates of home value losses must be independent. This is verified using the Randomization condition, since the houses were randomly sampled. The distribution of home value losses must be Normal. A histogram of home value losses in the sample would be checked to verify this, using the Nearly Normal condition. Even if the histogram is not unimodal and symmetric, the sample size of 36 should allow for some small departures from Normality.

b) $\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 9560 \pm t_{35}^{*}\left(\frac{1500}{\sqrt{36}}\right) \approx (9052.50, 10067.50)$

12. Home sales, again.

a) A larger standard deviation in home value losses would increase the width of the confidence interval.

b) Your classmate is correct. A lower confidence level results in a narrower interval.

c) A larger sample would reduce the standard error, since larger samples result in lower variability in the distribution of means than smaller samples, which makes the interval narrower. This is more statistically appropriate, since we could narrow the interval without sacrificing confidence. However, it may be difficult or expensive to increase the sample size, so it may not be practical.

Section 17.3

13. Home sales revisited.

We are 95% confident that the interval $9052.50 to $10,067.50 contains the true mean loss in home value. That is, 95% of all random samples of size 36 will contain the true mean.

14. Salaries.

We are 95% confident that the true mean is somewhere in this interval, but it is equally likely to be anywhere in this interval because the interval itself is random. In particular, the true mean is not more likely to be in the middle of the interval. We are not sure where in this interval the true mean will actually be located.

15. 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.

16. 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!

16. (continued)

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.

Section 17.4

17. Framingham revisited.

a) To find a 95% confidence interval for the mean, we need to exclude the most extreme 5% of bootstrap sample means, 2.5% from each side. Using the 2.5th percentile and the 97.5th percentile, the interval is 232.434 to 237.130.

b) We are 95% confident that the true mean cholesterol level is between 232.4 and 237.1 mg/dL.

c) We must assume that the individual values in the sample are independent. In this case, the assumption is valid, since the sample was random.

18. Student survey revisited.

a) To find a 98% confidence interval for the mean, we need to exclude the most extreme 2% of bootstrap sample means, 1% from each side. Using the 1st percentile and the 99th percentile, the interval is 702.498 to 797.358.

b) We are 95% confident that the true mean number of Facebook friends a student has is between 702.5 and 797.4.

c) We must assume that the individual values in the sample are independent. In this case, the assumption is valid, since the sample was random.

Section 17.5

19. Shoe sizes revisited.

The interval is a range of possible values for the mean shoe size. The average is not a value that any individual in the population will have, but an average of all the individuals.

20. Bird counts.

The confidence interval contains a range of values for the mean. The average number of species counted does not have to be a whole number.

21. 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 $1467.

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.

22. Snow.

a) Not correct. The confidence interval is not about the winters in the sample.

b) Not correct. The confidence interval does not predict what will happen in any one winter.

c) Not correct. The confidence interval is not based on a sample of days.

22. (continued)

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.

Chapter Exercises

23. 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.

24. Crawling.

a) We are 95% confident that the interval 30.65 to 32.89 weeks contains the true mean age at which babies begin to crawl.

b) The width of the interval is about 32.89 – 30.65 = 2.24 weeks. The margin of error is half of that, about 1.12 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.

25. 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.

26. Credit card charges.

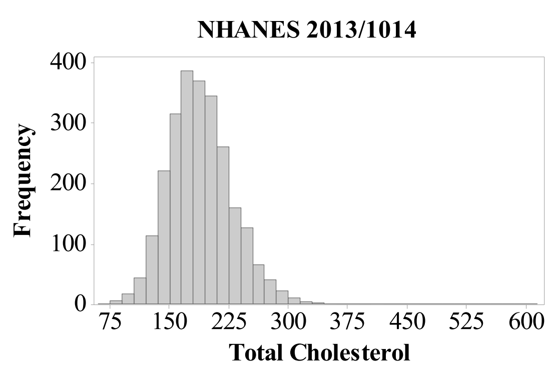
The analysts did not find the confidence interval useful because the conditions for inference were not met. There is one cardholder who spent over $3,000,000 on his card. This made the standard deviation, and therefore the standard error, huge. The *t*-interval is too wide to be of any use.

27. Cholesterol.

a) We cannot apply the Central Limit Theorem to describe the distribution of cholesterol measurements. The Central Limit Theorem is about means and proportions, not individual observations.

b) **Randomization condition:** Although not specifically stated, it is safe to assume that the 2515 adults are a random sample of US adults.

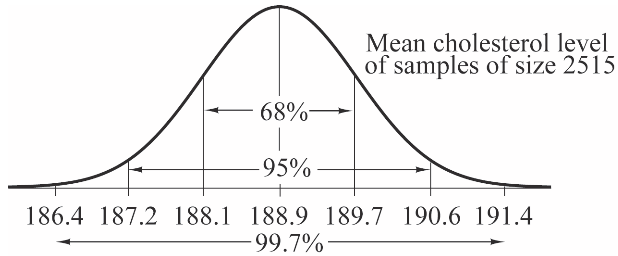
Nearly Normal Condition: The distribution of the sample of cholesterol levels is skewed to the right, with several outliers on both ends. However, the sample size is very large, and the Central Limit Theorem will allow us to us a Normal model to describe the mean cholesterol level of samples of 2515 adults.



