

Quantitative Methods for Business and Finance

Dr. Brian Travers



Contents

1	Fractions, Decimals and Percents	5
1.1	Fractions	7
1.1.1	Reducing Fractions	8
1.1.2	Adding and Subtracting Fractions	8
1.1.3	Multiplying and Dividing Fractions	10
1.1.4	Exercises	12
1.1.5	Solutions	13
1.2	Decimals with	14
1.2.1	Decimals and Fractions	14
1.2.2	Decimals and Percents	15
1.2.3	Exercises	16
1.2.4	Solutions	17
1.3	Representing Change	18
1.3.1	Exercises	20
1.3.2	Solutions	21
2	Functions	23
2.1	What Is A Function?	25
2.1.1	Exercises	26
2.2	Notation	28
2.2.1	Set Builder Notation	28
2.2.2	$y = f(x)$ Notation	28
2.2.3	Evaluating the Function $y = f(x)$	29
2.2.4	Exercises	30
2.2.5	Solutions	31
2.3	Functions from Tables	32
2.3.1	Exercises	33
2.3.2	Solutions	34
2.4	Graphs of Functions	35
2.4.1	Exercises	37
3	Linear Functions	41
3.1	Linear Functions	43
3.1.1	Slope	43
3.1.2	Special Cases for Slope	44

3.1.3	The Slope-Intercept Form of a Linear Function	45
3.1.4	Intercepts	46
3.1.5	Parallel and Perpendicular Lines	47
3.1.6	Standard Form $ax + by = c$	49
3.1.7	Is the point on the line?	50
3.1.8	Exercises	51
3.1.9	Solutions	52
3.2	Graphing Linear Functions	53
3.2.1	Plotting Using a t -chart	53
3.2.2	Plotting Using Slope-Intercept Form	55
3.2.3	Using the TI-Series Calculator	57
3.2.4	Exercises	58
3.2.5	Solutions	59
3.3	Finding the Intersection Point of Linear Functions	62
3.3.1	How to Find the Intersection	62
3.3.2	Do We Have to Have an Intersection?	64
3.3.3	Exercises	65
3.3.4	Solutions	66
3.4	Applications of Linear Functions	68
3.4.1	Linear Depreciation	68
3.4.2	Cost and Revenue	68
3.4.3	Straight Line Depreciation	69
3.4.4	Supply and Demand	70
3.4.5	Exercises	72
3.4.6	Solutions	73
3.5	Graphing Inequalities	75
3.5.1	Exercises	82
3.5.2	Solutions	83
3.6	Justification of Slopes of Perpendicular Lines	85
4	Applications of Linear Functions	87
4.1	Linear Programming: The Set-up	89
4.1.1	What Is Linear Programming?	89
4.1.2	What Are We Looking For?	89
4.1.3	Exercises	93
4.1.4	Solutions	95
4.2	Linear Programming: Maximization and Minimization	98
4.2.1	Optimizing Given Lines and Points	98
4.2.2	Why the Vertices Are the Only Points We Care About	100
4.2.3	More Involved Examples	101
4.2.4	Word Problems	104
4.2.5	Exercises	109
4.2.6	Solutions	110
4.3	Linear Programming on the TI-84	114

5	Exponential and Logarithmic Functions	119
5.1	Exponential Functions	121
5.1.1	What Makes a Function Exponential?	121
5.1.2	Types of Exponential Functions	122
5.1.3	The Graph of Exponential Functions	123
5.1.4	Examples	125
5.1.5	Exercises	128
5.1.6	Solutions	129
5.2	Logarithmic Functions	131
5.2.1	What is a Logarithm?	131
5.2.2	Properties of Logarithms	131
5.2.3	Graphs of Logarithmic Functions	133
5.2.4	Using Logarithms	134
5.2.5	Exercises	138
5.2.6	Solutions	139
5.3	Interest	141
5.3.1	Simple Interest	141
5.3.2	Compound Interest	142
5.3.3	Exercises	145
5.3.4	Solutions	146
6	Introduction to Statistics	147
6.1	Calculating Statistics	149
6.2	Standard Deviation	150
6.3	The 5-Number Summary	153
6.4	Statistics Using the TI-Series Calculator	154
6.5	Percentiles	155
6.6	Misuses of Statistics	157
6.7	Exercises	160
6.8	Solutions	162
7	Visual Representations of Data	163
7.1	Pie Charts	165
7.1.1	Drawbacks to Pie Charts	167
7.2	Bar Graphs	168
7.2.1	Creating Bar Graphs	168
7.2.2	Usage of Bar Graphs	169
7.3	Distributions	170
7.4	Dot Plots	171
7.4.1	Creating Dot Plots	171
7.4.2	The Pros and Cons	172
7.5	Stem-and-Leaf Plots	172
7.5.1	Creating Stem-and-Leaf Plots	172
7.6	Histograms	174
7.6.1	Creating Histograms	175
7.6.2	Histograms on the TI-Series Calculator	177

7.7	Box Plots	178
7.7.1	Box-and-Whisker Plots on the TI-Series Calculator	180
7.7.2	Comparisons with Box-and-Whisker Plots	182
7.7.3	Exercises	183
7.7.4	Solutions	185
8	Scatter plots and Regression	191
8.1	Scatter plots	193
8.1.1	Points of Clarification for Correlation Coefficients	202
8.2	Linear Regression	203
8.2.1	Important things to note about regression lines	205
8.2.2	What good is having this regression line?	206
8.2.3	Residuals	207
8.3	Lurking Variables	216
8.4	Exercises	218
8.5	Solutions	220
9	Probability	223
9.1	The Basic Laws of Probability	225
9.1.1	Definitions and Examples	225
9.1.2	Subsets	227
9.1.3	Back to Discrete Probability	229
9.1.4	The Basic Rules	229
9.1.5	Conditional Probability	233
9.1.6	Exercises	235
9.1.7	Solutions	236
9.2	Tree Diagrams and the Multiplication Rule	238
9.2.1	Exercises	241
9.2.2	Solutions	242
9.3	Binomial Probabilities	244
9.3.1	Rules for Using Binomial Probabilities	244
9.3.2	What Happens As n Gets Larger	246
9.3.3	Central Tendencies of Binomial Distributions	248
9.3.4	Cumulative Binomial Probabilities	248
9.3.5	Binomial Probabilities on the TI-Series Calculator	249
9.3.6	Exercises	253
9.3.7	Solutions	254
10	Normal Distributions	255
10.1	The Normal Distribution Curve	257
10.2	The Empirical Rule	258
10.3	Using z -Scores	262
10.3.1	Probabilities with z -Scores	263
10.3.2	Probabilities Using Technology	265
10.3.3	What do the probabilities mean?	269
10.4	The Central Limit Theorem	270

10.4.1	The Standard Normal Distribution	274
10.5	Exercises	275
10.6	Solutions	277
11	Rates of Change and Derivatives	281
11.1	Average Rates of Change	283
11.1.1	Exercises	287
11.1.2	Solutions	288
11.2	Instantaneous Rates of Change	289
11.2.1	Exercises	292
11.2.2	Solutions	293
11.3	Derivatives	295
11.3.1	When Derivatives Do Not Exist	298
11.3.2	Exercises	300
11.3.3	Solutions	301
11.4	Extrema Points	302
11.4.1	Applications	306
11.4.2	Exercises	312
11.4.3	Solutions	313

Chapter 1

Fractions, Decimals and Percents

1.1 Fractions

The importance of fractions is known by most students. You can't solve many real-life problems with only integers, so fractions, decimals and percents are necessary in working on such problems. Fractions can be as simple as making a problem a little more difficult to solve or can stop a student dead in their tracks. However, they do not need to make things harder or be a cause for stress. You just need to remember what we talk about here and before you know it, fractions will simply be something you use to solve problems and not something to be afraid of.

There are three types of fractions with which we may need to work. They are:

1. proper fractions: fractions where the denominator is larger than the numerator.
For example, $\frac{2}{5}$ is a proper fraction.
2. improper fractions: fractions where the denominator is smaller than the numerator.
For example, $\frac{12}{5}$ is an improper fraction.
3. mixed numbers: fractions that consist of a whole number part and a proper fraction.
For example, $2\frac{2}{5}$ is a mixed number.

We may need to convert back and forth between mixed number and improper fractions. If we have an improper fraction, we need to figure out how many times the denominator goes into the numerator evenly to get the equivalent mixed number. The remainder becomes the fractional part.

Example 1.1.1 Convert $\frac{31}{7}$ into a mixed number.

Solution If we find $31 \div 7$, we get 4 with a remainder of 3. So, the mixed number equivalent to $\frac{31}{7}$ is $4\frac{3}{7}$.

Conversely, we may want to convert from a mixed number into an improper fraction. In fact, we prefer improper fractions for our applications. When we are working with linear expressions, it is easier for us to deal with slopes as improper fractions. When we are dealing with matrices, it can be confusing and appear as multiple entries if we express the value as a mixed number. So, this skill is an important one for our purposes.

To convert a mixed number to an improper fraction, we need to change the whole number part into a fraction and then add the whole part to the fractional part. We will use the following example to illustrate how to do this.

Example 1.1.2 Convert $3\frac{4}{5}$ into an improper fraction.

Solution The question we need to ask ourselves is ‘How many fifths are there in 3 wholes?’ In practice, what we want to do is multiply the denominator by the whole number and add this to the numerator. In theory what we are doing is writing 3 as an equivalent fraction with the same denominator as the fractional part of the mixed number and then adding the two fractions. So, for this problem we get

$$3\frac{4}{5} = \frac{5 \times 3 + 4}{5} = \frac{19}{5}$$

1.1.1 Reducing Fractions

Often, it is useful for us to write fractions in simplest form. This means that we express the fraction so that there are no common factors in the numerator and denominator. To write a fraction in simplest form, we need to determine if there is a number that divides both the numerator and denominator evenly and if so, divide both to arrive at an equivalent fraction with no commonality.

Example 1.1.3 Write $\frac{21}{56}$ in simplest form.

Solution $21 = 3 \times 7$ and $56 = 8 \times 7$. So, they both are divisible by 7. Therefore, we get

$$\frac{21}{56} = \frac{3 \times 7}{8 \times 7} = \frac{3}{8}$$

1.1.2 Adding and Subtracting Fractions

If we want to add or subtract fractions, they must have a common denominator. The smallest such denominator is called the least common denominator, or LCD. It is the smallest number which both denominators evenly divide. As long as the denominator is the same for both fractions, it does not have to be the LCD. However, if we do not have the LCD then the answer we get will be in a different form. This is not wrong, though, and both fractions will reduce to the same value.

Example 1.1.4 Find $\frac{1}{8} + \frac{5}{8}$.

Solution Both fractions have the same denominator, so all we have to do is add the numerators and keep the common denominator.

$$\frac{1}{8} + \frac{5}{8} = \frac{6}{8}$$

The least common multiple of the numerator and denominator is not 1, so this fraction can be reduced.

$$\frac{6}{8} = \frac{2 \times 3}{2 \times 4} = \frac{3}{4}$$

If the denominators are the same then it does not matter if we are adding or subtracting; the process is the same other than the operation in question.

Example 1.1.5 Find $\frac{1}{8} - \frac{5}{8}$.

Solution Since the denominators are the same, we simply subtract the numerators and keep the common denominator. Then, reduce if necessary.

$$\frac{1}{8} - \frac{5}{8} = -\frac{4}{8} = -\frac{1}{2} \quad (1.1)$$

The following example will illustrate how to add and subtract fractions when the denominators are not the same.

Example 1.1.6 Find $\frac{1}{6} + \frac{5}{8}$.

Solution Since the denominators are different, we need to determine a common denominator. We will solve this in two different ways to show that we get the same answer whether we find the least common denominator or just any common denominator once we reduce the fractions.

Method 1 : Least common denominator

The least common denominator for these fractions is the least common multiple of 6 and 8. The question we need to ask ourselves here is ‘What is the smallest number that 6 and 8 both divide evenly?’ The number we are looking for is 24. We now need to write both fractions with a denominator of 24.

$$\begin{aligned} \frac{1}{6} \times \frac{4}{4} &= \frac{4}{24} \\ \frac{5}{8} \times \frac{3}{3} &= \frac{15}{24} \end{aligned}$$

Now that we have the common denominators, we simply add and reduce (if necessary) as we did before.

$$\frac{4}{24} + \frac{15}{24} = \frac{19}{24}$$

Method 2 : Any common denominator

Often, it is easier to find any common denominator when combining fractions. We very well may need to reduce the sum or difference regardless of the denominator we use, so we may not want to take the time to find the LCD. One easy way to find a common denominator is to multiply the two denominators together. This will give us at worst a common denominator that results in a fraction that probably needs to be reduced. Here, this denominator will be 48.

$$\begin{aligned} \frac{1}{6} \times \frac{8}{8} &= \frac{8}{48} \\ \frac{5}{8} \times \frac{6}{6} &= \frac{30}{48} \end{aligned}$$

So, when we add these fractions and reduce, we get the same sum as before.

$$\frac{8}{48} + \frac{30}{48} = \frac{38}{48} = \frac{19}{24}$$

As before, if we want to subtract two fractions that have different denominators, we find the common denominator using either method from above and then subtract (and reduce if necessary) the resulting fractions.

