

3.3: Measures of Placement

3.3.1 Z-Scores

A **z-score** is the number of standard deviations an observation x is above or below the mean. Z-scores are used to compare placement of a value compared to the mean.

If x is an observation from a sample then the standardized value of x is the z-score where $z = \frac{x - \bar{x}}{s}$

If x is an observation from a population then the standardized value of x is the z-score where $z = \frac{x - \mu}{\sigma}$

If the z-score is negative, x is less than the mean. If the z-score is positive, x is greater than the mean.

There are no shortcuts on the calculator in Excel for z-score, but you can find s and \bar{x} then simply subtract and divide.

The number of standard deviations that a data value is from the mean is frequently used when comparing position of values. If a z-score is zero, then the data value is the same as the mean. If the z-score is one, then the data value x is one standard deviation above the mean. If the z-score is -3.5 , then the data value is three and a half standard deviations below the mean. The shaded area in Figure 3-20 represents one standard deviation from the mean.

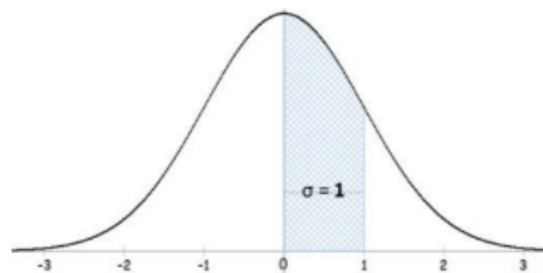


Figure 3-20

For a random sample, the mean time to make a cappuccino is 2.8 minutes with a standard deviation of 0.86 minutes. Find the z-score for someone that makes their cappuccino in 4.95 minutes.

Solution

$z = \frac{x - \bar{x}}{s} = \frac{4.95 - 2.8}{0.86} = 2.5$. Their time is 2.5 standard deviations above average.

On a math test, a student scored 45. The class average was 50 with a standard deviation of 3. The same student scored an 80 on a history test, and the class average was 85 with a standard deviation of 2.5. Which exam did the student perform better on compared with the rest of the class?

Solution

$$z_m = \frac{45 - 50}{3} = -1.67 \quad z_h = \frac{80 - 85}{2.5} = -2$$

Test scores are “better” when they are larger, so whichever has the largest z-score did better. The student did better on the math test than the history test, compared to the rest of the class.

Be careful with the word “better,” depending on the context, better may be smaller rather than larger. For example, golf scores, time running a race, and cholesterol levels would be better if they were smaller values.

The length of a human pregnancy has a mean of 272 days. A pregnancy lasting 281 days or more has a z-score of one. How many standard deviations above the mean is a pregnancy lasting 281 days or more?

Solution

One, since by definition the z-score is the number of standard deviations from the mean.

The length of a human pregnancy has a mean of 272 days. A pregnancy lasting 281 days or more has a z-score of one. What is the standard deviation of human pregnancy length?

Solution

We know the z-score = 1 and mean = 272. Replace these two numbers in the z-score formula then solve for the standard deviation.

$$1 = \frac{281-272}{\sigma} \Rightarrow 1 = \frac{9}{\sigma} \Rightarrow \sigma = 9$$

3.3.2 Percentiles

Along with the center and variability, another useful numerical measure is the ranking of a number. A **percentile** is a measure of ranking. It represents a location measurement of a data value to the rest of the values. Many standardized tests give the results as a percentile. Doctors use percentiles graphs to show height and weight standards.

Interpreting Percentiles

The pth percentile is the value that separates the bottom p% from the upper (100 – p)% of the ordered (smallest to largest) data. For example, the 75th percentile is the value that separates the bottom 75% from the upper 25% of the data. There are several methods used to find percentiles. You may get different percentile values depending on which software or calculator you use. For example, Excel has two methods, both of which are not the same method as the TI calculators.

What does a score of the 90th percentile represent?

Solution

This means that 90% of the scores were at or below this score. (A person did the same as or better than 90% of the test takers.)

What does a score of the 70th percentile represent?

Solution

This means that 70% of the scores were at or below this score.

Percentile versus Score

If the test was out of 100 points and you scored at the 80th percentile, what was your score on the test? You do not know! All you know is that you scored the same as or better than 80% of the people who took the test. If all the scores were low, you could have still failed the test. On the other hand, if many of the scores were high you could have gotten a 95% or so.

Note there is more than one method to find percentiles. This rounding rule in Excel is not the same as used on your TI calculators.

Finding a Percentile:

Step 1: Arrange the data in order from lowest to highest.

Step 2: Substitute into the formula $i = \frac{(n+1) \cdot p}{100}$ where n = sample size and p = percentile.

Step 3A: If i is a whole number, count out i places from the lowest number to find the percentile. For example, if you get i = 3, then the 3rd value is the percentile.

Step 3B: If i is not a whole number, then take the weighted average between the ith and ith + 1 data value as the percentile. For example, if i = 3.25, this would be 25% of the distance between the 3rd and the 4th data values as the percentile. Percentile = ith data value + (ith + 1 data value – ith data value)*(0.##) where ## is the remainder percent.