27. (continued)

Therefore, the sampling distribution model for the mean cholesterol level of 2515 US adults is $N\left(188.9, \frac{41.6}{\sqrt{2515}}\right)$ or $N(188.9, 0.83)$.

c) The Normal model is shown below.

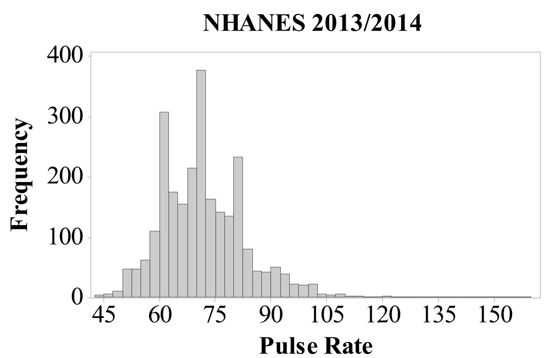


28. Pulse rates.

a) We cannot apply the Central Limit Theorem to describe the distribution of pulse rates. The Central Limit Theorem is about means and proportions, not individual observations.

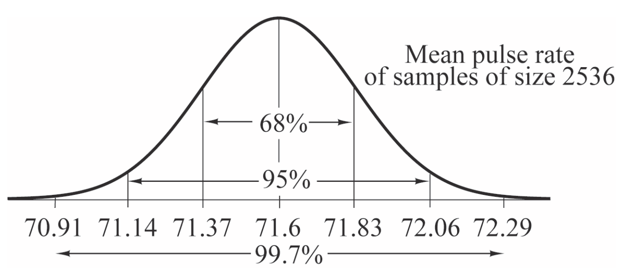
b) **Randomization condition:** Although not specifically stated, it is safe to assume that the 2536 adults are a random sample of US adults.

Nearly Normal Condition: The distribution of the sample of pulse rates is skewed to the right, with outliers on the upper end. However, the sample size is very large, and the Central Limit Theorem will allow us to us a Normal model to describe the mean cholesterol level of samples of 2536 adults.



Therefore, the sampling distribution model for the mean pulse rate of 2536 US adults is $N\left(71.6, \frac{11.5}{\sqrt{2536}}\right)$ or $N(71.6, 0.23)$.

c) The Normal model is shown below.



29. Normal temperature.

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

30. Parking.

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

31. 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.

31. (continued)

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.

32. 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.

33. 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.

34. Michelson.

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 27) 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.

35. Flight on time 2016.

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.
Nearly Normal condition: The histogram looks unimodal, and slightly skewed to the left. Since the sample size is 270, this should not be of concern.

b) The on-time departure rates in the sample had a mean of 78.099%, and a standard deviation in of 5.010%. Since the conditions have been satisfied, construct a one-sample t -interval, with 270 – 1 = 269 degrees of freedom, at 90% confidence.

$$\overline{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 78.099 \pm t_{269}^{*}\left(\frac{5.010}{\sqrt{270}}\right) \approx (77.596, 78.602)$$

c) We are 90% confident that the interval from 77.60% to 78.60% contains the true mean monthly percentage of on-time flight departures.

36. Flight on time 2016 revisited.

a) **Randomization condition:** Since there is no time trend, the monthly delay rates should be independent. This is not a random sample, but should be representative.
Nearly Normal condition: The histogram looks unimodal and symmetric.

b) The delay rates in the sample had a mean of 19.696%, and a standard deviation in of 4.214%. Since the conditions have been satisfied, construct a one-sample t -interval, with 270 – 1 = 269 degrees of freedom, at 99% confidence.

$$\overline{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 19.696 \pm t_{269}^{*}\left(\frac{4.217}{\sqrt{270}}\right) \approx (19.03, 20.362)$$

c) We are 99% confident that the interval from 19.03% to 20.36% contains the true mean monthly percentage of delayed flights.

37. Farmed salmon, second 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.

38. 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.

39. Pizza.

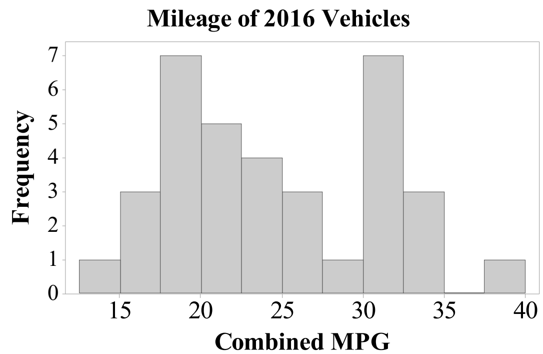
Because even the lower bound of the confidence interval is above 220 mg/dL, we are quite confident that the mean cholesterol level for those who eat frozen pizza is a level that indicates a health risk.