If the fractions we want to combine are mixed numbers, we want to first write them as improper fractions and then add or subtract the resulting values.

Example 1.1.7 Find $2\frac{1}{2} - 1\frac{3}{5}$.

Solution First, convert the mixed numbers to improper fractions.

$$2\frac{1}{2} = \frac{5}{2}$$
$$1\frac{3}{5} = \frac{8}{5}$$

Now we find a common denominator.

$$\frac{5}{2} \times \frac{5}{5} = \frac{25}{10}$$
$$\frac{8}{5} \times \frac{2}{2} = \frac{16}{10}$$

Now, we subtract as we did before.

$$\frac{25}{10} - \frac{16}{10} = \frac{9}{10}$$

If we wanted to find this difference without using improper fractions, we can combine the whole numbers and then combine the fractions separately and combine the two resulting values. We will get the same result, but it will require more work to get the answer.

1.1.3 Multiplying and Dividing Fractions

We actually already multiplied fractions when we were finding common denominators. To multiply fractions, all we need to do is multiply straight across; that is, multiply the numerators together and make it the numerator of the product and then multiply the denominators together and make this product the denominator of the answer. Notice that we may still need to reduce the result that we get.

Example 1.1.8 Find $\frac{7}{12} \times \frac{3}{5}$.

Solution When we multiply straight across, we get (don't forget to reduce)

$$\frac{7}{12} \times \frac{3}{5} = \frac{21}{60} = \frac{3 \times 7}{3 \times 20} = \frac{7}{20}$$

We could also 'cross-cancel' the fractions before we multiply. 3 and 12 share the common factor of 3, so we can reduce these values before we multiply because one is in the numerator and the other is in the denominator.

$$\frac{7}{12} \times \frac{3}{5} = \frac{7}{4} \times \frac{1}{5} = \frac{7}{20}$$

When dividing fractions, we are really multiplying by the reciprocal of the second fraction. The familiar phrase that we use to remind us of how to do this is 'flip-and-multiply'.

Example 1.1.9 Find $\frac{5}{8} \div \frac{3}{4}$

Solution We take the reciprocal of the second fraction and then revert to the multiplication rules.

$$\frac{5}{8} \div \frac{3}{4} = \frac{5}{8} \times \frac{4}{3} = \frac{5}{2} \times \frac{1}{3} = \frac{5}{6}$$

Note that if the quotient we obtained was not in simplified form, we would want to reduce the fraction.

1.1.4 Exercises

1. Convert the given fraction to simplest form.

(a) $\frac{27}{45}$

(b) $\frac{36}{100}$

(c) $\frac{81}{24}$

(d) $\frac{37}{15}$

2. Convert the given improper fraction to a mixed number.

(a) $\frac{45}{8}$

(b) $\frac{100}{12}$

(c) $\frac{232}{25}$

(d) $\frac{450}{18}$

3. Convert the following mixed numbers to improper fractions.

(a) $3\frac{2}{5}$

(b) $2\frac{9}{13}$

(c) $-3\frac{2}{9}$

(d) $11\frac{7}{14}$

4. Combine the following fractions using the indicated operation(s).

(a) $\frac{2}{3} + \frac{5}{8}$

(b) $\frac{3}{5} - \frac{7}{10}$

(c) $\frac{9}{15} - \frac{5}{8}$

(d) $2\frac{3}{4} + 1\frac{3}{5}$

(e) $\frac{3}{5} \times \frac{4}{7}$

(f) $\frac{5}{4} \div \frac{3}{5}$

(g) $4\frac{1}{3} \times 2\frac{3}{4}$

(h) $2\frac{5}{6} \div 3\frac{1}{3}$

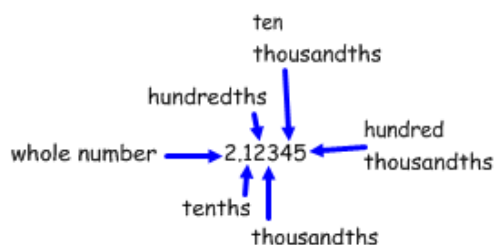
1.1.5 Solutions

1. (a) $\frac{27}{45} = \frac{3}{5}$
(b) $\frac{36}{100} = \frac{9}{25}$
(c) $\frac{81}{24} = \frac{27}{8}$
(d) $\frac{37}{15}$ is not reducible.
2. (a) $\frac{45}{8} = 5\frac{5}{8}$
(b) $\frac{100}{12} = 8\frac{4}{12} = 8\frac{1}{3}$
(c) $\frac{232}{15} = 9\frac{7}{15}$
(d) $\frac{450}{18} = 25$
3. (a) $3\frac{2}{5} = \frac{17}{5}$
(b) $2\frac{9}{13} = \frac{35}{13}$
(c) $-3\frac{2}{9} = -\frac{29}{9}$
(d) $11\frac{7}{14} = 11\frac{1}{2} = \frac{23}{2}$
4. (a) $\frac{2}{3} + \frac{5}{8} = \frac{16}{24} + \frac{15}{24} = \frac{31}{24}$
(b) $\frac{3}{5} - \frac{7}{10} = \frac{6}{10} - \frac{7}{10} = -\frac{1}{10}$
(c) $\frac{9}{15} - \frac{5}{8} = \frac{3}{5} - \frac{5}{8} = \frac{24}{40} - \frac{25}{40} = -\frac{1}{40}$
(d) $2\frac{3}{4} + 1\frac{3}{5} = \frac{11}{4} + \frac{8}{5} = \frac{55}{20} + \frac{32}{20} = \frac{87}{20}$
(e) $\frac{3}{5} \times \frac{4}{7} = \frac{12}{35}$
(f) $\frac{5}{4} \div \frac{3}{5} = \frac{5}{4} \times \frac{5}{3} = \frac{25}{12}$
(g) $4\frac{1}{3} \times 2\frac{3}{4} = \frac{13}{3} \times \frac{11}{4} = \frac{143}{12}$
(h) $2\frac{5}{6} \div 3\frac{1}{3} = \frac{17}{6} \div \frac{10}{3} = \frac{17}{6} \times \frac{3}{10} = \frac{17}{20}$

1.2 Decimals with ...

1.2.1 Decimals and Fractions

The only difference between decimals and fractions is the form in which we are writing the value. Depending on the situation with which we are working, however, we may need the number expressed in one form or another. So, we need to make sure we can convert back and forth. Before we look at conversions, let's make sure we remember the names of the place values.



Example 1.2.1 Write .4 as a fraction.

Solution Since there is only one decimal place, write the number we are given as the numerator (without the decimal point) and make the denominator a 10.

$$.4 = \frac{4}{10} = \frac{2}{5}$$

In general, we make the number we are given the numerator (after removing the decimal point) and make the denominator a 1 followed by a 0 for each decimal place.

Example 1.2.2 Write .456 as a fraction

Solution Since there are 3 decimal places, we need a denominator of 1000. This would give us

$$.456 = \frac{456}{1000} \tag{1.2}$$

Notice that the resulting fraction is not in simplified form, but we will not worry about that in this section since we already talked about that.

When converting a fraction into a decimal, the quickest way to do so is to divide the denominator into the numerator. We can also write the fraction so that the denominator is a power of 10 and then move the decimal point to find the equivalent decimal. This second method can only be done, however, when the factors of the denominator are only powers of 2 and 5. The next two examples will illustrate when we can and cannot use this second method.

Example 1.2.3 Write $\frac{3}{8}$ as a decimal.

Solution The denominator is 8, which is equal to $2 \times 2 \times 2$. If we multiply this by $5 \times 5 \times 5$, then we will have $10 \times 10 \times 10 = 1000$ in the denominator. In order to change the denominator, we also need to change the numerator by multiplying that by the same number, for otherwise we change the value of the fraction.

We only want to change the form and to do so we multiply by '1', just like we did when we were finding common denominators in the last section.

$$\frac{3}{8} \times \frac{125}{125} = \frac{375}{1000} = .375$$

If we would have divided 8 into 3 we would have gotten the exact same answer.

Example 1.2.4 Write $\frac{2}{7}$ as a decimal.

Solution The denominator has a factor besides 2 or 5, so it cannot be written as a fraction like we did above. The quickest way for us to write as a decimal is to divide. When we do, we get .285714... Remember, the dots means that the decimal does not terminate.

How do we know how many decimal places to round to? When dealing with money problems, always round to 2 decimal places. If it not a money problem, using two more decimal places than what we are performing the operations with is sufficient, 3-4 is better as the more decimal places you use, the more accurate your approximation will be.

1.2.2 Decimals and Percents

Similar to the relationship between decimals and fractions, decimals and percents also represent the same number. The conversion between the two is done by moving the decimal point. If we want to convert from a decimal to a percent, we move the decimal 2 places to the right. If we want to convert from a percent to a decimal, we move the decimal point 2 places to the left.

Example 1.2.5 Write 3.421 as a percent.

Solution Percent means 'per 100'. When we are given a decimal, we can set up a proportion to find the percent representation, but we will see that this reduces to moving the decimal point.

$$\begin{aligned}\frac{3.421}{1} &= \frac{p}{100} \\ (3.421)(100) &= (1)(p) \\ 342.1 &= p\end{aligned}$$

So, we get that $3.421 = 342.1\%$. And this could have been done by moving the decimal point two places to the right.

Example 1.2.6 Write 23.21% as a decimal.

Solution Since percent means 'per 100', we divide a percent by 100 to get the decimal representation. Since this is the same as moving the decimal places two places to the left, we here get $23.21\% = .2321$.

1.2.3 Exercises

1. Write $.32$ as a fraction in lowest terms.
2. Write $.156$ as a fraction in lowest terms.
3. Write 2.535 as a fraction in lowest terms.
4. Write $\frac{5}{12}$ as a decimal.
5. Write $\frac{12}{15}$ as a decimal.
6. Write $2\frac{3}{10}$ as a decimal.
7. Write $3\frac{7}{20}$ as a decimal.
8. Explain why $\frac{5}{6}$ cannot be written as a terminating decimal.
9. Write $.412$ as a percent.
10. Write 21.54 as a percent.
11. Write 21.54% as a decimal.
12. Write $.002\%$ as a decimal.

1.2.4 Solutions

1. $.32 = \frac{32}{100} = \frac{8}{25}$
2. $.156 = \frac{156}{1000} = \frac{39}{250}$
3. $2.535 = \frac{2535}{1000} = \frac{507}{200}$
4. $\frac{5}{12} = .41\bar{6}$
5. $\frac{12}{15} = \frac{4}{5} = .8$
6. $2\frac{3}{10} = 2.3$
7. $3\frac{7}{20} = 3.35$
8. We cannot write $\frac{5}{6}$ as a terminating decimal because the factorization of 6 is 2×3 , which contains integers besides just 2 and 5.
9. $.412 = 41.2\%$
10. $21.54 = 2154\%$
11. $21.54\% = .2154$
12. $.002\% = .00002$

1.3 Representing Change

Sometimes, it is sufficient to know a quantity and a context in order to understand a situation. If we say ‘1200 students graduated from SSU last year’, it makes complete sense. But if we say ‘38 more students graduated this year than last’, we have a very incomplete view of the number of graduates. Sure, it is great that more students graduated, but how much of an increase is it really when compared to the number of graduates? How does the increase in graduates relate to the number of enrolled students? It would be far less impressive a number if we learned that there were 100 more students in this year’s list of potential graduates versus learning that there were only 45 more potential graduates. What this comes down to is the difference between absolute change and percent change.

Definition 1.3.1 The absolute change is the difference between the new and old quantities and includes direction by indicating sign.

$$\text{absolute change} = \text{new quantity} - \text{old quantity}$$

Example 1.3.2 Ticket prices at Fenway Park were \$28 in section 42 last season and this season they cost \$32. So, the absolute change in price is

$$\$32 - \$28 = \$4$$

Definition 1.3.3 The percent change is a ratio of the absolute change as compared to the original quantity.

$$\text{percent change} = \frac{\text{absolute change}}{\text{original quantity}} \times 100$$

Note: The $\times 100$ above is so that the final answer is represented in percent form. Without this, our answer would be the corresponding decimal.

Example 1.3.4 From last year to this year, the percent change in those Red Sox tickets is

$$\frac{4}{28} \times 100 \approx 14.3\%$$

By having the percent change, it gives us a frame of reference - an extra \$4 might not sound like a lot, but over 14% more makes it sound like a much larger increase. Imagine if they increased 14% each year? We will revisit this when we get to exponential functions.

Example 1.3.5 There were 38 more graduates at Salem State last year than the previous year. If there were 1200 graduates this year, what is the percent change?