Compute the 10th percentile of the random sample of 13 ages: 15, 18, 22, 25, 26, 31, 33, 35, 38, 46, 51, 53, and 95.

Solution

Data is already ordered, so next find $i = \frac{(n+1) \cdot p}{100} = \frac{14 \cdot 10}{100} = 1.4$. Since i is not a whole number use Step 3B, take the weighted average of 40% of the way between the 1st and 2nd values. This would be $15 + (18 - 15) \cdot 0.4 = 16.2$ and this is your 10th percentile, $P_{10} = 16.2$.

In Excel use =PERCENTILE.EXC(array, k) where array is the cell reference to where the data is located and k is the percentile as a decimal between 0 and 1. Note you do not have to sort the data prior to typing it in to Excel.

For this example, if you type in the data into column A, then use the formula =PERCENTILE.EXC(A1:A13, 0.1) = 16.2.

	A	B	C	D	E
1	Age				
2	15		=PERCENTILE.EXC(A2:A14,0.1)		
3	18				
4	22				
5	25				
6	26				
7	31				
8	33				
9	35				
10	38				
11	46				
12	51				
13	53				
14	95				

3.3.3 Quartiles

There are special percentiles called quartiles. Quartiles are numbers that divide the data into fourths. One fourth (or a quarter) of the data falls between consecutive quartiles. There are three quartiles Q_1 , Q_2 , and Q_3 that subsequently divide the ordered data into the 4 pieces of approximately equal size, or 25% each. Thus, 25% of the values are less than Q_1 , 25% of the data values are between Q_1 and Q_2 , 25% of the data values are between Q_2 and Q_3 , and 25% are of the data values are greater than Q_3 .

Use the dollar as an example. If we make change for a dollar, we would get four quarters to make one dollar. Hence, quarter for quartiles.



To find the quartiles use the same rules as percentiles where we:

1. Arrange the observations from smallest to largest and use the previous percentile rule.
2. Then find all three quartiles.
 - $Q_1 = \text{first quartile} = 25^{\text{th}} \text{ percentile}$
 - $Q_2 = \text{second quartile} = \text{median} = 50^{\text{th}} \text{ percentile}$
 - $Q_3 = \text{third quartile} = 75^{\text{th}} \text{ percentile}$

Compute all three quartiles for the random sample of 13 ages: 15, 18, 22, 25, 26, 31, 33, 35, 38, 46, 51, 53, and 95.

Solution

For the first quartile $i = \frac{(n+1) \cdot p}{100} = \frac{14 \cdot 25}{100} = 3.5$. Since i is not a whole number take the weighted average of half way between the 3rd and 4th data values $22 + (25 - 22) \cdot 0.5 = 23.5$, so $Q_1 = 23.5$.

In Excel you could use the percentile formula, but there is also a quartile formula: =QUARTILE.EXC(array, quartile), where array is the cell reference to the data and quartile is either 1, 2 or 3 for the 3 possible quartiles.

In this example we would have =QUARTILE.EXC(A1:A13, 1) = 23.5.

To find the second quartile: $i = \frac{(n+1) \cdot p}{100} = \frac{14 \cdot 50}{100} = 7$. Since i is a whole number use the 7th value for Q_2 , so $Q_2 = 33$.

Or use the Excel formula =QUARTILE.EXC(A1:A13, 2) = 33.

For the third quartile, $i = \frac{(n+1) \cdot p}{100} = \frac{14 \cdot 75}{100} = 10.5$. Since i is not a whole number use the weighted average of half way between the 10th and 11th values $46 + (51 - 46) \cdot 0.5 = 48.5$.

Or use the Excel formula =QUARTILE.EXC(A1:A13, 3) = 48.5, so $Q_3 = 48.5$.

The high school graduating class of 2016 in Oregon had the following ACT quartile scores. Interpret what the number 26 under the Composite column represents.

Quartile	English	Mathematics	Reading	Science	Composite
Q ₃ (75th Percentile)	25	26	27	25	26
Q ₂ (50th Percentile)	21	21	22	22	21
Q ₁ (25th Percentile)	16	17	17	18	17

https://www.act.org/content/dam/act/unsecured/documents/P_38_389999_S_S_N00_ACT-GCPR_Oregon.pdf

Solution

From the report we can see that the third quartile for composite score is 26, this means that 75% of Oregon students that took the ACT exam scored 26 or below.

Other Types of Percentiles

Quintiles break a data set up into five equal pieces. We will not be using these, but be aware that percentiles come in different forms.

Deciles break a data set up into ten equal pieces and are found using the percentile rule. For example, the 6th decile = $D_6 = 60^{\text{th}}$ percentile.

Use the dollar as an example. If we make change for a dollar, we would get ten dimes to make one dollar. Hence, a dime might help you remember deciles.



Earlier in Example 2-14, we made an ogive using Excel with the following sample of 35 ages. Use the ogive to find the age for the 8th decile.