40. Golf balls.

Because 250 feet per second is in the interval, the USGA has no grounds from this experiment to reject the new ball for having too high a mean initial velocity.

41. Fuel economy 2016 revisited.

a) **Randomization condition:** The 35 cars were not selected randomly. We will have to assume that they are representative of all 2016 automobiles.
Nearly Normal condition: The distribution doesn't appear to be unimodal and symmetric, but the sample size is reasonably large.



The mileages in the sample had a mean of 24.3429 mpg, and a standard deviation in of 6.53021 mpg. Since the conditions have been satisfied, construct a one-sample t -interval, with 35 – 1 = 34 degrees of freedom, at 95% confidence.

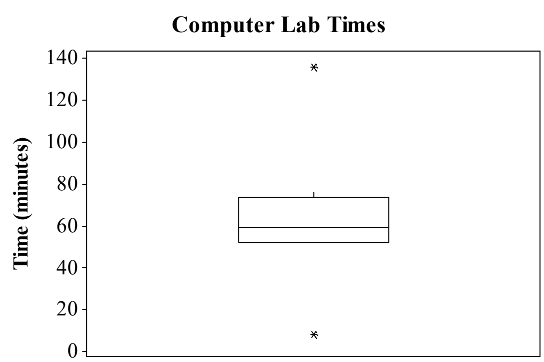
$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 24.3429 \pm t_{34}^{*}\left(\frac{6.53021}{\sqrt{35}}\right) \approx (22.1, 26.6)$$

We are 95% confident that the interval from 22.1 to 26.6 contains the true mean mileage of 2016 automobiles.

b) The data is a mix of small, mid-size, and large vehicles. Our histogram provides some evidence that there may be at least two distinct groups with regards to mileage. Without knowing how the data were selected, we are cautious about generalizing to all 2016 cars.

42. Computer lab fees.

a) The 8 minute and 136 minute times were extreme outliers. If the outliers are included, the conditions for inference, specifically the Nearly Normal condition, are not met.



b) With outliers: $\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 63.25 \pm t_{11}^{*}\left(\frac{28.927}{\sqrt{12}}\right) \approx (44.9, 81.6)$

Without outliers: $\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 61.5 \pm t_{9}^{*}\left(\frac{9.595}{\sqrt{10}}\right) \approx (54.6, 68.4)$

In either case, we would be reluctant to conclude that the mean is above 55 minutes. The small sample size and the presence of two large outliers causes us to be cautious about conclusions from this sample.

43. Waist size.

a) The distribution of waist size of 250 men is unimodal and slightly skewed to the right. A typical waist size is approximately 36 inches, and the standard deviation in waist sizes is approximately 4 inches.

b) All of the histograms show distributions of sample means centered near 36 inches. As n gets larger the histograms approach the Normal model in shape, and the variability in the sample means decreases. The histograms are fairly Normal by the time the sample reaches size 5.

44. CEO compensation.

a) The distribution of total compensation for the CEOs for the 800 largest U.S. companies is unimodal, but skewed to the right with several large outliers.

b) All of the histograms are centered near \$10,000,000. As n gets larger, the variability in sample means decreases, and histograms approach the Normal shape. However, they are still visibly skewed to the right, with the possible exception of the histogram for $n = 200$.

c) This rule of thumb doesn't seem to be true for highly skewed distributions.

45. Waist size, revisited.

a)

n	Observed Mean	Theoretical Mean	Observed Standard Deviation	Theoretical Standard Deviation
2	36.314	36.33	2.855	$4.019/\sqrt{2} \approx 2.842$
5	36.314	36.33	1.805	$4.019/\sqrt{5} \approx 1.797$
10	36.341	36.33	1.276	$4.019/\sqrt{10} \approx 1.271$
20	36.339	36.33	0.895	$4.019/\sqrt{20} \approx 0.899$

b) The observed values are all very close to the theoretical values.

c) For samples as small as 5, the sampling distribution of sample means is unimodal and symmetric. The Normal model would be appropriate.

d) The distribution of the original data is nearly unimodal and symmetric, so it doesn't take a very large sample size for the distribution of sample means to be approximately Normal.