Solution Notice that here we are given the new quantity of 1200 graduates and the absolute change but we are not given the original quantity. Fortunately for us, we can find the original quantity so that we can find the percent change we want.

$$\text{New quantity} = \text{Original quantity} + \text{increase in graduates}$$

$$1200 = \text{Original quantity} + 38$$

$$\text{Original quantity} = 1162$$

So, now that we know the original number of graduates, we can find the percent change.

$$\frac{38}{1162} \times 100 = 3.27\%$$

Example 1.3.6 *You want to buy a new television during the big sales right before the Super Bowl and see that a \$1999 TV is \$350 off. What do you expect to see for the percent off on the website?*

Solution We can solve this using the same idea of percent change. We know the original price and we know the absolute change, but we have to remember that in this situation, it is a negative quantity because it will cost less. So, we have

$$\frac{-\$350}{\$1999} \times 100 \approx -17.5\%$$

1.3.1 Exercises

1. If a salesman has \$212,000 in sales last year but only had \$193,000 in sales this year, what is the absolute change?
2. If a salesman has \$212,000 in sales last year but only had \$193,000 in sales this year, what is the percent change?
3. If a seat for a regular season game at Fenway Park costs \$42 but it costs \$95 during the playoffs, what is the percent change in price?
4. One salesman says he had a better month because he sold \$25,000 in product last month and \$32,000 this month. Another salesman says he was better because he went from \$30,000 in sales last month to \$36,000 in sales. How would you decide?
5. If you invested \$2000 and had a percent change of 35%, what is the absolute change?
6. If you invested \$5000 and had an absolute change of -\$1500, what is the percent change?

1.3.2 Solutions

1. $193,000 - 212,000 = -\$19,000$

2. $\frac{-19,000}{212,000} \times 100 = -8.962\%$

3. $\frac{95-42}{42} \times 100 \approx 126\%$ increase in price.

4.

$$\frac{32,000 - 25,000}{25,000} \times 100 = 28\%$$

$$\frac{36,000 - 30,000}{30,000} = 20\%$$

The first salesman did better from a percent change standpoint.

5. $2000 \times .35 = \$700$. So, your return is \$2700 with an absolute change of \$700.

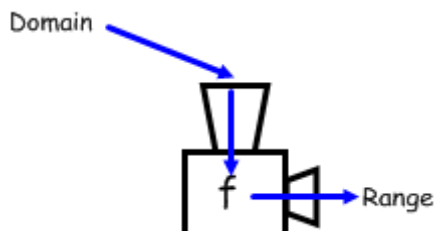
6. $\frac{-1500}{5000} \times 100 = -30\%$

Chapter 2

Functions

2.1 What Is A Function?

Formally speaking, a **function** is a correspondence between a set of input values, called the **domain**, and the set of output values, called the **range**, such that each element of the domain correspond to exactly one element of the range. This can seem confusing, but it doesn't have to be so bad. In simple English, a function can be thought of like a 'machine' that takes in one set of values and produces an output of another set of values. This 'machine', or function, has the special property that for each input, the output is unique.



A function could also be thought of as a rule that assigns a unique output value to each member of the input set.

Example 2.1.1 Explain why the given table represents a function. State a possible rule with the domain and range.

<i>Boston</i>	<i>Massachusetts</i>
<i>Baltimore</i>	<i>Maryland</i>
<i>Smithtown</i>	<i>New York</i>

Solution This is a function because each city corresponds to exactly one state. The domain would be {Boston, Baltimore, Smithtown} and the range is {Massachusetts, Maryland, New York}. The rule here could be that we input a city name and the output would be the state containing that city.

Example 2.1.2 Determine if the given table represents a function.

<i>Salem</i>	<i>Massachusetts</i>
<i>Lynn</i>	<i>Massachusetts</i>
<i>Saugus</i>	<i>Massachusetts</i>

Solution This is a function. Notice that every input from our domain (city names) corresponds to exactly one state. Nothing in the definition of a function says that the outputs must be unique too - if that were the case, we would say we have a 1 – 1 function.

Example 2.1.3 Determine if the given table represents a function.

2	4
-2	4
3	9
3	-9

Solution This is not a function. It looks like we have the function $f(x) = x^2$, but then when we get to the last two pairs, we see (3, 9) and (3, -9), which is to say that 3 maps to different output values and this violates the definition of a function.

2.1.1 Exercises

1. Determine if the given table could represent a function.

x	1	2	3	4	5
y	2	5	8	11	14

2. Determine if the given table could represent a function.

x	1	2	3	4	5
y	3	3	3	3	3

3. Determine if the given table could represent a function.

x	1	2	3	2	5
y	2	4	6	8	10

4. Determine if the given table could represent a function.

x	1	2	3	2	5
y	2	4	6	4	10

5. Determine if the given table could represent a function.

Name	Owen	Jack	Edward	Cameron
Home Town	Salem	Danvers	Lynn	Salem

6. Fill in a value for a that would make the given table represent a function.

x	1	3	5	7	9
y	2	4	a	8	10

7. Fill in a value for b that would make the given table not represent a function.

x	4	3	2	b	1
y	12	9	6	12	9

1. Since the y values change by 3 units as the x values change by 1 unit, this could represent a function. Since the change is constant, it appears to be the linear function $y = 3x - 1$.
2. Since all of the y values are the same, this could represent the constant function $y = 3$.
3. Since we have the points $(2, 4)$ and $(2, 8)$ here, we have different outputs for the same input. Therefore, this table could not represent a function.
4. Here, we have exactly one output for each input, so this could represent a function. Note that we have 2 repeated twice, but its output value is a 4 in each case. Therefore, we do not violate the definition of a function.
5. Since each name on the list is unique, this table could represent a function. It is OK that we have outputs repeated.
6. The y values are increasing by 2 as the x values are increasing by 2. Sticking with this pattern, we would need $a = 6$ to follow this pattern. But, any real number for a would give us a function as all of the given x values are different.
7. Since we are looking for the input value that gives an output of 12 and we already have the pair $(4, 12)$ in the table, we want $b = 4$ here. But, any choice for b besides 1, 2 or 3 would work as we would not be violating the definition of a function provided we do not have one x value mapping to two different y values.

2.2 Notation

There are a many different ways for us to express functions symbolically. We will talk about two of these ways here.

2.2.1 Set Builder Notation

We often denote functions by describing the elements in the set. This can be done by using set builder notation. The general format for this is

$$A = \{x \mid \text{properties}\}$$

which, in English, says ‘ A is the set of all x such that x has the listed properties’.

Example 2.2.1 *Using set builder notation, describe the set of all even numbers.*

Solution We can use any capital letter to name the set. We often use a letter that makes some kind of logical sense in the context of the problem. So we could write $E = \{x \mid x \text{ is an even number}\}$.

This would be far from a technical definition but it is a way to express the set. The following is a more precise way to describe a set of numbers.

Example 2.2.2 *Using set builder notation, describe the set of all odd numbers.*

Solution $O = \{x \mid x = 2n + 1 \text{ where } n \text{ is an integer}\}$.

2.2.2 $y = f(x)$ Notation

When describing sets, we often use set builder notation, but when describing mathematical functions, we use $f(x)$ to denote a function where x is the variable. The letter ‘ f ’ is often used simply because the word function begins with f but we can actually use any letter we choose. When working on word problems we often choose a letter that makes sense in the context of the problem. Similarly, we use the letter ‘ x ’ for the variable much of the time but we can choose any letter we’d like here also.

Example 2.2.3 *Write the function describing the area of a square with side of length s inches.*

Solution Since we are talking about area, we will use ‘ A ’ for the function name and ‘ s ’ for the variable. This would give us $A(s) = s^2$.

When talking about the kinds of functions that we graph, we often use ‘ $y =$ ’. This is actually the same as using ‘ $f(x) =$ ’, except for the fact that we are specifying that x is the variable in the second notation. For example, if we are given the function $y = 2x + 1$, we are describing the line with y -intercept $(0, 1)$ and slope $m = 2$. For each x we choose, we get a distinct y value. If we change the y to $f(x)$, we get $f(x) = 2x + 1$. This gives is a linear function with the same y -intercept and slope as the other function. There is no difference besides the way we are expressing the function.

2.2.3 Evaluating the Function $y = f(x)$

Suppose we wanted to find the value of the function $f(x) = 2x + 1$ at the point where $x = 4$. We do so by substituting into the expression on the right and we denote it by replacing the x with 4 on the left. That is, $f(4)$ is the value of the function f when $x = 4$.

$$f(4) = 2(4) + 1 = 9$$

We can also express the answer as the ordered pair $(4, 9)$.

Example 2.2.4 *Suppose that the distance a falling object has travelled after t seconds is given by $d(t) = 16t^2$ feet. What is the distance travelled of the object after 10 seconds? After 20 seconds? How long would it take for the object to fall 14400 feet?*

Solution Notice that we used d as the function name, since we are talking about the distance the object has travelled, and that we used t for the variable, since we are concerned with the distance travelled after t seconds.

At $t = 10$ seconds, we have $d(10) = 16(10)^2 = 1600$ feet.

At $t = 20$ seconds, we have $d(20) = 16(20)^2 = 6400$ feet.

For the final part of the question, we are looking for the value of t so that $d(t) = 14400$. Since we have an expression for d in terms of t , we can replace $d(t)$ with what it equals. This will give us an equation we can solve.

$$\begin{aligned} 16t^2 &= 14400 \\ t^2 &= 900 \\ t &= 30 \text{ seconds} \end{aligned}$$

Some of you may be asking why we only have one answer here. We took the square root of both sides, so why don't we consider the negative value as well? When we take the square root of a real number, we have to take the positive and negative root, since the square of both will equal the same number. That is, $3^2 = (-3)^2 = 9$. In the case of our word problem, however, it makes no sense to talk about negative time, so we only consider the positive square root. This root is called the principle square root.

We have to be careful, however, with this usage of \pm . We introduced the square root in order to solve the problem, so we had to consider both answers and then dismiss the negative root because of context. But if we are given the square root in the problem, there is only one correct answer and that is the principle root.

Example 2.2.5 *If $x^2 = 9$ then $x = \pm 3$ but $\sqrt{9} = 3$ only.*

2.2.4 Exercises

1. Use set builder notation to define the function of all integers that are one less than three times a given integer.
2. Use set builder notation to describe the set of players on a baseball team.
3. Write a function describing the volume of a circle with radius r .
4. Write a function describing one more than the cube of an integer x .
5. Given the function $f(x) = x^2 + 3x$, find
 - (a) $f(2)$
 - (b) $f(-3)$
 - (c) $f(t)$
 - (d) $f(x+2)$
6. Suppose the height of an object dropped from a height of 300 feet is given by $h(t) = -16t^2 + 300$ where t is the time in seconds. Find the height of the object at
 - (a) $t = 3$ seconds
 - (b) $t = 4$ seconds
 - (c) Explain why we cannot evaluate at $t = -1$.
 - (d) How long does it take for the ball to hit the ground? (Hint: the ball hits the ground when the height is what?)

2.2.5 Solutions

1. $S = \{x \mid x = 3n - 1 \text{ for } n \in \mathbb{Z}\}$
2. $P = \{p \mid p \text{ is a player on the team}\}$
3. $V = \pi r^2$ where r is the radius of the circle.
4. $f(x) = x^3 + 1$
5.
 - (a) $f(2) = 2^2 + 3(2) = 10$
 - (b) $f(-3) = (-3)^2 + 3(-3) = 0$
 - (c) $f(t) = t^2 + 3t$
 - (d) $f(x+2) = (x+2)^2 + 3(x+2) = x^2 + 4x + 4 + 3x + 6 = x^2 + 7x + 10$
6.
 - (a) $h(3) = -16(3)^2 + 300 = -144 + 300 = 156$ feet
 - (b) $h(4) = -16(4^2) + 300 = -256 + 300 = 44$ feet
 - (c) We cannot evaluate at $t = -1$ because that would be negative time. We can evaluate functions at negative input values, in general, but in this context, a negative input value is not in our domain.
 - (d) We need to find the time(s) when the height is 0 feet.

$$\begin{aligned} -16t^2 + 300 &= 0 \\ 16t^2 &= 300 \\ t^2 &= \frac{300}{16} \\ t &= \pm \frac{\sqrt{300}}{4} \approx \pm 4.33 \end{aligned}$$

Since we cannot have negative time here either, the ball would hit the ground after approximately 4.33 seconds.

2.3 Functions from Tables

As we saw earlier, we can represent functions using tables. This can be particularly useful if we have a small set of data to represent or if we don't have an easily defined rule.

Example 2.3.1 Consider the closing figures for the Dow Jones Industrial Average⁴ in the table below:

Date	8/18	8/19	8/22	8/23	8/24	8/25	8/26
Dollars	18,597.70	18,552.57	18,529.42	18,547.30	18,481.48	18,448.41	18,395.40

This is certainly a function as each day can only have one dollar amount for its closing figure. We have as our domain the set of dates and the range is the closing figures. But we don't have a rule we can assign to this function as there are too many variables to be considered.

Example 2.3.2 The following table gives price per child for entrance to a park for a school field trip. The park charges a flat \$100 for hosting the event and additionally charges the following:

Number of Children	Price per Child
1	\$5
2	\$10
5	\$25
10	\$50

Since each quantity of children has its own price, we have a function here. The domain is the number of children and the range is the set of prices. And, since it works out to \$5 per child no matter how many children there are, we have a constant increase in price as the number of children increases. When we factor in the flat fee, we have a function $C(n) = 100 + 5n$.