46 47 49 25 46 22 42 24 46 40 39 27 25 30 33 27 46 21 29 20 26 25 25 26 35 49 33 26 32 31 39 30 39 29 26

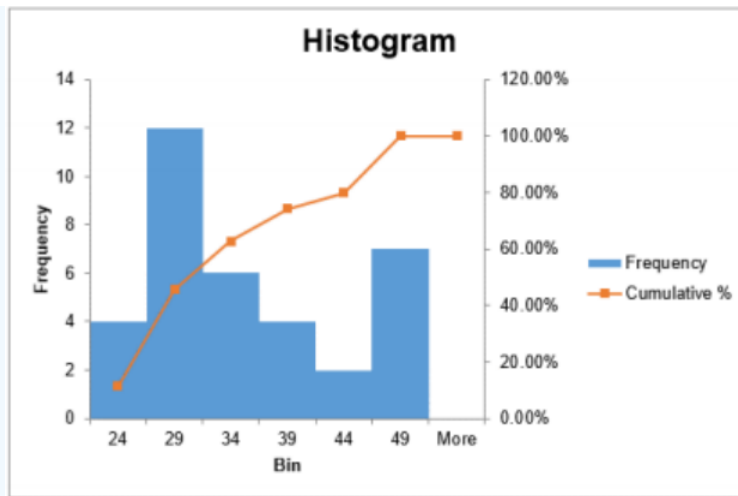


Figure 3-21

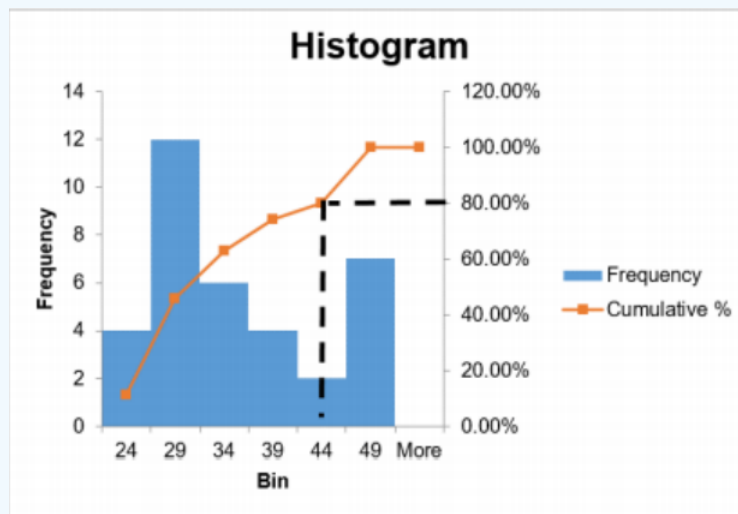


Figure 3-22

Solution

The cumulative % represents the cumulative relative frequencies which are equivalent to the percentiles for each class. The red line is the ogive and the percentiles correspond to the vertical axis on the right side.

If we wanted to know what age the 80th percentile was in the sample, we could use the ogive to get an approximate value. Starting on the right at 80% make a horizontal line until you hit the red cumulative % line, and then make a vertical line from there down to the axis to get the approximate age. See Figure 3-22. In this case, the 80th percentile = 8th decile would be approximately 44.

3.3.4 Five Number Summary & Outliers

If you record the quartiles together with the minimum and maximum values from a data set, you have five numbers. These five numbers are known as the five-number summary consisting of the minimum, the first quartile (Q_1), the median (Q_2), the third quartile (Q_3), and the maximum (in that order).

The **interquartile range**, IQR, is the difference between the first and third quartiles, Q_1 and Q_3 . Half of the data (50%) falls in the interquartile range. If the IQR is “large,” the data is spread out and if the IQR is “small,” the data is closer together.

$$\text{The interquartile range (IQR)} = Q_3 - Q_1$$

Not only does the IQR give a range of the middle 50% of the data, but is also used to determine outliers in a sample.

To find these outliers we first find what are called a lower and upper limit sometimes called fences.

The lower limit, or inner fence, is $Q_1 - (1.5 \cdot \text{IQR})$. Any values that are less than the lower limit is considered an outlier.

Similarly, the upper limit, or outer fence, is $Q_3 + (1.5 \cdot \text{IQR})$. Any values that are more than the upper limit are considered outliers.

If all the numbers in the sample fall between the lower and upper limit, including the endpoints, then there are no outliers in the sample. Any values outside these limits would be considered outliers.

3.3.5 Modified Box-and-Whisker Plot

A boxplot (or box-and-whisker plot) is a graphical display of the five-number summary. A boxplot can be drawn vertically or horizontally. The modified boxplot shows outliers, whereas a regular boxplot does not show outliers. The basic format of the plot is a box drawn from Q_1 to Q_3 , a vertical line drawn inside the box for the median, and horizontal lines (called whiskers) extending out of the middle of each end of the box to the minimum and maximum. The box should not touch the number line. The modified boxplot extends the left line to the smallest value greater than the lower fence, and extends the right line to the largest value less than the upper fence. Dots, circles or asterisks represent any outlier. We will make modified boxplots for this course. Like always, label the tick marks on the number line and give the graph a title.