46. CEOs, revisited.

a)

n	Observed Mean	Theoretical Mean	Observed Standard Deviation	Theoretical Standard Deviation
30	10,251.73	10,307.31	3359.64	$17{,}964.62/\sqrt{30} \approx 3279.88$
50	10,343.93	10,307.31	2483.84	$17{,}964.62/\sqrt{50} \approx 2540.58$
100	10,329.94	10,307.31	1779.18	$17{,}964.62/\sqrt{100} \approx 1796.46$
200	10,340.37	10,307.31	1260.79	$17{,}964.62/\sqrt{200} \approx 1270.29$

b) The observed values are all very close to the theoretical values.

c) All the sampling distributions are still quite skewed, with the possible exception of the sampling distribution for $n = 200$, which is still somewhat skewed. The Normal model would not be appropriate.

d) The distribution of the original data is strongly skewed, so it will take a very large sample size before the distribution of sample means is approximately Normal.

47. GPAs.

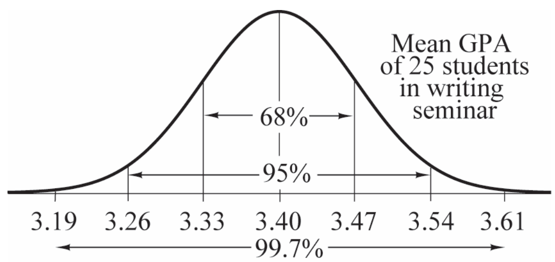
Randomization condition: Assume that the students are randomly assigned to seminars.

Independence assumption: It is reasonable to think that GPAs for randomly selected students are mutually independent.

Nearly Normal condition: The distribution of GPAs is roughly unimodal and symmetric, so the sample of 25 students is large enough.

The mean GPA for the freshmen was $\mu = 3.4$, with standard deviation $\sigma = 0.35$. Since the conditions are met, the Central Limit Theorem tells us that we can model the sampling distribution of the mean GPA with a Normal model, with $\mu_{\bar{y}} = 3.4$ and standard deviation $SD(\bar{y}) = \frac{0.35}{\sqrt{25}} \approx 0.07$.

The sampling distribution model for the sample mean GPA is approximately $N(3.4, 0.07)$.



48. Home values.

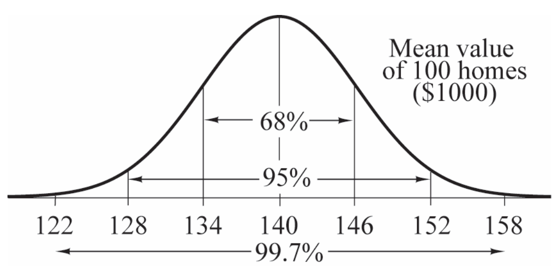
Randomization condition: Homes were selected at random.

Independence assumption: It is reasonable to think that assessments for randomly selected homes are mutually independent.

Nearly Normal condition: A sample of 100 homes is large enough for the Central Limit Theorem to apply.

The mean home value was $\mu = \$140,000$, with standard deviation $\sigma = \$60,000$. Since the conditions are met, the Central Limit Theorem tells us that we can model the sampling distribution of the mean home value with a Normal model, with $\mu_{\bar{y}} = \$140,000$ and standard deviation $SD(\bar{y}) = 60,000/\sqrt{100} = \$6000$.

The sampling distribution model for the sample mean home values is approximately $N(140000, 6000)$.



49. Lucky spot?

a) Smaller outlets have more variability than the larger outlets, just as the Central Limit Theorem predicts.

b) If the lottery is truly random, all outlets are equally likely to sell winning tickets.

50. Safe cities.

The standard deviation of the sampling model for the mean is $\sigma/\sqrt{n}$. So, cities in which the average is based on a smaller number of drivers will have greater variation in their averages and will be more likely to be both safest and least safe.

51. Pregnancy.

a) $z = \dfrac{y-\mu}{\sigma} = \dfrac{270-266}{16} = 0.25$ and $z = \dfrac{y-\mu}{\sigma} = \dfrac{280-266}{16} = 0.875$; According to the Normal model, approximately 21.1% of all pregnancies are expected to last between 270 and 280 days.

b) According to the Normal model, the longest 25% of pregnancies are expected to last approximately 276.8 days or more.

$$z = \frac{y-\mu}{\sigma}$$
$$0.674 = \frac{y-266}{16}$$
$$y \approx 276.8 \text{ days}$$

c) **Randomization condition:** Assume that the 60 women the doctor is treating can be considered a representative sample of all pregnant women.

Independence assumption: It is reasonable to think that the durations of the patients' pregnancies are mutually independent.

Nearly Normal condition: The distribution of pregnancy durations is Normal.

The mean duration of the pregnancies was $\mu = 266$ days, with standard deviation $\sigma = 16$ days. Since the distribution of pregnancy durations is Normal, we can model the sampling distribution of the mean pregnancy duration with a Normal model, with $\mu_{\bar{y}} = 266$ days and standard deviation

$$SD(\bar{y}) = \frac{16}{\sqrt{60}} \approx 2.07 \text{ days.}$$

d) $z = \dfrac{y-\mu}{\sigma/\sqrt{n}} = \dfrac{260-266}{16/\sqrt{50}} \approx -2.899$; According to the Normal model, with mean 266 days and standard deviation 2.07 days, the probability that the mean pregnancy duration is less than 260 days is 0.002.