When comparing the last two examples, 2.3.1 is an example where it is not reasonable to expect a function in the form of an equation but 2.3.2 is an example where we have the ability to find a formula that will allow us to find other values.

⁴Data obtained from <https://finance.yahoo.com>

2.3.1 Exercises

1. Find the equation of the function represented by the given table.

x	y
2	6
4	10
6	14
8	18

2. Find the equation of the function represented by the given table.

x	y
3	-2
6	1
7	2
10	5

3. Find the equation of the function represented by the given table.

x	y
1	2
2	-2
3	-6
4	-10

4. Explain why the daily average temperature over the course of one year is a function with variable t representing the day.
5. Explain why we cannot have a function with the input being a person's name and the output being their age.

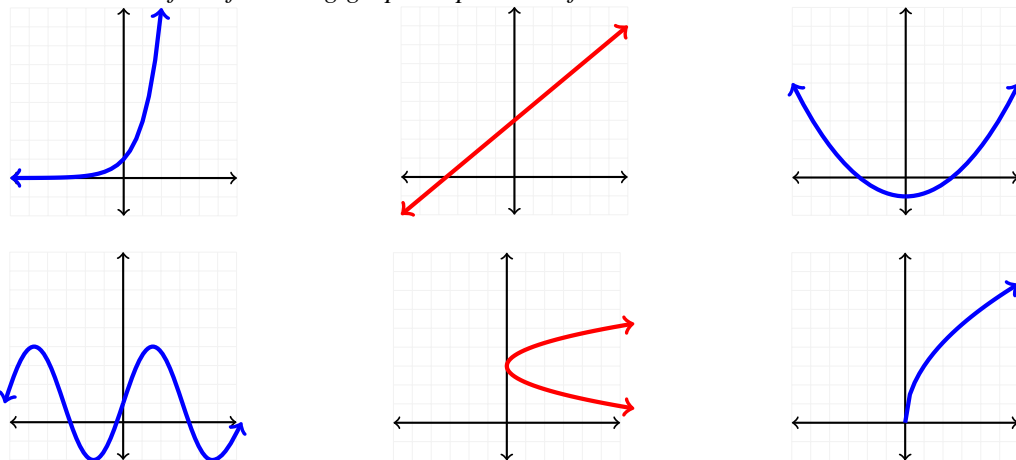
2.3.2 Solutions

1. $y = 2x + 2$
2. $y = x - 5$
3. $y = -4x + 6$.
4. This situation would be represented by a function because we can only get one average temperature per day and each day would only occur once. Therefore, our inputs would all be unique and we would only get one output value for each of these inputs.
5. Unless every person had a unique name, we would get more than one person with a given name. This would mean that everyone with a certain name would have to be the same age in order for this situation to represent a function. Since not every person with the same name is of the same age, however, this situation cannot be represented by a function.

2.4 Graphs of Functions

We will focus on different types of functions later in this text, but when we look at a graph, we want to be able to identify whether or not it is a function regardless of the formula. So, what are we looking for to make this determination?

Example 2.4.1 Which of the following graphs represent a function?

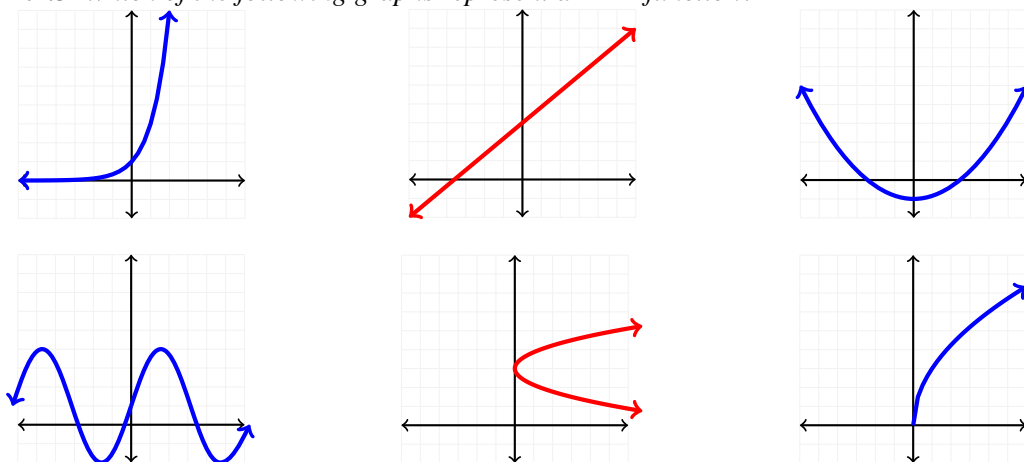


Solution All of the given graphs are functions, with the exception of the fifth one. How do we know this? When thinking of a function, we know that there can be only one output for any input. Visually, this is saying that if we drew a vertical line over the graph, we cannot find anywhere that the line would touch the curve more than once. For all but the middle graph in the bottom row, we can do this; for the middle graph in the bottom row, we would touch the graph twice for all x values greater than 0. So, this is not a function.

Definition 2.4.2 *The Vertical Line Test*

If a graph represents a function, every vertical line drawn over the graph must touch the curve at no more than one point.

Example 2.4.3 Which of the following graphs represent a 1 – 1 function?



Solution We have not formally defined a 1 – 1 function, so let’s do that here.

Definition 2.4.4 *A function f is 1 – 1 if whenever $f(a) = f(b)$, $a = b$. That is, the only way two output values can be the same is if the input values were the same as well.*

So for which of the functions in our example does this property hold? it is true for the first two graphs and for the last one. When we look at a graph, we are looking to see that there is only one y value for each x value, which is the same as only having one y value on any horizontal line.

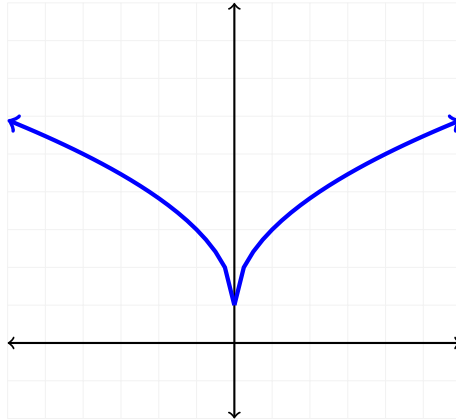
Definition 2.4.5 *Horizontal Line Test*

If a graph represents a 1 – 1 function, every horizontal line drawn over the graph can only touch the graph at most at one point.

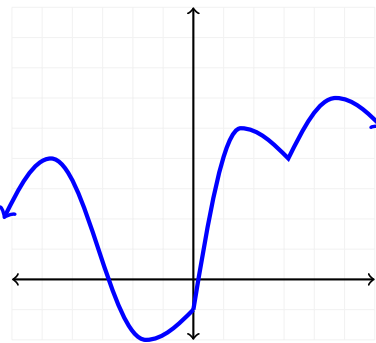
Note: we would only touch the graph in the middle of the bottom row with a horizontal line at most one, but we didn’t bother to consider this here because we already determined that this graph did not represent a function.

2.4.1 Exercises

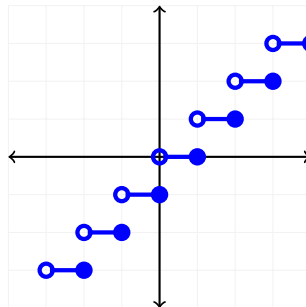
1. Determine if the given graph represents a function. If so, determine if the function is $1 - 1$. Explain your answer.



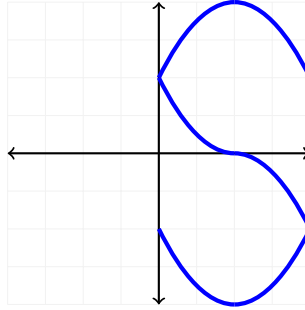
2. Determine if the given graph represents a function. If so, determine if the function is $1 - 1$. Explain your answer.



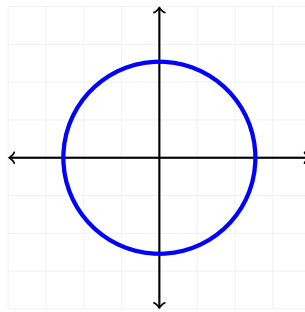
3. Determine if the given graph represents a function. If so, determine if the function is $1 - 1$. Explain your answer.



4. Determine if the given graph represents a function. If so, determine if the function is $1 - 1$. Explain your answer.



5. Determine if the given graph represents a function. If so, determine if the function is $1 - 1$. Explain your answer.



1. This graph represents a function because each input has exactly one output. It is not $1 - 1$, however, as we get two different x values that give the same y value.
2. This graph represents a function because each input has exactly one output. It is not $1 - 1$, however, as we get two different x values that give the same y value.
3. This graph represents a function because each input has exactly one output. It is not $1 - 1$, however, as we get two different x values that give the same y value.
4. This is not a function because we get multiple y values for one x value.
5. This is not a function because we get multiple y values for one x value.

Chapter 3

Linear Functions

3.1 Linear Functions

A **linear function** is simply a straight line. In this section, we will discuss how to find the equation of a linear function.

Definition 3.1.1 A linear function is an equation of the form $y = mx + b$ where m represents the slope of the line and b represents the y -intercept. Linear equations are an example of what is known as a 2-parameter family because there are two values that determine the orientation of the line.

3.1.1 Slope

Definition 3.1.2 The slope of a linear function tells us how steep the line is and whether it is increasing or decreasing. You might hear the phrase ‘rise over run’ when slope is discussed. This refers to the change in the y value with respect to the change in the x value.

When we want to find the slope, we need two points to determine this value. The formula we use is

$$\text{slope} = m = \frac{\text{rise}}{\text{run}} = \frac{\Delta y}{\Delta x} = \frac{y_2 - y_1}{x_2 - x_1}$$

Example 3.1.3 Find the slope of the line containing the points $(3, 2)$ and $(-3, 12)$.

Solution It does not matter which point you consider point 1 and which you consider point 2 as long as you are consistent; that is, as long as x_1 and y_1 come from the same ordered pair, you can choose either point as your first point. So without loss of generality, consider $(3, 2)$ to be point 1.

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{12 - 2}{-3 - 3} = \frac{10}{-6} = -\frac{5}{3}$$

So, the slope of the line containing these two points is $-\frac{5}{3}$. Notice however, that we did not make it into a mixed number. For the purpose of plotting linear functions, it is better to leave the slope as an improper fraction.

What would have happened if we chose $(-3, 12)$ as point 1?

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{2 - 12}{3 - (-3)} = \frac{-10}{6} = -\frac{5}{3}$$

We again get $-\frac{5}{3}$ for the slope of the line. So, we can see here that it does not matter which ordered pair we choose as point 1 as long as we are consistent.

Example 3.1.4 Find the slope of the line passing through the points $(-2, -1)$ and $(4, 2)$.

Solution Without loss of generality, we choose $(-2, -1)$ as point 1.

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{2 - (-1)}{4 - (-2)} = \frac{3}{6} = \frac{1}{2}$$

Why is slope so important? Slope represents a constant rate of change when we talk of linear functions. But, what if we could take the rate of change between two point so close together that we can't tell the difference between the points? What if we could take the rate of change at only one point? If we could, we would have the *instantaneous* rate of change. Another name for the instantaneous rate of change is the derivative, which gives us the ability to find extrema points of functions and so solve optimization problems with regards to marginality, among other ideas. We will revisit this later, but for now, it is important to remember why we want to stress the importance of slope.

3.1.2 Special Cases for Slope

Example 3.1.5 Find the slope of the line containing the points $(-2, 3)$ and $(4, 3)$.

Solution Without loss of generality, choose $(-2, 3)$ as point 1.

$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{3 - 3}{4 - (-2)} = \frac{0}{6} = 0$$

Notice here that the y -coordinates for the two points are both 3. Because of this, it doesn't matter what the x -values are as long as they are not the same too (or it is the same point); for any x -values, we get a rise of 0. When this occurs, we get **horizontal line**. A horizontal line has a slope of 0 and has as its equation $y = b$, where b is the common y -value.

Example 3.1.6 Find the slope of the line passing through the points $(2, 2)$ and $(2, -1)$.

Solution Without loss of generality, choose $(2, 2)$ as point 1.

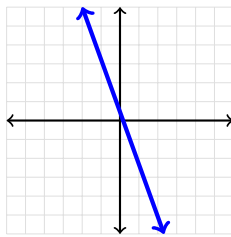
$$m = \frac{y_2 - y_1}{x_2 - x_1} = \frac{-1 - 2}{2 - 2} = \frac{-3}{0} = \text{undefined}$$

Here, the denominator is 0. This is saying that there is no 'run' for this line. When we get a 0 in the denominator, we have a **vertical line**. The line is still straight, but the slope is undefined. Lines of this type are given by the equation $x = a$ where a is the common x value of the two points. And, remember the definition of a function - vertical lines are not functions. But, they are important because they often serve as constraints in real-life applications.