A **boxplot** is a graph of the 5-number summary, see Figure 3-23.

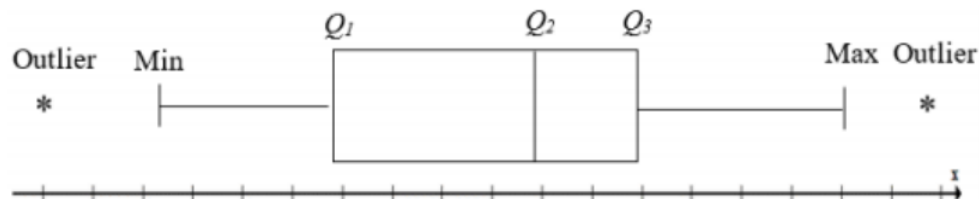


Figure 3-23

It is important to note that when you are making the boxplot the limits for finding outliers are not graphed in the plot, they were only used to find the outliers. The whiskers would go to the next largest (or smallest) value in the data set after you removed the outlier(s).

If the sample has a symmetrical distribution, then the boxplot will be visibly symmetrical. If the data distribution has a left skew or a right skew, the line on that side of the boxplot will be visibly long in the direction of skewness. If the four quartiles are all about the same distance apart, then the data are likely a near uniform distribution. If a boxplot is symmetrical, and both outside lines are noticeably longer than the Q_1 to median and median to Q_3 distance, the distribution is then probably bell-shaped.

Make a modified box-and-whisker plot for the random sample of 13 ages: 15, 18, 22, 25, 26, 31, 33, 35, 38, 46, 51, 53, and 95.

Solution

Use Excel to compute the three quartiles as:

$$Q_1 = \text{QUARTILE.EXC}(A1:A13, 1) = 23.5$$

$$Q_2 = \text{QUARTILE.EXC}(A1:A13, 2) = 33$$

$$Q_3 = \text{QUARTILE.EXC}(A1:A13, 3) = 48.5.$$

The 5-number summary values are 15, 23.5, 33, 48.5 and 95. Each of these numbers will need to be incorporated into the box-and-whisker plot, and any outliers to graph the modified box-and-whisker plot.

To find the outliers, first find the IQR, and then find the lower and upper limits. The $\text{IQR} = Q_3 - Q_1 = 48.5 - 23.5 = 25$. The lower limit is $Q_1 - (1.5 \cdot \text{IQR}) = 23.5 - 1.5(25) = -14$. The upper limit is $Q_3 + (1.5 \cdot \text{IQR}) = 48.5 + 1.5(25) = 86$.

Any value in our data set that is not between the lower and upper limits $[-14, 86]$ is an outlier. By observation, we have one number that is outside the range so the outlier is 95.

The whiskers would be drawn to the next largest (or smallest) value in the data set after you removed the outlier(s). For this example, the next largest value in the data set is 53.

Now put that all together to get the following graph in Figure 3-24

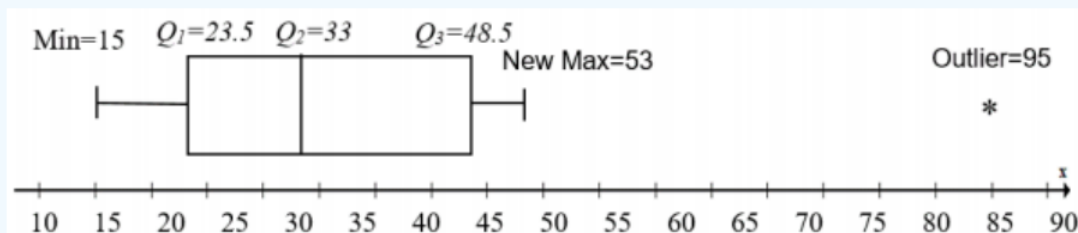
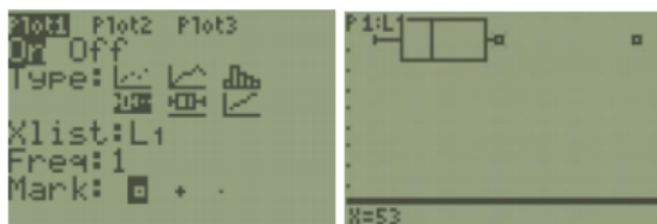


Figure 3-24

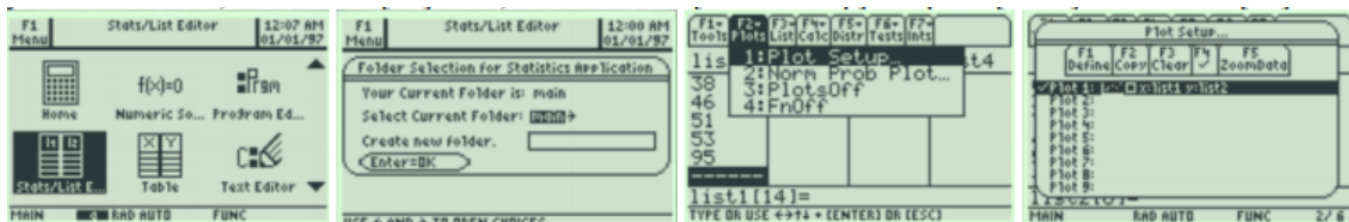
The TI-calculator and newer versions of Excel will make a modified boxplot. Note that the quartile rules used in the TI calculators are slightly different than in Excel and what is presented in this content. They do not use a weighted mean between values, just half way between values.