52. Rainfall.

a) $z = \dfrac{y-\mu}{\sigma} = \dfrac{40-35.4}{4.2} \approx 1.095$; According to the Normal model, Ithaca is expected to get more than 40 inches of rain in approximately 13.7% of years.

b) According to the Normal model, Ithaca is expected to get less than 31.9 inches of rain in driest 20% of years.

$$z = \frac{y-\mu}{\sigma}$$
$$-0.842 = \frac{y-35.4}{4.2}$$
$$y \approx 31.9$$

52. (continued)

c) **Randomization condition:** Assume that the 4 years in which the student was in Ithaca can be considered a representative sample of all years.

Independence assumption: It is reasonable to think that the rainfall totals for the years are mutually independent.

Nearly Normal condition: The distribution of annual rainfall is Normal.

The mean rainfall was $\mu = 35.4$ inches, with standard deviation $\sigma = 4.2$ inches. Since the distribution of yearly rainfall is Normal, we can model the sampling distribution of the mean annual rainfall with a Normal model, with $\mu_{\bar{y}} = 35.4$ inches and standard deviation $SD(\bar{y}) = \dfrac{4.2}{\sqrt{4}} = 2.1$ inches.

d) $z = \dfrac{y-\mu}{\sigma/\sqrt{n}} = \dfrac{30-35.4}{4.2/\sqrt{4}} \approx -2.571$; According to the Normal model, with mean 35.4 inches and standard deviation 2.4 inches, the probability that those four years averaged less than 30 inches of rain is 0.005.

53. Pregnant again.

a) The distribution of pregnancy durations may be skewed to the left since there are more premature births than very long pregnancies. Modern practice of medicine stops pregnancies at about 2 weeks past normal due date by inducing labor or performing a Caesarean section.

b) We can no longer answer the questions posed in parts (a) and (b). The Normal model is not appropriate for skewed distributions. The answer to part (c) is still valid. The Central Limit Theorem guarantees that the sampling distribution model is Normal when the sample size is large.

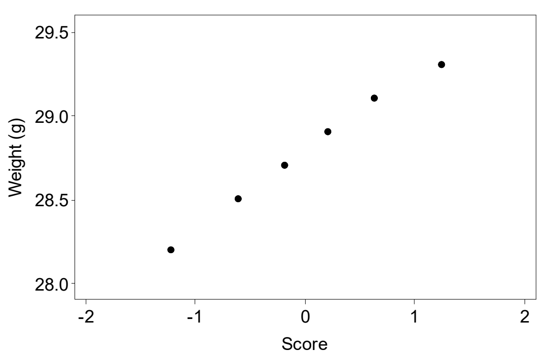
54. At work.

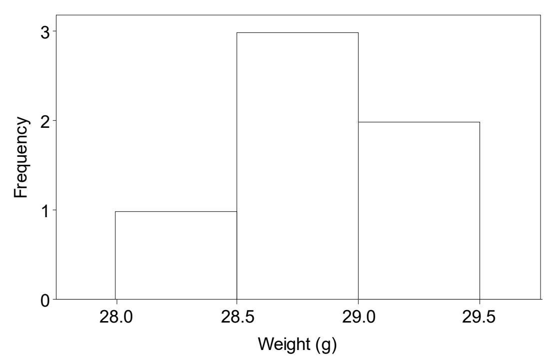
a) The distribution of length of time people work at a job is likely to be skewed to the right, because some people stay at the same job for much longer than the mean plus two or three standard deviations. Additionally, the left tail cannot be long, because a person cannot work at a job for less than 0 years.

b) The Central Limit Theorem guarantees that the distribution of the mean time is Normally distributed for large sample sizes, as long as the assumptions and conditions are satisfied. The Central Limit Theorem doesn't help us with the distribution of individual times.

55. 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.
Nearly Normal condition: The histogram of the weights of chips in the sample is nearly normal.





55. (continued)

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.

$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 28.78 \pm t_{5}^{*}\left(\frac{0.40}{\sqrt{6}}\right) \approx (28.36,\ 29.21)$$

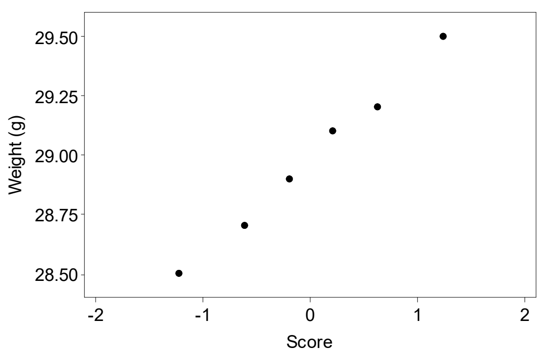
d) We are 95% confident that the mean weight of the contents of Ruffles bags is between 28.36 and 29.21 grams.