For linear equations that have a slope that is defined and not equal to 0, there are two possibilities; the slope can be positive or negative. But what does knowing the sign of the slope tell us? From this we know whether the function will be increasing or decreasing. That is, we will know whether the function will rise or fall as we go from left to right.

Example 3.1.7 Describe the orientation of the linear function $y = -3x + 1$.

Solution Since the slope is negative, the line will fall to the right.



3.1.3 The Slope-Intercept Form of a Linear Function

Now that we know how to find the slope of a line, we want to be able to find the equation of a line passing through some given points. One way to accomplish this is to use the **slope-intercept equation**, $y = mx + b$. The important aspect of this equation is that we can graph the line directly from this equation (we will get to that later in the chapter). The two pieces that we need to find to express a line in this form are the slope, m , and the y -intercept, b . The y -intercept is the point at which the line crosses the y -axis. It is given by the ordered pair $(0, b)$. If we know the slope and one point the line passes through, we can find the value of b .

Example 3.1.8 Find the y -intercept of the line with slope $m = 3$ that passes through the point $(-3, 9)$.

Solution The given point, as all ordered pairs are, is in the form (x, y) . Using this point and the given slope, we know three of the four unknown values of the equation $y = mx + b$. We know $m = 3$, $x = -3$ and $y = 9$. We have left to find the value b so that we have the y -intercept, $(0, b)$, for our equation.

$$\begin{aligned}y &= mx + b \\9 &= 3(-3) + b \\9 &= -9 + b \\b &= 18\end{aligned}$$

That is, the y -intercept of this line is $(0, 18)$. This means that the linear function in question crosses the y -axis this point.

An important distinction - people commonly say b is the y -intercept of a linear function, but this is not true. b is the value that represents the y -intercept, but the y -intercept is a point and must be expressed as the ordered pair $(0, b)$.

Example 3.1.9 Find the y -intercept of the line passing through the points $(3, 2)$ and $(6, -4)$.

Solution We are not given the slope here as we were in the last example. However, we know how to find the slope of a line given two points.

$$\frac{y_2 - y_1}{x_2 - x_1} = \frac{-4 - 2}{6 - 3} = \frac{-6}{3} = -2$$

Now that we have the slope, we can choose either point to find the y -intercept. Here we will choose $(6, -4)$.

$$\begin{aligned}y &= mx + b \\-4 &= -2(6) + b \\-4 &= -12 + b \\b &= 8\end{aligned}$$

If we would have chosen the other point, we will end up with the same y -intercept because the two points are on the same line.

$$\begin{aligned}y &= mx + b \\2 &= -2(3) + b \\2 &= -6 + b \\b &= 8\end{aligned}$$

So, the y -intercept is the point $(0, 8)$.

What we have done in these examples is show how to find the respective pieces of a linear function. When we put these all together, we have the equation of a linear function.

Example 3.1.10 Find the equation of the linear function passing through the points $(-3, 5)$ and $(2, 3)$.

Solution We first need to find the slope of the line.

$$\frac{5-3}{-3-2} = \frac{2}{-5} = -\frac{2}{5}$$

Now we can find the y -intercept. We can choose either point as the one we want to use to solve for b . Here we will choose the point $(2, 3)$.

$$\begin{aligned}y &= mx + b \\3 &= -\frac{2}{5}(2) + b \\3 &= \frac{-4}{5} + b \\\frac{5}{5}(3) &= -\frac{4}{5} + b \\\frac{15}{5} &= -\frac{4}{5} + b \\\frac{19}{5} &= b\end{aligned}$$

Why did we multiply by $\frac{5}{5}$ in the 4th step above? The reason is that we have to find a common denominator in order to add 3 and $\frac{4}{5}$. The least common denominator is 5. We need to make the left hand side have the same denominator as the right hand side but do not want to change the relationship we have in the equation. So, we multiply by '1', or in this case $\frac{5}{5}$ to accomplish this task.

Now that we have all of the pieces we need, we can write the equation in slope-intercept form.

$$y = mx + b \Rightarrow y = -\frac{2}{5}x + \frac{19}{5}$$

3.1.4 Intercepts

While on the subject of intercepts, sometimes it is important for us to be able to find not only where a function crosses the y -axis, but the x -axis as well. We will see why a little later when we get to graphing linear equations, but for the time being we want to make sure we can find them.

- To find the y -intercept, set $x = 0$ in the equation and solve for y . If we solve and get $y = b$ then the y -intercept is $(0, b)$.
- To find the x -intercept, set $y = 0$ in the equation and solve for x . If we solve and get $x = a$ then the x -intercept is $(a, 0)$.

Example 3.1.11 Find the x and y intercept of $y = -2x - 3$.

Solution When in slope-intercept form, we can just get the value of b from the given equation. But when in standard form, we need to either write the equation in slope-intercept form or do what is explained above.

x -intercept:

$$0 = -2x - 3 \Rightarrow 2x = -3 \Rightarrow x = -\frac{3}{2}$$

So, the x -intercept is $(-\frac{3}{2}, 0)$.

y -intercept:

$$y = -2(0) - 3 \Rightarrow y = -3$$

So, the y -intercept is $(0, -3)$.

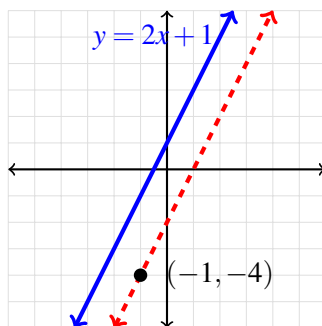
3.1.5 Parallel and Perpendicular Lines

Sometimes, it is valuable for us to be able to find a line that is either parallel or perpendicular to a given line. The important thing to note in each of these cases is the relationship between the slopes of the two lines.

Parallel lines have the property that they lie in the same plane but never touch or meet. Lines that are parallel necessarily have the same slope but a different y -intercept, since if the y -intercept was the same then would be the same line. By knowing two lines are parallel, we additionally only need one point that the second line passes through for us to be able to find the equation of the new line.

Example 3.1.12 Find the equation of the line parallel to $y = 2x + 1$ that passes through the point $(-1, -4)$.

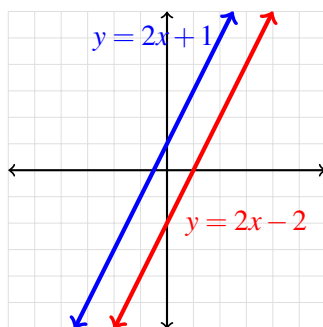
Solution We are looking for the equation of the line that is parallel to $y = 2x + 1$. So, we are roughly looking for the following line.



As we can see from the picture, any line parallel to the line $y = 2x + 1$ must have the same slope. What we need is the specific line that has the right y -intercept so that it will pass through the point $(-1, -4)$.

$$\begin{aligned} y &= mx + b \\ -4 &= 2(-1) + b \\ -4 &= -2 + b \\ b &= -2 \end{aligned}$$

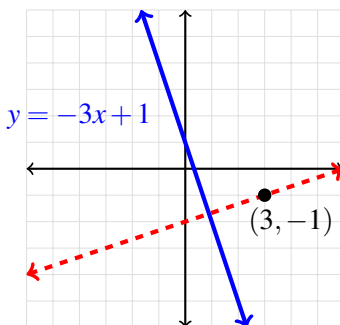
So, taking the y-intercept to be $(0, -2)$ gives the particular parallel line we are looking for, namely $y = 2x - 2$.



From a given line, we can also find the equation of the line that is perpendicular to that line. **Perpendicular lines** are those that intersect at right (90°) angles. The slopes will not be the same here, however. Instead we need the slopes to be ‘opposites’ of each other; that is, we need the slopes to be negative reciprocals of each other. (For justification, see the section at the end of the chapter.)

Example 3.1.13 Find the equation of the line perpendicular to $y = -3x + 1$ and passing through the point $(3, -1)$.

Solution We are looking for the equation of the line that passes through $y = -3x + 1$ that passes through $(3, -1)$ and intersects the original line at a right angle.

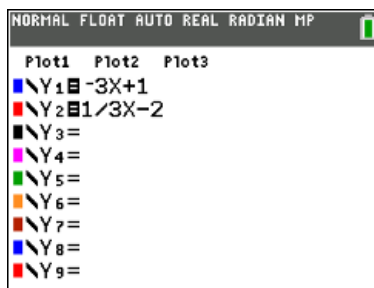


Since the line we are looking for is perpendicular to our original line, we know that the slope has to be the negative reciprocal of the slope of the given line. So, since the slope of the given line is -3 , the slope of the new line will be $\frac{1}{3}$. Knowing this, we can use our slope-intercept form of a line and the given point to find the y-intercept and ultimately the equation of the line we want.

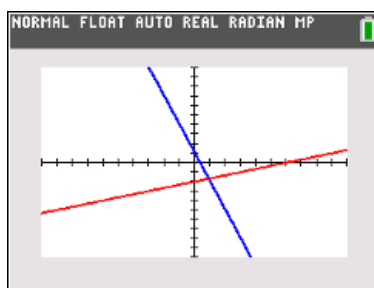
$$\begin{aligned} y &= mx + b \\ -1 &= \frac{1}{3}(3) + b \\ -1 &= 1 + b \\ b &= -2 \end{aligned}$$

So, y-intercept is -2 and the equation we are looking for is $y = \frac{1}{3}x - 2$.

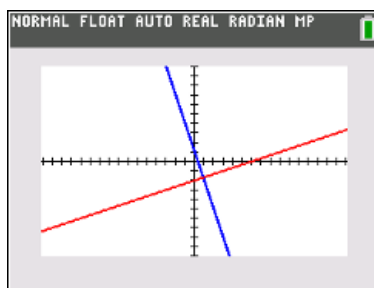
If we wanted to plot these together, we could do so by hand or we could use the TI-series graphing calculator.



This would give us the pair of lines.



When plotting these equations in the calculator, this is what we get. Notice that the two lines do not look perpendicular even though the equations put in the calculator are supposed to be. This is because of the view on the calculator screen. You have to be careful with perceptions when using the calculator. In order to get a more accurate view on the calculator screen, we can use the **ZSquare** option. Once the graph is on the screen, press **ZOOM** and then select **ZSquare**.



Now the lines look perpendicular.

3.1.6 Standard Form $ax + by = c$

We are not always given the equation in slope-intercept form. Earlier it was mentioned that we want it in this form because we can graph the line directly from the equation. The other way that we can express the same line is in **standard form**. The equation is essentially the same, but some algebra is required to convert from one form to the other and we want no fractions in this form.

Example 3.1.14 Convert the equation $y = -\frac{4}{5}x + \frac{19}{5}$ into standard form.

Solution We need to put the equation into standard form $ax + by = c$. To do this, we will first get rid of the fractions by multiplying by a common denominator and then rewriting by moving the terms to the correct sides.

$$\begin{aligned}y &= -\frac{4}{5}x + \frac{19}{5} \\5y &= 5\left(-\frac{4}{5}x\right) + 5\left(\frac{19}{5}\right) \\5y &= -4x + 19 \\4x + 5y &= 19\end{aligned}$$

Example 3.1.15 Convert the equation $2x + 3y = 5$ into slope-intercept form.

Solution We want our solution in the form $y = mx + b$. So we will use some algebraic steps to accomplish our goal.

$$\begin{aligned}2x + 3y &= 5 \\3y &= -2x + 5 \\y &= -\frac{2}{3}x + \frac{5}{3}\end{aligned}$$

3.1.7 Is the point on the line?

If we are given a linear function, then all points that satisfy that equation are said to be on the line. A **line** is a series of points that all have the relationship that the slope between any two points on the line is the same. If we want to see if a point is on a line, we need to substitute the x value into the equation and see if the answer we obtain is the y value from the ordered pair. If it is the same then the point is on the line. If it is not the same then it is not a point on the line.

Example 3.1.16 Determine if the point $(3, 12)$ is on the line $y = 3x + 4$.

Solution If we substitute $x = 3$ into $y = 3x + 4$, we get $y = 3(3) + 4 = 13$ which is not equal to 12. Therefore, the point $(3, 12)$ is not on the line $y = 3x + 4$.

Example 3.1.17 Determine if the point $(4, -7)$ is on the line $y = -2x + 1$.

Solution If we substitute $x = 4$ into $y = -2x + 1$, we get $y = -2(4) + 1 = -7$. So, the point $(4, -7)$ is on the line $y = -2x + 1$.

3.1.8 Exercises

1. Put $y = \frac{2}{3}x + 1$ into standard form.
2. Put $y = -x + 3$ into standard form.
3. Put $17x + 13y = 11$ into slope-intercept form.
4. Put $x - y - 5 = 0$ into slope-intercept form.
5. Find the x and y intercept of $x - y = 3$.
6. Find the x and y intercept of $y = -\frac{2}{3}x + 6$.
7. Find the x and y intercept of $y = -6x$.
8. Find the x and y intercept of $x + 5 = 0$.
9. Determine if $(1, 1)$ is on the line $3x + 5y = 15$.
10. Determine if $(-1, 2)$ is on the line $2y + 3x = 1$.
11. Determine if $(5, 3)$ is on the line $\frac{1}{3}x - \frac{1}{5}y = -1$.
12. Determine if $(4, 6)$ is on the line $y = \frac{3}{4}x + 3$.
13. The slope of a vertical line is _____.
14. The slope of a horizontal line is _____.