TI-84: First, enter your data in to list 1. Next, press 2^{nd} > STAT PLOT, then choose the first plot. Note that your calculator may say Plot1...Off or show a different type of graph then the screenshot. Using your arrow keys, turn the plot on. Choose the modified boxplot which is the first of the two boxplot options with the small dots to the right of the whiskers representing outliers. Make sure your Xlist: is on L_1 , keep frequency as a one, and any mark will work, but the square shows up best.

Here is screen shot from the calculator for the last example. You can use Trace on the boxplot from the TI-84 calculator below to see where each quartile, whisker and outlier are.



TI-89: Enter the data into the Stat/List editor under list 1. Press [APP] then scroll down to Stat/List Editor; on the older style TI-89 calculators, go into the Flash/App menu, and then scroll down the list. Make sure the cursor is in the list, not on the list name, and type the desired values pressing [ENTER] after each one. To clear a previously stored list of data values, arrow up to the list name you want to clear, press [CLEAR], and then press enter. After you enter the data, select Press [F2] Plots, scroll down to [1: Plot Setup] and press [Enter]. Then select [F1] Define.



Use your arrow keys to select Mod Box Plot for Type, and then scroll down to the x-variable box. Press 2^{nd} [VarLink] this key is above the + sign. Then arrow down until you find your List1 name under the Main file folder.

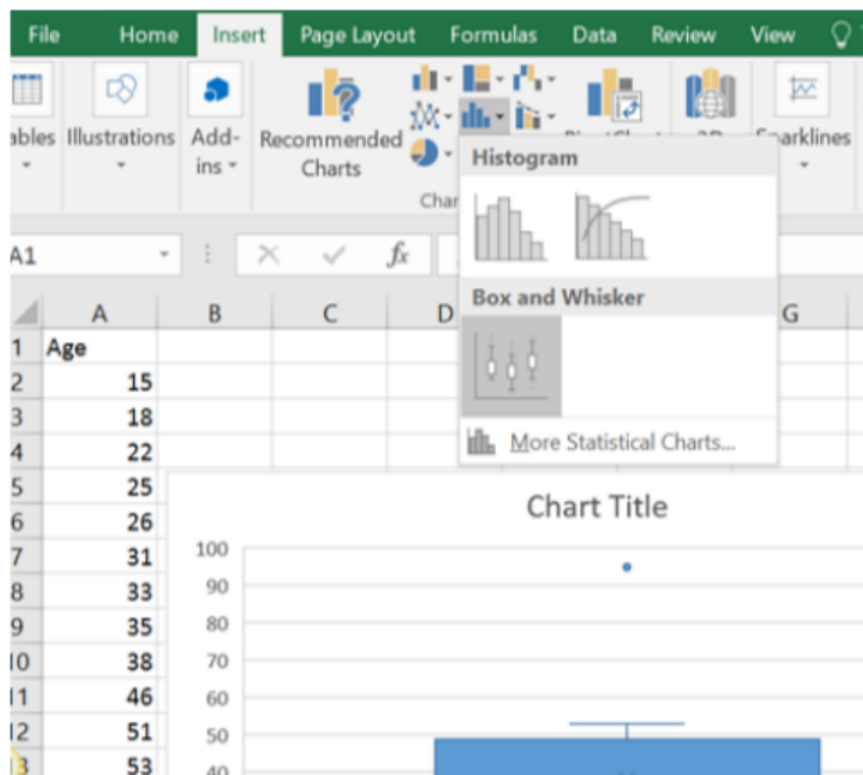
Then press [Enter] and this will bring the name List1 back to the menu. You will now see that Plot1 has a small picture of a boxplot. To view the boxplot, press [F5] Zoom Data.



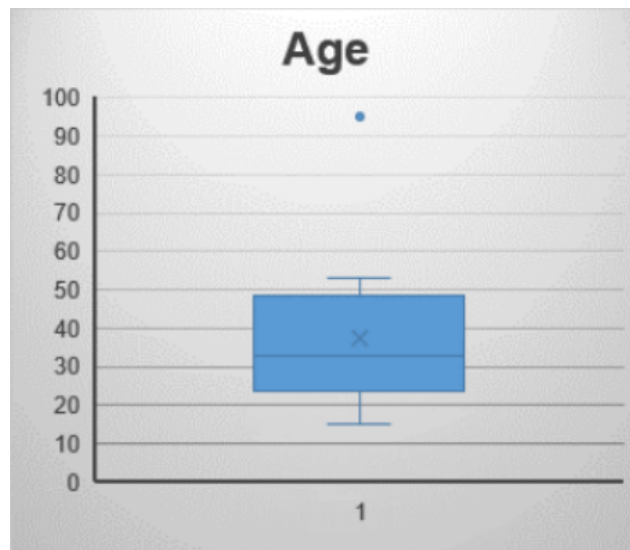
Select [F3:Trace] to see the five-number summary and any outliers. Use the left and right arrow keys to move to the other values.