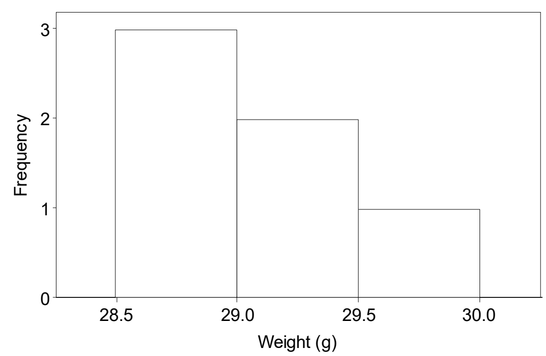
e) Since the interval is above the stated weight of 28.3 grams, there is evidence that the company is filling the bags to more than the stated weight, on average.

f) The sample size of 6 bags is too small to provide a good basis for a bootstrap confidence interval.

56. 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.
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.

$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 28.98 \pm t_{5}^{*}\left(\frac{0.36}{\sqrt{6}}\right) \approx (28.61,\ 29.36)$$

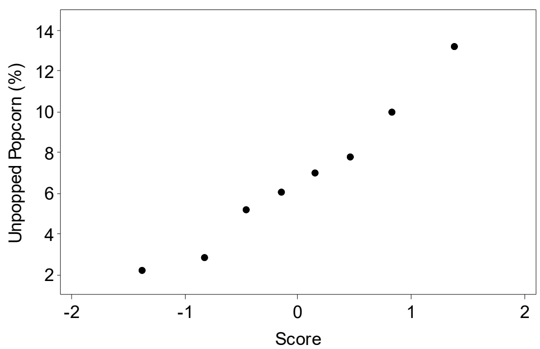
d) We are 95% confident that the interval 28.61 to 29.36 grams contains the true mean weight of the contents of Doritos bags.

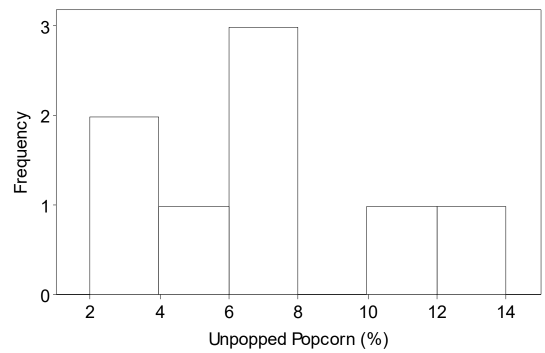
e) Since the interval is above the stated weight of 28.3 grams, there is evidence that the company is filling the bags to more than the stated weight, on average.

f) The sample size of 6 bags is too small to provide a good basis for a bootstrap confidence interval.

57. Popcorn.

Randomization condition: The 8 bags were randomly selected.
Nearly Normal condition: The histogram of the percentage of unpopped kernels is unimodal and roughly symmetric.





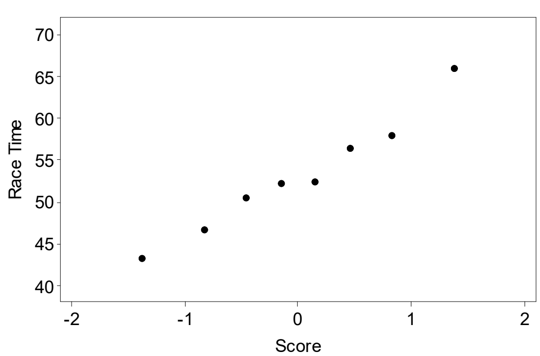
The bags in the sample had a mean percentage of unpopped kernels of 6.775 percent and a standard deviation in percentage of unpopped kernels of 3.637 percent. Since the conditions for inference are satisfied, we can use a one-sample t-interval, with 8 – 1 = 7 degrees of freedom, at 95% confidence.

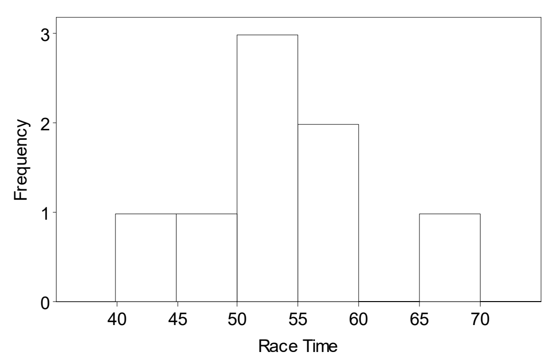
$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 6.775 \pm t_{7}^{*}\left(\frac{3.637}{\sqrt{8}}\right) \approx (3.7344,\ 9.8156)$$

We are 95% confident that the true mean percentage of unpopped kernels is contained in the interval 3.73% to 9.82%. Since 10% is not contained in the interval, there is evidence that Yvon Hopps has met his goal of an average of no more than 10% unpopped kernels.