3.1.9 Solutions

1. $-2x + 5y = 5$
2. $x + y = 3$
3. $y = -\frac{17}{13}x + \frac{11}{13}$
4. $y = x - 5$
5. x -intercept: $x - 0 = 3 \Rightarrow x = 3$ giving x -intercept $(3, 0)$.
 y -intercept: $0 - y = 3 \Rightarrow y = -3$ giving y -intercept $(0, -3)$.
6. x -intercept: $0 = -\frac{2}{3}x + 6 \Rightarrow x = 4$ giving x -intercept $(4, 0)$.
 y -intercept: $y = -\frac{2}{3}(0) + 6 \Rightarrow y = 6$ giving y -intercept $(0, 6)$.
7. Here, because the function passes through the origin, the x -intercept and the y -intercept are both $(0, 0)$.
8. When we rewrite this equation, we get $x = -5$. This is a vertical line, so there is no y -intercept and the x -intercept is $(-5, 0)$.
9. $3(1) + 5(1) = 8 \neq 15$, so $(1, 1)$ is not on the line $3x + 5y = 15$.
10. $2(2) + 3(-1) = 1$, so $(-1, 2)$ is on the line $2y + 3x = 1$.
11. $\frac{1}{3}(5) - \frac{1}{5}(3) = \frac{5}{3} - \frac{3}{5} = \frac{16}{15} \neq -1$, so $(5, 3)$ is not on the line $\frac{1}{3}x - \frac{1}{5}y = -1$.
12. $\frac{3}{4}(4) + 3 = 6 = 3 + 3 = 6$, so $(4, 6)$ is on the line $y = \frac{3}{4}x + 3$.
13. The slope of a vertical line is undefined.
14. The slope of a horizontal line is always 0.

3.2 Graphing Linear Functions

Now that we know how to obtain the equation of a linear function, there are many things we could do with the function. We could determine if points are on the line. We could use the line to solve a word problem. Here we will concern ourselves with plotting linear functions. There will be three methods illustrated in this section.

Note: If you are going to plot any line by hand, it is highly recommended that you use graph paper and a ruler to make the lines straight and the points in the correct positions.

3.2.1 Plotting Using a t -chart

The first method for plotting a function is by finding some points that satisfy the equation and using those points to determine the shape of the curve. In order to do so, we need to select some x values and determine the corresponding y values. The common way to keep track of these points is by using a t -chart.

Example 3.2.1 Plot the linear function $y = 2x + 3$.

Solution The first thing we need to do is set up a t -chart.

x	$y = 2x + 3$
-3	
-2	
-1	
0	
1	
2	
3	

Now we need to find the value of the function at each of the points in the chart. We do so by substituting each of the x -values into the original equation.

$$x = -3: y = 2(-3) + 3 = -3$$

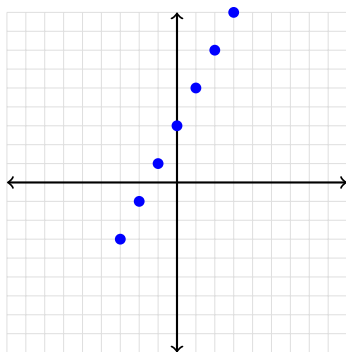
$$x = -2: y = 2(-2) + 3 = -1$$

$$x = -1: y = 2(-1) + 3 = 1$$

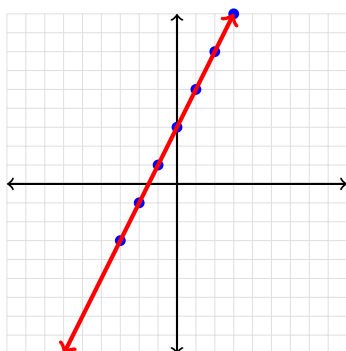
If we continue in this manner then the chart will give us seven points that are on the line $y = 2x + 3$.

x	$y = 2x + 3$
-3	-3
-2	-1
-1	1
0	3
1	5
2	7
3	9

Once we have these points, we can plot them



and then draw the straight line that passes through the points.



Do we really need this many points to plot a linear function? To answer this, we need to ask ourselves how many points are necessary to determine a straight line. The answer is two.

Example 3.2.2 Plot the linear function $2x - 3y = -1$.

Since this equation is in standard form, it will be easier for us if we convert it to slope-intercept form. We do not have to do this, since we could substitute x values into the equation in standard form, but then we will have to perform a series of algebraic steps to solve for the y value. For convenience sake, we will write the equation first in slope-intercept form.

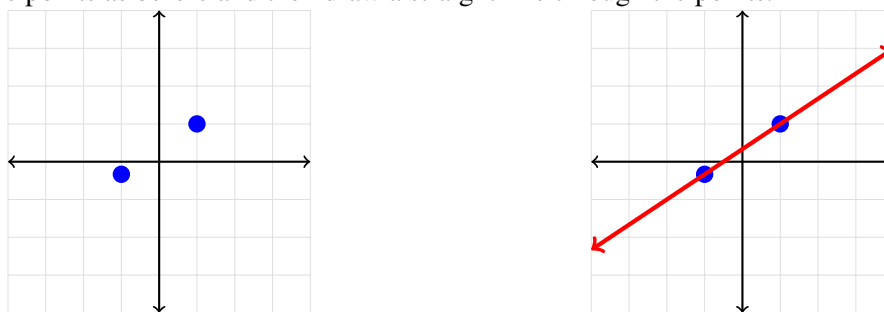
Solution We put the equation into slope-intercept form by using algebra, as we did in the last section.

$$\begin{aligned}
 2x - 3y &= -1 \\
 -3y &= -2x - 1 \\
 y &= \frac{-2}{-3}x - \frac{1}{-3} \\
 y &= \frac{2}{3}x + \frac{1}{3}
 \end{aligned}$$

Now we find the value of the function for at least two x values. We will do so for $x = 1$ and $x = -1$.

x	$y = \frac{2}{3}x + \frac{1}{3}$
-1	$-\frac{1}{3}$
1	1

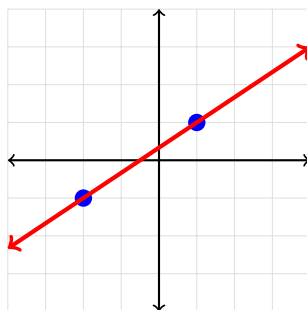
Now we plot the points as before and then draw a straight line through the points.



When we are trying to plot whole numbers, we can get a pretty accurate graph. But when we are forced to plot fractions, we have to estimate where the points should be, even when we are using graph paper, and this can give us a representation that is not accurate. We have to keep this in mind when drawing graphs so that when we are making the graphs to help us solve a problem, we do not rely solely on the picture. One way to aid in this is to make good choices about what to select for x values. In our last example, using $x = 1$ worked fine, but $x = -1$ did not give us an integer solution. With a little trial and error, we see that if we chose $x = -2$ instead, we would have

$$y = \frac{2}{3}(-2) + \frac{1}{3} = -\frac{4}{3} + \frac{1}{3} = -1$$

which is much easier to precisely plot.



3.2.2 Plotting Using Slope-Intercept Form

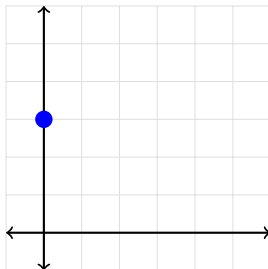
Another way we can plot a linear function is to use the values from the slope-intercept form and directly plot the function. We know two important things when the function is in the form; we know the slope and the y -intercept. This method has three steps.

1. Plot the y -intercept $(0, b)$.
2. Plot a second point using the slope and thinking of it as $\frac{\text{rise}}{\text{run}}$. We will, as a matter of convention, always consider a negative in the slope to be part of the numerator. This means that a negative slope will have a 'downward rise'. The run will always be to the right.
3. Draw a straight line that passes through these two points.

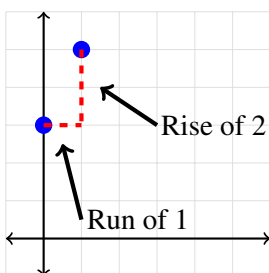
Example 3.2.3 Graph the linear function $y = 2x + 3$.

Solution We know from having the equation in slope intercept form that the y-intercept is the point (0,3) and the slope is $m = 2$.

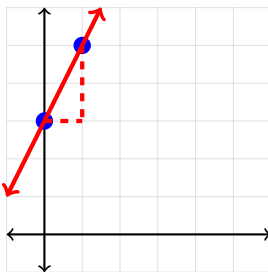
1. Plot the y-intercept (0,3).



2. Plot a second point by using the slope but think of it as a 'rise of 2' and a 'run of 1'.



3. Draw a straight line passing through the two points.

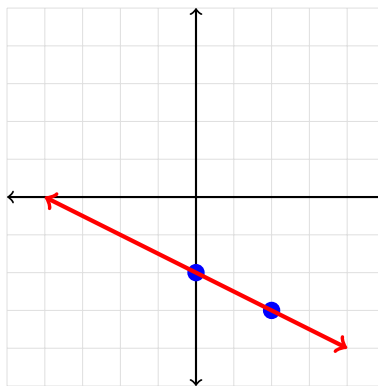


Example 3.2.4 Graph the linear function $-x - 2y = 4$.

Solution We first put the equation in slope-intercept form and then we can graph the line as we did in the last example.

$$\begin{aligned} -x - 2y &= 4 \\ -2y &= x + 4 \\ y &= -\frac{1}{2}x - 2 \end{aligned}$$

Now we can plot the point (0, -2) and then 'rise -1' and 'run 2' to the right.

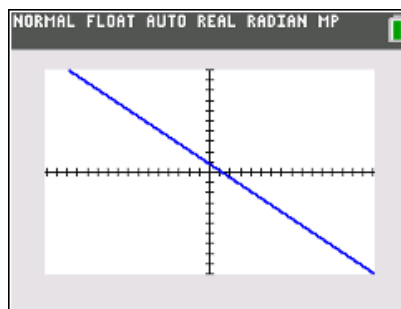
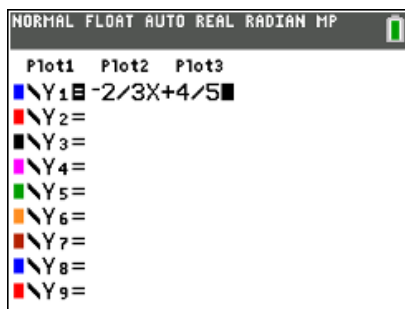


3.2.3 Using the TI-Series Calculator

We can also use the calculator to graph out linear functions. To do this, we first press the **Y=** key. This is found on the top right of the calculator, right below the screen. Then, on the line Y_1 , type the function, being sure to use standard order of operations when applicable. Once that is done, press the **GRAPH** key and your line will appear on the screen.

Example 3.2.5 Graph the linear function $y = -\frac{2}{3}x + \frac{4}{5}$.

Solution Here is how we can plot this line using the calculator.



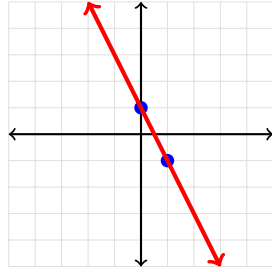
The negative here is not the subtraction key; if you try to use it you will get a syntax error. Instead, for a negative in front like this, use the $(-)$ in the bottom row.

3.2.4 Exercises

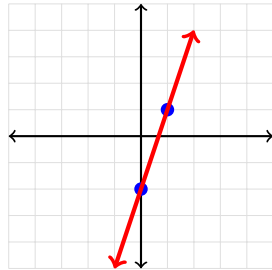
1. Plot $y = -2x + 1$ using a t -chart.
2. Plot $y = 3x - 2$ using a t -chart.
3. Plot $y = \frac{1}{2}x + 2$ using a t -chart.
4. Plot $y = -x - 1$ using a t -chart.
5. Plot $y = -3x - 4$ by using the slope and y -intercept.
6. Plot $-3y = x$ by using the slope and y -intercept.
7. Plot $y = \frac{5}{2}x + \frac{5}{2}$ by using the slope and y -intercept.
8. Plot $y = -2$ by using the slope and y -intercept.
9. Plot $y = -\frac{2}{9}x + \frac{11}{13}$ using the calculator.
10. Plot $y = -x + \frac{37}{50}$ using the calculator.
11. Plot $y = .27x - .527$ using the calculator.
12. Plot $y = .2079 - .11x$ using the calculator.