Excel: Note this example is on a PC running Excel 2019. Older versions of Excel may not have a boxplot option. First, type your sample data into column A in any order. Highlight the data, and then select the Insert tab. Under the graphing options, the picture shaped liked a histogram called statistical charts, select Box and Whisker.

You can change the formatting options and add the chart title as needed.



Below is the finished Excel boxplot. Note that Excel does a vertical boxplot rather than the traditional horizontal number line.



Excel marks an \times just above the median where the mean would fall. Usually one would not include the mean on a boxplot. Remember that when the mean is greater than the median the distribution is usually skewed to the right.

When a boxplot has outliers only on one side, then we can also say the distribution is skewed in the direction of the outlier, which also indicates that these ages are skewed to the right.

Side by side boxplots are great at comparing quartiles and distribution shapes for several samples using the same units and scale.

There are four franchises in different parts of town. Compare the weekly sales over a year for each of the four franchises. Compare the boxplots shown in Figure 3-25.



Figure 3-25

Solution

We can see that Store 2 in Figure 3-25 has the highest sales since the median for this store is higher than the third quartile for all the other stores. Store 2 also has sales that are more consistent from week to week with the smaller range and has a symmetric distribution. The lowest performing store, Store 1, has the lowest median sales and is skewed to the right. Both Stores 3 and 4 have moderate sales and are skewed left.

3.3.6 Empirical Rule

Before looking at the process for computing probabilities, it is somewhat useful to look at the Empirical Rule which gives the approximate proportion of data points under a bell-shaped curve between two points. The Empirical Rule is just an approximation,

more precise methods for finding these proportions will be demonstrated in later sections.

The Empirical Rule should only be used with bell-shaped data.



The Empirical Rule (also called the 68-95-99.7 Rule)

In a bell-shaped distribution with mean μ and standard deviation σ ,

- Approximately 68% of the observations fall within 1 standard deviation (σ) of the mean μ .
- Approximately 95% of the observations fall within 2 standard deviations (2σ) of the mean μ .
- Approximately 99.7% of the observations fall within 3 standard deviations (3σ) of the mean μ .

Note that we are using notation for the population mean μ and the population standard deviation σ , but the rule would also work using the sample mean and sample standard deviation.

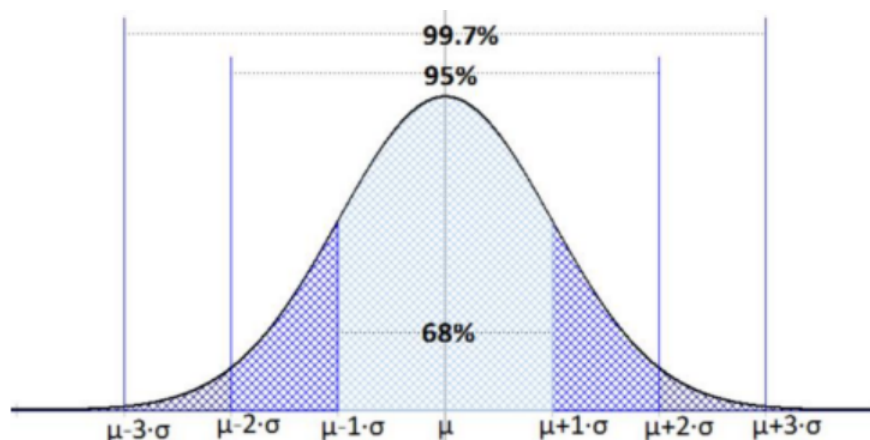


Figure 3-26

In 2009 the average SAT mathematics score was 501, with a standard deviation of 116. Assume that SAT scores are bell-shaped.

- Approximately what proportion of students scored between 269 and 733 on the 2009 SAT mathematics test?
- Approximately what proportion of students scored between 385 and 617 on the 2009 SAT mathematics test?
- Approximately what proportion of students scored at least 617 on the 2009 SAT mathematics test?

Solution

a) The key word is bell shaped so we can use the Empirical Rule. Start by finding the z-scores $z = \frac{x-\mu}{\sigma}$ for both endpoints given in the question.

A z-score by definition is the number of standard deviations a data value is from the mean.