58. Ski wax.

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.





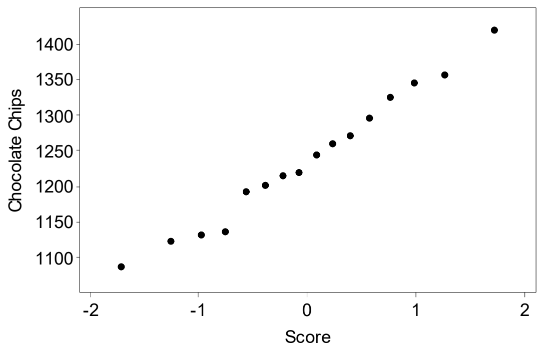
The times in the sample had a mean of 53.1 seconds and a standard deviation of 7.029 seconds. Since the conditions for inference are satisfied, we can use a one-sample t-interval, with 8 – 1 = 7 degrees of freedom, at 95% confidence.

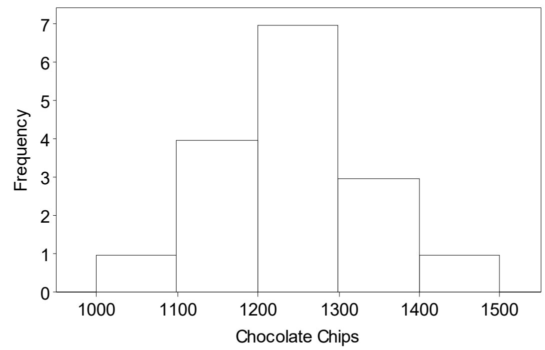
$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 53.1 \pm t_{7}^{*}\left(\frac{7.029}{\sqrt{8}}\right) \approx (47.224,\ 58.976)$$

We are 95% confident that the true mean course time is contained in the interval 47.22 seconds to 58.98 seconds. This interval contains 55 seconds, so he cannot be sure that this wax will help him average under 55 seconds.

59. Chips ahoy.

a) **Randomization condition:** The bags of cookies were randomly selected.
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 a mean number of chips of 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.

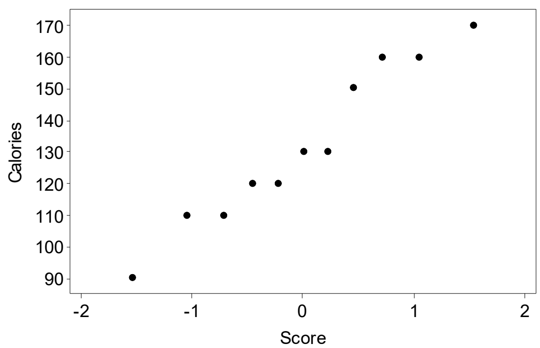
c) Since the confidence interval is well above 1000, there is strong evidence that the mean number of chips per bag is well above 1000.

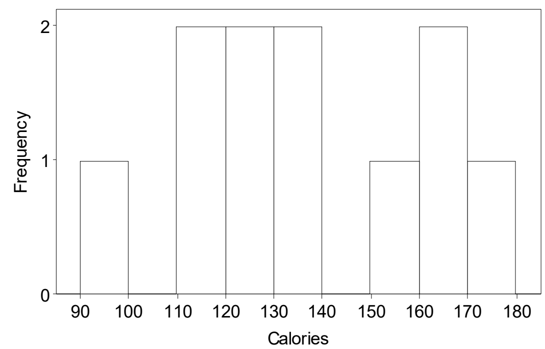
However, since the "1000 Chip Challenge" is about individual bags, not means, the claim made by Nabisco may not be true. If the mean was around 1188 chips, the low end of our confidence interval, and the standard deviation of the population was about 94 chips, our best estimate obtained from our sample, 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.

60. 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.
Independence assumption: The Randomization Condition is designed to check the reasonableness of the assumption of independence. We had some trouble verifying this condition. 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 plausibly unimodal and symmetric.

60. (continued)





b) The brands in the sample had a mean calorie content of 131.82 calories, and a standard deviation of 25.23 calories. Since the conditions for inference have been satisfied, use a one-sample *t*-interval, with 11 – 1 = 10 degrees of freedom, at 95% confidence.

$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 131.82 \pm t_{10}^{*}\left(\frac{25.23}{\sqrt{11}}\right) \approx (114.87, 148.77)$$

c) We are 95% confident that the mean calorie content in a serving of vanilla yogurt is between 114.87 and 148.77 calories. The reported average of 120 calories is plausible. The 95% confidence interval contains 120 calories.