3.2.5 Solutions

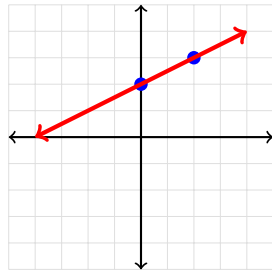
1.
$$\begin{array}{r|l} x & -2x+1 \\ \hline 0 & 1 \\ 1 & -1 \end{array}$$



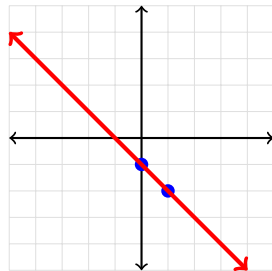
2.
$$\begin{array}{r|l} x & 3x-2 \\ \hline 0 & -2 \\ 1 & 1 \end{array}$$



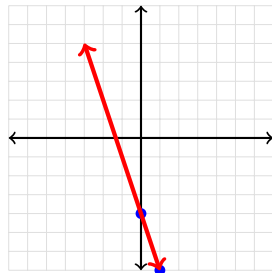
3.
$$\begin{array}{r|l} x & \frac{1}{2}x+2 \\ \hline 0 & 2 \\ 2 & 3 \end{array}$$



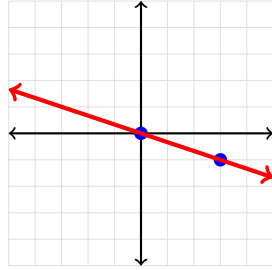
4.
$$\begin{array}{r|l} x & -x-1 \\ \hline 0 & -1 \\ 1 & -2 \end{array}$$



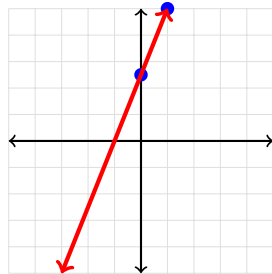
5. Since we are in slope-intercept form, we can see that the slope $m = -3$ and the y -intercept is $(0, -4)$.



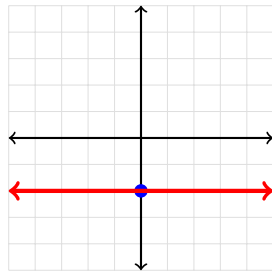
- Rewriting, we get $y = -\frac{1}{3}x$,
 6. which has a slope $m = -\frac{1}{3}$ and a y-intercept of the origin, $(0,0)$.



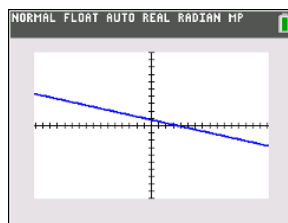
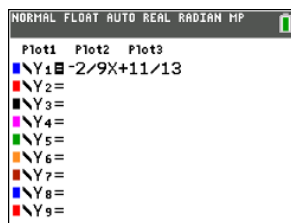
- Since we are in slope-intercept form, we see the slope is $m = \frac{5}{2}$ and the y-intercept is $(\frac{5}{2})$.



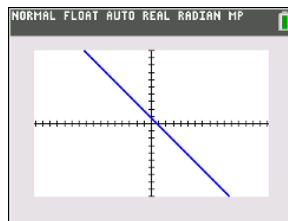
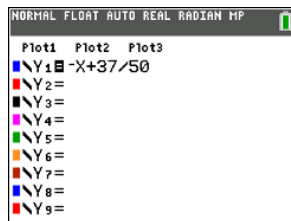
- This is a horizontal line, so the
 8. slope is $m = 0$ and the y-intercept is $(0, -2)$.



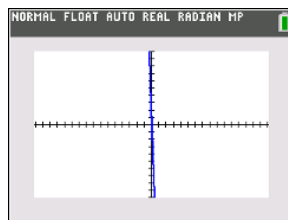
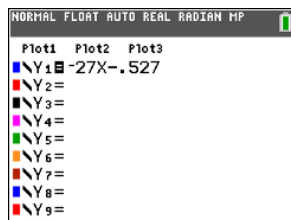
9.



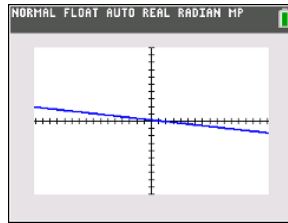
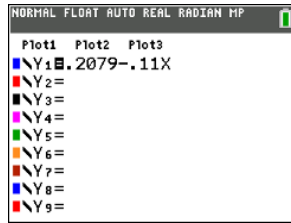
10.



11.



12.



3.3 Finding the Intersection Point of Linear Functions

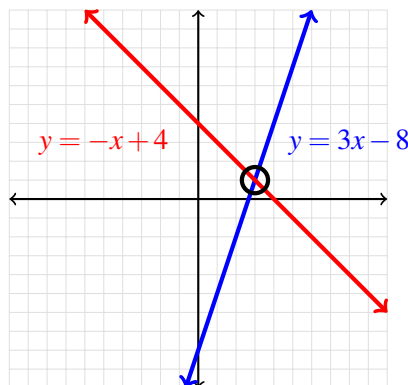
Now that we know how to find the equation of and how to graph a linear function, what are we to use this for? Before we can get to applications, we need to learn how to find the intersection of a pair of lines. Often, the intersection will be the break even point or will be a point of optimization.

3.3.1 How to Find the Intersection

The essential question here is this: if we have a pair of linear functions, how can we find the intersection point. First note that there is at most one point at which a pair of non-identical lines can intersect because the lines are straight. So we are looking for a point that is common to both lines.

Example 3.3.1 Find the intersection point of the linear functions $y = -x + 4$ and $y = 3x - 8$.

Solution If we graph the two functions together, we can get an idea of what the intersection point will be. To graph two functions together, we use any of the processes discussed earlier in this chapter for the first equation and then repeat the process for the second equation and graph both lines on the same set of axes. When plotting them, we should get



Using the graph, we can approximate the point at which the lines intersect. In a case such as this when the x and y values are both integers we can get a pretty good estimate of the point, which is $(3, 1)$. But if this point does not have integer values, then estimating will not give us a good enough answer. Especially when we are dealing with real-life applications, approximations are not good enough. We need to get as close as we can to the exact value. To do this we solve by setting the lines equal to each other.

Why does this work? If we want to find the intersection point of two lines, the point necessarily will be the only ordered pair that both lines share. In other words, the x and y values will have to be the same at that point and that point only. So, if the equations are $y = m_1x + b_1$ and $y = m_2x + b_2$, then we can find what we are looking for by solving for x algebraically in the equation $m_1x + b_1 = m_2x + b_2$.

Note: If the given lines are in standard form and we want to solve using this method, we first have to solve for one of the variables. Where it is not wrong to write the equations in the form $x =$ and solve, it is more customary to write the equations in the form $y =$ to solve. The reason is because the equations will be in

slope-intercept form so that we have them in a convenient way to plot them.

Here, we have $y = -x + 4$ and $y = 3x - 8$ so we want to know when $-x + 4 = 3x - 8$.

$$\begin{aligned} -x + 4 &= 3x - 8 \\ 4x + 4 &= -8 \\ 4x &= 12 \\ x &= 3 \end{aligned}$$

As we expected, we got the same $x = 3$. But now we need to find the corresponding y value. It does not matter which equation we use since the y value will be the same for both.

$$\begin{aligned} y &= -x + 4 \\ y &= -3 + 4 \\ y &= 1 \end{aligned}$$

Again, as we expected, we got the y value we were looking for and an intersection point of $(3, 1)$.

Another way we can ask this same question is by asking for the solution to a **system of equations**. A **system of equations** is simply multiple equations for which we want to find common points between the lines in question. This common point will be called a **solution**.

Example 3.3.2 Find a solution to the given system of equations.

$$\begin{cases} y = 3x + 2 \\ y = -2x + 1 \end{cases}$$

Solution The solution to this system is found by setting the equations equal to each other, as we did before.

$$\begin{aligned} 3x + 2 &= -2x + 1 \\ 5x + 2 &= -1 \\ 5x &= -1 \\ x &= -\frac{1}{5} \end{aligned}$$

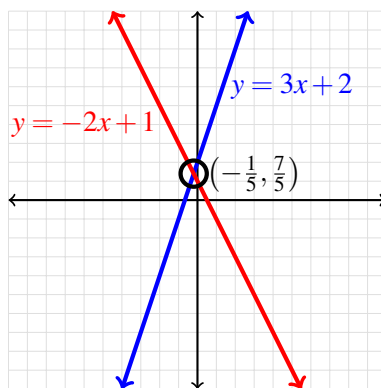
Now we need to find the corresponding y value.

$$\begin{aligned} y &= 2x + 1 \\ y &= -2\left(-\frac{1}{5}\right) + 1 \\ y &= \frac{2}{5} + 1 \end{aligned}$$

We need a common denominator, so ...

$$\begin{aligned} y &= \frac{2}{5} + \frac{5}{5} \\ y &= \frac{7}{5} \end{aligned}$$

The solution to this system is the point $\left(-\frac{1}{5}, \frac{7}{5}\right)$



3.3.2 Do We Have to Have an Intersection?

We stated earlier that if distinct lines intersect, they will do so at exactly one point. But this is not the only thing that can happen, as we will see in the next two examples.

Example 3.3.3 Find a solution to the given system of equations.

$$\begin{cases} y = 3x + 1 \\ y = 3x + 2 \end{cases}$$

Solution When we set these two equations equal to each other, we get

$$\begin{aligned} 3x + 1 &= 3x + 2 \\ 1 &\neq 2 \end{aligned}$$

This is certainly impossible and what it really means is that the two lines do not intersect, which in turn means that the lines are parallel. Notice that they have the same slope but a different y -intercept. Your answer here would be that there was no solution.

Example 3.3.4 Find a solution to the given system of equations.

$$\begin{cases} y = 3x + 1 \\ 2y = 6x + 2 \end{cases}$$

Solution When we divide each term of the second equation by 2 so that we have slope-intercept form, we see that the two lines are identical. This means that any point on one line will be on the other line, and since a line is comprised of an infinite number of points, we have an infinite number of solutions.

3.3.3 Exercises

For each of the exercises below, find the point of intersection of the sets of lines, if one exists.

1. $y = 2x + 1$ and $y = -x + 10$
2. $y = -3x + 2$ and $y = -x + 1$
3. $y = x + 1$ and $y = x + 2$
4. $y = \frac{1}{2}x + 1$ and $y = -\frac{3}{2}x + 3$
5. $y = .3x + 1.2$ and $y = -.2x + 3.7$
6. $y = \frac{1}{2}x$ and $y = \frac{2}{3}x$
7. $y = \frac{2}{3}x + 1$ and $y = \frac{3}{5}x - 2$
8. $x + y = 2$ and $x + 2y = 2$
9. $x - 2y = -6$ and $-\frac{1}{2}x + y = 3$
10. $-2x + y = 23$ and $-11x + y = 27$

3.3.4 Solutions

1. $y = 2x + 1$ and $y = -x + 10$

$$2x + 1 = -x + 10$$

$$3x = 9$$

$$x = 3$$

$$y = 2(3) + 1 = 7$$

So, the intersection is the point $(3, 7)$.

2. $y = -3x + 2$ and $y = -x + 1$

$$-3x + 2 = -x + 1$$

$$1 = 2x$$

$$\frac{1}{2} = x$$

$$y = -\frac{1}{2} + 1 = \frac{1}{2}$$

The intersection point is $(\frac{1}{2}, \frac{1}{2})$.

3. $y = x + 1$ and $y = x + 2$

Notice that these lines have the same slope but different y-intercepts. This means that they are parallel lines and therefore do not intersect.

4. $y = \frac{1}{2}x + 1$ and $y = -\frac{3}{2}x + 3$

$$\frac{1}{2}x + 1 = -\frac{3}{2}x + 3$$

$$x + 2 = -3x + 6$$

$$4x = 4$$

$$x = 1$$

$$y = \frac{1}{2}(1) + 1 = \frac{3}{2}$$

Therefore, the intersection point is $(1, \frac{3}{2})$.

5. $y = .3x + 1.2$ and $y = -.2x + 3.7$

$$.3x + 1.2 = -.2x + 3.7$$

$$3x + 12 = -2x + 37$$

$$5x = 25$$

$$x = 5$$

$$y = .3(5) + 1.2 = 2.7$$

The intersection point is therefore $(5, 2.7)$.

6. $y = \frac{1}{2}x$ and $y = \frac{2}{3}x$

Since both lines pass through the origin, the intersection point is $(0, 0)$.

7. $y = \frac{2}{3}x + 1$ and $y = \frac{3}{5}x - 2$

$$\begin{aligned} \frac{2}{3}x + 1 &= \frac{3}{5}x - 2 \\ 10x + 15 &= 9x - 30 \\ x &= -45 \\ y &= \frac{2}{3}(-45) + 1 = -29 \end{aligned}$$

The intersection point is $(-45, -29)$.

8. $x + y = 2$ and $x + 2y = 2$

$$\begin{array}{r} x + y = 2 \\ -(x + 2y = 2) \\ \hline -y = 0 \\ x + 0 = 2 \Rightarrow x = 2 \end{array}$$

This gives the intersection point at $(2, 0)$.

9. $x - 2y = -6$ and $-\frac{1}{2}x + y = 3$

$$\begin{array}{r} x - 2y = -6 \\ 2\left(-\frac{1}{2}x + y = 3\right) \\ \hline 0 + 0 = 0 \end{array}$$

This shows that the lines are the same, just written in different forms. So, if we let y be any value t , then $x = 2t - 6$, giving the ordered pair $(2t - 6, t)$ for any real number t .