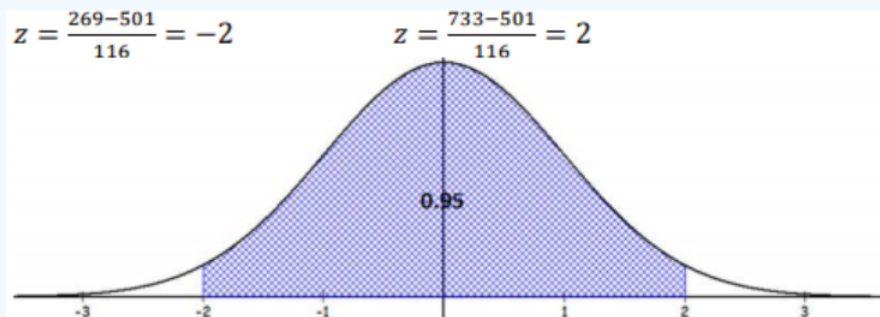


Figure 3-27

The two z-scores show that the test scores of 269 and 733 are two standard deviations from the mean. Using the second bulleted item in the Empirical Rule the answer would be approximately 95% of the math SAT scores will fall between 269 and 733.

b) Take the z-scores of the endpoints to get: $z = \frac{385-501}{116} = -1$, $z = \frac{617-501}{116} = 1$

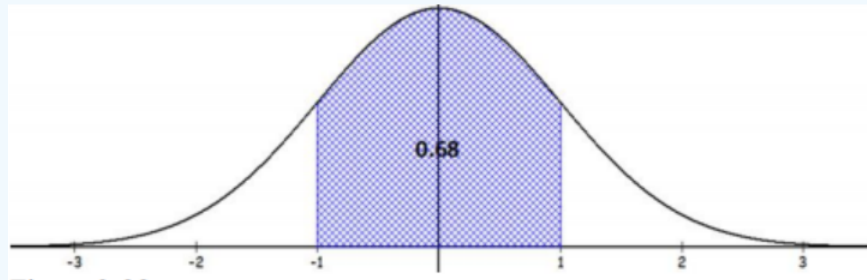


Figure 3-28

The two z-scores show that the test scores of 385 and 617 are one standard deviation from the mean. Using the first bulleted item in the Empirical Rule the answer would be approximately 68% of the math SAT scores will fall between 385 and 617.

c): Start by taking the z-score of 617, we get $z = \frac{617-501}{116} = 1$.

Since a bell-shaped curve is symmetric and we can assume that 100% of the population is represented then we can subtract the middle from the whole to get $100\% - 68\% = 32\%$. If we divide this outside area by two, $\frac{32}{2} = 16\%$, we would expect 16% in each tail area. See Figure 3-29.

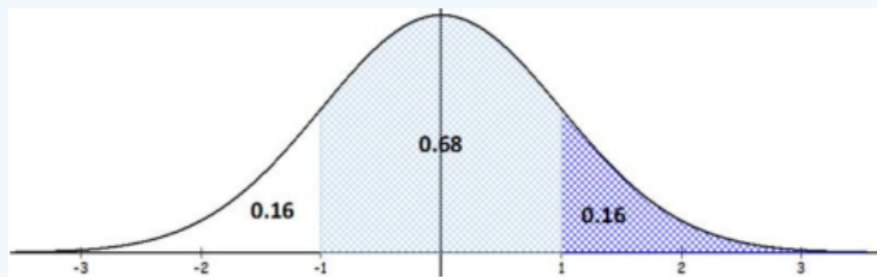


Figure 3-29

The answer would be approximately 16% of students scored at least 617 on the 2009 SAT mathematics test.

If you were to get a z-score that is not -3 , -2 , -1 , 1 , 2 or 3 then you would not be able to apply the Empirical Rule. We also need to ensure that our population has a bell-shaped curve before using the Empirical Rule.

3.3.7 Chebyshev's Theorem

One way of estimating the proportion of values from any data set within a certain number of standard deviations is Chebyshev's Theorem. Pafnuty Chebyshev (Чебышёв) was a Russian mathematician who proved several important theorems. One that we will use for this chapter is called Chebyshev's Inequality. The Empirical Rule only works for bell-shaped distributions. However, we can use Chebyshev's Inequality for any distribution shape.



[Chebyshev](#)

Chebyshev's Inequality: The proportion (percent or fraction) of values from a data set that will fall within k standard deviations of the mean will be at least $\left(\left(1 - \frac{1}{(z)^2} \right) \cdot 100 \right) \%$, where z is a real number that has an absolute value greater than 1 (z is not necessarily an integer).

The average number of acres for farms in a certain country is 443 with a standard deviation of 42 acres. At least what percent of farms will have between 338 and 548 acres?

Solution

The question gives no indication for the distribution shape for number of farm acres so we will use Chebyshev's Inequality instead of the Empirical Rule. The easiest way to start is to find the z -score of the lower and upper bounds given in the question where $\mu = 443$ and $\sigma = 42$.

$$z = \frac{338-443}{42} = -2.5 \text{ and } z = \frac{548-443}{42} = 2.5$$

Use either z -score, but it is easier to use the positive value, and substitute into the formula:

$$\left(\left(1 - \frac{1}{(z)^2} \right) \cdot 100 \right) \% = \left(\left(1 - \frac{1}{(2.5)^2} \right) \cdot 100 \right) \% = 84\%$$

At least 84% of the farms will have between 338 and 548 acres.

The average quiz score for a statistics course is 15.2 with a standard deviation of 3.15. What are the quiz scores that would have at least 75% of the scores between?