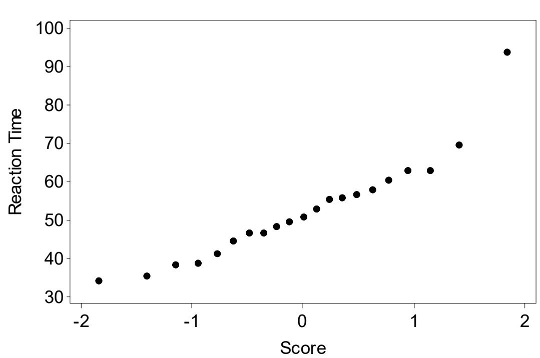
61. Maze.

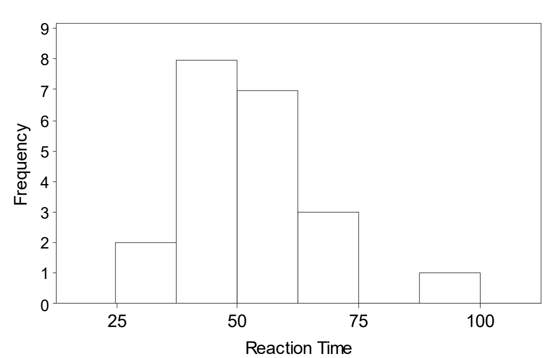
a) The rats in the sample finished the maze with a mean time of 52.21 seconds and a standard deviation in times of 13.5646 seconds. Since the conditions for inference are satisfied, we can construct a one-sample *t*-interval, with 21 – 1 = 20 degrees of freedom, at 95% confidence.

$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 52.21 \pm t_{20}^{*}\left(\frac{13.5646}{\sqrt{21}}\right) \approx (46.03, 58.38)$$

We are 95% confident that the true mean maze completion time is contained in the interval 46.03 seconds to 58.38 seconds.

b) **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.





61. (continued)

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 construct a one-sample t-interval, with 20 – 1 = 19 degrees of freedom, at 95% confidence.

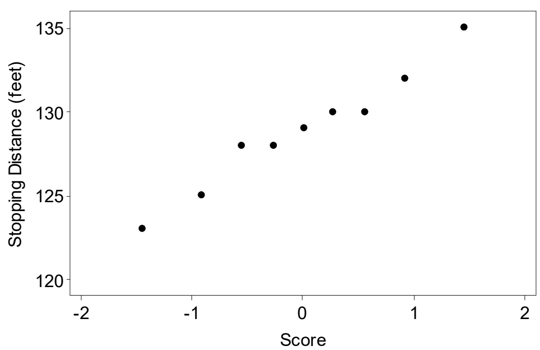
$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 50.13 \pm t_{19}^{*}\left(\frac{9.90}{\sqrt{20}}\right) \approx (45.49, 54.77)$$

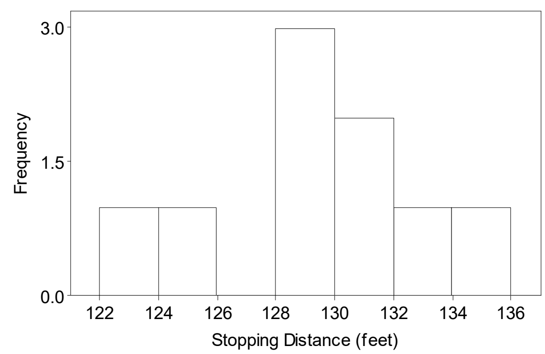
We are 95% confident that the true mean maze completion time is contained in the interval 45.49 seconds to 54.77 seconds.

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.

62. Stopping distance 60.

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 (without the outlier of 102 feet) had a mean of 128.889 feet, and a standard deviation of 3.55121 feet. Since the conditions for inference are satisfied, we can construct a one-sample t-interval, with 9 – 1 = 8 degrees of freedom, at 95% confidence.

$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 128.889 \pm t_{8}^{*}\left(\frac{3.55121}{\sqrt{9}}\right) \approx (126.2, 131.6)$$

We are 95% confident that the true mean stopping distance is contained in the interval 126.2 feet to 131.6 feet. This interval does not contain 125 feet, so the company cannot be confident that the new tread pattern meets their standard. The new tread pattern should not be adopted.

63 Golf Drives 2015.

a) $\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 288.69 \pm t_{198}^{*}\left(\frac{9.28}{\sqrt{199}}\right) \approx (287.39, 289.99)$

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 2016 data to represent the population of all professional golfers, past, present, and future.

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

64. Wind power.

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

b) 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 construct a one-sample *t*-interval, with 1114 – 1 = 1113 degrees of freedom, at 95% confidence.

$$\bar{y} \pm t_{n-1}^{*}\left(\frac{s}{\sqrt{n}}\right) = 8.019 \pm t_{1113}^{*}\left(\frac{3.813}{\sqrt{1114}}\right) \approx (7.7948, 8.2432)$$

We are 95% confident that the true mean wind speed is contained in the interval 7.79 mph and 8.24 mph.

Since 8 miles per hour is contained in the interval, 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.