Solution

This question gives a percent and we need to work backward. We will use Chebyshev's Inequality since there is no mention of a bell-shaped distribution. Use algebra to solve for z in the following formula

$$\left(\left(1 - \frac{1}{(z)^2} \right) \cdot 100 \right) \% = 75\%.$$

Start by dividing both sides by 100% to get rid of the %. We then have $1 - \frac{1}{(z)^2} = 0.75$.

Next, add $\frac{1}{(z)^2}$ to both sides of the equation and subtract 0.75 from both sides of the equation to get $0.25 = \frac{1}{(z)^2}$.

Multiply both sides by z^2 , divide both sides by 0.25 and simplify to get $(z)^2 = \frac{1}{0.25} \Rightarrow z^2 = 4$.

Take the square root of both sides $\sqrt{z^2} = \sqrt{4} \Rightarrow z = \pm 2$.

This means that according to Chebyshev's Theorem at least 75% of the data will fall within two standard deviations from the mean. Next, we need to find what quiz scores are two standard deviations from the mean.

Find the mean ± 2 standard deviations, by computing:

$$\mu - 2 \cdot \sigma = 15.2 - 2 \cdot 3.15 = 8.9$$

$$\mu + 2 \cdot \sigma = 15.2 + 2 \cdot 3.15 = 21.5$$

At least 75% of the students scored between 8.9 and 21.5 on the quiz.

The general formulas for finding the endpoints are: Lower endpoint: $a = \mu - z \cdot \sigma$

Upper endpoint: $b = \mu + z \cdot \sigma$

If the distribution of quiz scores were bell shaped we would have a larger percent (95%) between 8.9 and 21.5. Chebyshev's Inequality assumes the distribution is skewed and says "at least" which would then also be correct if more students fell within the range.

Since Chebyshev's Inequality works for any shape of a distribution we should only use it between two values, not strictly below or above a point. If we were interested in below a certain point, we would not know if we had the fat or skinny tail of a skewed distribution. The following picture shows a positively skewed distribution with 1.5 standard deviations from the mean shaded in red. Chebyshev's Inequality would guarantee at least

$\left(\left(1 - \frac{1}{(1.5)^2} \right) \cdot 100 \right) \% = 56\%$ of the data in the shaded area would fall within 1.5 standard deviations from the mean.

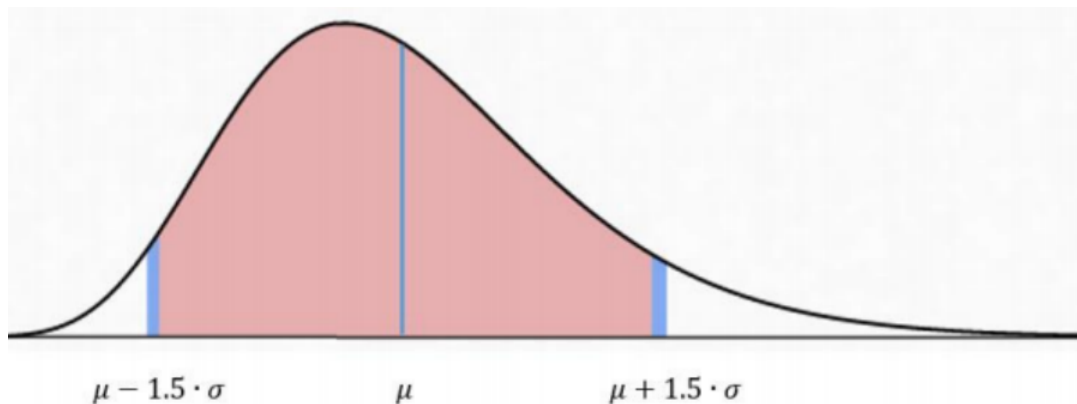


Figure 3-30

Summary:

Use the mode for nominal, ordinal, interval, and ratio data, since the mode is just the data value that occurs most often. You are just counting the data values. The median can be found on ordinal, interval, and ratio data, since you need to put the data in order. As long as there is order to the data you can find the median. The mean can be found on interval and ratio data, since you must have numbers to add together. The mean is pulled in the direction of outliers. By comparing the mean to the median, you can decide if a distribution is symmetric or skewed. The range, variance, standard deviation and coefficient of variation are used to measure how spread out a data set is from the middle. When comparing two data sets with different units or scale use the coefficient of variation. Z-scores tell you how many standard deviations a data point is away from the mean. Quartiles are special percentiles that are used to find the interquartile range, identify outliers, and make a box-and-whisker plot. Use the Empirical Rule when finding the proportion of a sample or population that fall within 1, 2 or 3 standard deviations on a bell-shaped curve. If the distribution is bell shaped then the Empirical Rule states that approximately 68% of the data will fall within one standard deviation, 95% within two standard deviations and 99.7% within three standard deviations. If you do not know the distribution shape, then use Chebyshev's Inequality to find the minimum proportion within $|z| > 1$ standard deviations from the mean.

This page titled [3.3: Measures of Placement](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Rachel Webb](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.