10. $-2x + y = 23$ and $-11x + y = 27$

$$\begin{array}{r} -2x + y = 23 \\ -(-11x + y = 27) \\ \hline 9x = -4 \\ x = -\frac{4}{9} \\ -2\left(-\frac{4}{9}\right) + y = 23 \\ y = \frac{199}{9} \end{array}$$

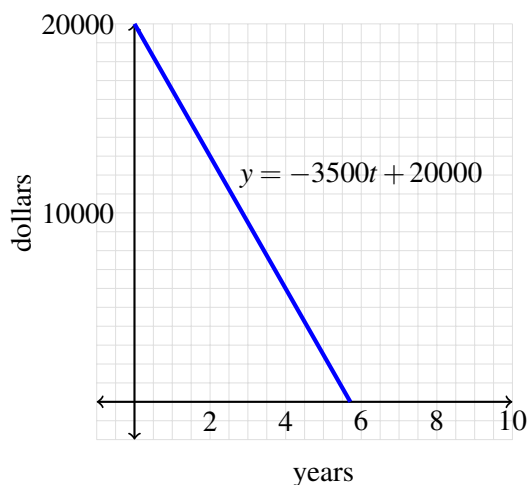
This gives the intersection point $\left(-\frac{4}{9}, \frac{199}{9}\right)$.

3.4 Applications of Linear Functions

3.4.1 Linear Depreciation

Example 3.4.1 Suppose that you buy a new car for \$20,000. If the car depreciates at a rate of \$3,500 per year, write a linear equation that represents this situation and use its graph to help interpret the x -intercept.

Solution If we let t represent the time in years then the car will be new at time $t = 0$. The price here is \$20,000, so the y -intercept of this function is $(0, 20000)$. Since the car depreciates at a rate of \$3,500 per year, the slope is $m = -3500$. So, the linear function that governs this situation would therefore be $V = -3500t + 20000$, where V is the value of the car in dollars after t years. If we plot this linear function, we get



We can see here that the line crosses the t -axis somewhere close to 6 years. If we solve this algebraically, we can see almost exactly where it crosses the axis.

$$\begin{aligned}0 &= -3500t + 20000 \\3500t &= 20000 \\t &= \frac{20000}{3500} \approx 5.71\end{aligned}$$

So, the t -intercept is $(5.71, 0)$. The amount of time it takes for the car to have no value is approximately 5.71 years.

3.4.2 Cost and Revenue

Example 3.4.2 A publisher quotes you a price of \$.25 per book to print your new manuscript with a fixed cost of \$250. How much will it cost to produce 10,000 copies of your book?

Solution Since the cost per book is constant, we see that this cost can be expressed as a linear function. This constant cost is our slope, so $m = .25$. And, the fixed cost is how much it will cost us to produce no books. This in practical terms is the y -intercept of our function and is the point $(0, 250)$. Therefore, $b = 250$.

Putting this together gives $C(x) = .25x + 250$ where x is the number of books to be printed.

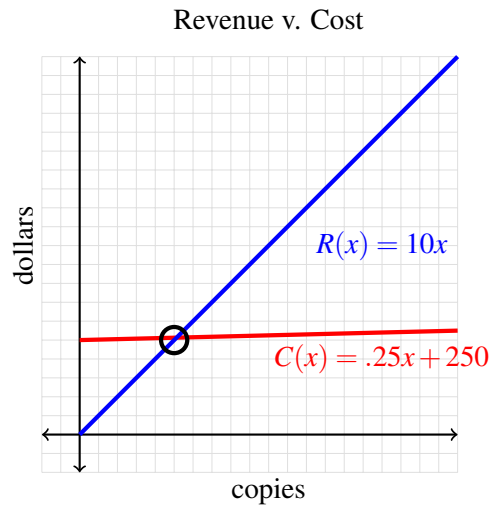
To find our cost for the order, we need to find $C(10000)$.

$$C(10000) = .25(10000) + 250 = \$2750.00$$

Example 3.4.3 *You plan to sell your book for \$10. How many copies would you have to sell before you broke even?*

Solution Since the price of the books will be the same for all of them, this is also a linear function and the slope will be 10. But, since you will make nothing if you sell no books, we have a y-intercept of $(0,0)$. This gives us a revenue function of $R(x) = 10x$ where x is the number of copies of the book sold.

To find when we would break even we set the revenue function equal to the cost function and solve for the quantity x .



$$\begin{aligned} 10x &= .25x + 250 \\ 9.75x &= 250 \\ x &= \frac{250}{9.75} \\ x &= 25.64 \end{aligned}$$

Since we cannot sell a fraction of a book, we would break even between 25 and 26 copies sold and then the rest would be profit.

3.4.3 Straight Line Depreciation

The straight line depreciation method simplifies how much we will lose per year in value by considering the loss to be a constant per year. This method is most useful when the economic benefits of the product are expected to be relatively even over the useful life of the product. For example, if you are making screen

printed t-shirts, the screener will roughly work as well on day one as it would at the end of its useful life. So we could find out what we would get as a salvage price when we sold it and calculate how long it would make sense to keep and operate the device.

Example 3.4.4 *Suppose you bought a screen printer for \$10000 and you know that you can sell it after 6 years of use for \$1000 (this is known as a residual value). What is the rate of depreciation?*

Solution To find the rate, we use the formula

$$\text{Depreciation per year} = \frac{\text{Cost} - \text{Residual Value}}{\text{Useful Life}}$$

So, here we have

$$\text{Depreciation per year} = \frac{10000 - 1000}{6} = \$1500$$

Now, we wanted to find the rate, so we have

$$\begin{aligned} \text{Depreciation per year} &= (\text{Cost} - \text{Residual Value}) \times \text{Rate of Depreciation} \\ 1500 &= 9000 \times \text{Rate of Depreciation} \\ \text{Rate of Depreciation} &= \frac{1}{6} \approx 16.67\% \end{aligned}$$

So, we lost roughly 16.67% of the value of the machine per year during the useful lifespan of the printer.

Example 3.4.5 *You are the CEO of a company that offers company cars to executives. Suppose that you buy a fleet of \$40,000 cars and intend for them to be kept for 3 years, at the end of which you will be able to sell those cars for \$10,000. What is the depreciation per year? What is the annual rate of depreciation? Assume that we can follow a linear model.*

Solution We will lose \$30,000 over the life of the investment in the cars, so over 3 years, this will come to \$10,000 per year using the linear model. To find the rate, we see

$$\begin{aligned} \text{Depreciation per year} &= (\text{Cost} - \text{Residual Value}) \times \text{Rate of Depreciation} \\ 10000 &= 30000 \times \text{Rate of Depreciation} \\ \text{Rate of Depreciation} &= \frac{1}{3} \approx 33.33\% \end{aligned}$$

3.4.4 Supply and Demand

When discussing supply and demand curves, we will use p for the y value and q for the x value. The p will represent price and q will represent quantity.

A **supply curve** is the equation that relates the price p that the manufacturer will sell the product for at quantity q . The greater the supply, the higher overall the price must be since it will cost the manufacturer more to make more items. The **demand curve** is the equation that relates the price p that the manufacturer must charge for a product when there are q units available. The greater the quantity the lower the price must be for the consumers to buy it. It is desirable for manufacturers to know what the ‘market price’ is. That is, they want to know how much to charge per item and what quantity to produce so that they will exactly cover

production costs. Companies use this information, along with other factors, to determine how many items to produce to maximize profit or to minimize costs. We will add in more parameters in the next chapter, but for now we will focus only on this break even point.

Example 3.4.6 *Suppose the supply curve for a particular product is given as $p = .002q + 1$, where q is in dollars. For this same product, the demand curve is given as $p = -.001q + 4$. We want to find what the market price should be.*

Solution The solution is found by solving the system of equations

$$\begin{cases} p = .002q + 1 \\ p = -.001q + 4 \end{cases}$$

To solve this, we do as we did in the previous examples and set the equations equal to each other.

$$\begin{aligned} .002q + 1 &= -.001q + 4 \\ .003q + 1 &= 4 \\ .003q &= 3 \\ q &= 1000 \end{aligned}$$

So, for this product, the quantity at which the supply and demand curves meet is 1000 units. To determine how much it will cost per item, we substitute this quantity into either curve.

$$\begin{aligned} p &= .002q + 1 \\ p &= .002(1000) + 1 \\ p &= 2 + 1 = 3 \end{aligned}$$

To produce 1000 items, the cost per item will be \$3.

3.4.5 Exercises

1. Suppose you bought a new iPhone for \$750 and the phone depreciates at a rate of \$35 per month. Use this information to write a linear equation modeling the situation and determine when you should buy a new phone so that your phone has some value at that point.
2. You buy a new car for \$25,000 and 2 years later, the car is valued at \$17,500. Assuming that the depreciation is linear, write an equation that models this situation and use that equation to find when the car is worth \$5,000.
3. You spend \$100 to print a short story you wrote. If you plan to sell it for \$4.99 per copy, how many copies would you need to sell in order to profit \$400?
4. It costs you \$.55 each to produce stickers to promote your new business and there is an associated \$75 set up fee. What is the cost of an order for 250 stickers? Write the equation that models the situation before solving.
5. You have done research on marketing your brand and have concluded that you need to pass out 1000 flyers at a cost of \$.15 each to reach your goal. The company you order from quotes you a price of \$160 to create and print your flyers. How much are they charging you to set up the design if there are no other costs besides setting up the design and the individual price per flyer?
6. Find the rate of depreciation of an investment that is at \$2,000 right now, but you invested \$5,000 four years ago.
7. Suppose the demand for a certain product is governed by the equation $p = -.05q + 11$ and the supply is governed by $p = .04q + 2$, where p represents the price in each case. Find the market price that will allow you to break even.
8. Your marketing research has shown you sell 1000 watches at \$50 each per month, but when the price is \$40, you sell 1500. Use this information to write the linear demand function for the sales of this watch.
9. Your store can sell 50 t-shirts of a particular style per week at a price of \$15 but only 30 per week at \$20. Your supplier is aware of your prices and decided that they will only sell you 25 shirts per week if you are going to charge \$15 but will sell you 60 per week if you are going to charge \$20. Find the associated linear supply and demand functions and find what price the shirts should be sold at so that there is neither a surplus or a shortage of them in stock.

3.4.6 Solutions

1. Using the given information, we see that the value of the phone is given by the function $V(t) = 750 - 35t$ where t is given in months since purchase. We want to know how many months to have the value be \$0 so that we can answer the question.

$$\begin{aligned}750 - 35t &= 0 \\35t &= 750 \\t &\approx 21.43 \text{ months}\end{aligned}$$

So, if we trade in at 21 months, there will be a small value to the phone. If we waited until 22 months, there would be no value.

2. Using two points, we can find the equation of a linear function.

$$m = \frac{17500 - 25000}{2 - 0} = -3750$$

Since one of the given points is the y -intercept, we can use that without further calculations to get the equation $V(t) = 25000 - 3750t$ where t is years since purchase. Now, we want to know when the car is worth \$5000.

$$\begin{aligned}25000 - 3750t &= 5000 \\3750t &= 20000 \\t &= 5.\bar{3}\end{aligned}$$

So, after 5 years, 4 months, the car is worth \$5000.

3. Since you spent \$100 to print the stories, you really want to make \$500 in order to profit \$400. That is, we want to know when $4.99x = 500$ where x is the number of copies sold.

$$\begin{aligned}4.99x &= 500 \\x &= 100.2\end{aligned}$$

Since we cannot sell part of a story, we would therefore need to sell 101 copies to make a profit of \$400.

4. Using the given information, we have $C(s) = 75 + .55s$ where s is the number of stickers produced. Then, $C(250) = 75 + .55(250) = \212.50 .
5. From the information given, we know $C(1000) = 125 = .15(1000) + s$, where s is the start up cost. Solving, we get

$$s = 125 - .15(1000) = 10$$

So, there is a \$10 set up charge.

6. We lose \$3000 over the life of the investment, so we have an annual depreciation of \$750/year. This gives

$$\begin{aligned}750 &= 3000 \times R \\R &= .25\end{aligned}$$

So, the rate of depreciation is 25%.

7. To find the market price, we find the intersection of these two curves.

$$\begin{aligned}-.05q + 11 &= .04q + 2 \\-.09q &= -9 \\q &= 100\end{aligned}$$

So, we would break even at 100 units.

8. We will use the points (50, 1000) and (40, 1500) to find the demand curve.

$$\begin{aligned}m &= \frac{1500 - 1000}{40 - 50} = -50 \\1000 &= -50(50) + b \\3500 &= b\end{aligned}$$

This gives the demand curve $p = 3500 - 50q$ where q is the number of units.

9. We first need the supply and demand equations. For the demand equations, we will use the points (15, 50) and (20, 30).

$$\begin{aligned}m &= \frac{50 - 30}{15 - 20} = -4 \\30 &= -4(20) + b \\b &= 110\end{aligned}$$

So, the demand equation is $q = -4p + 110$, where p is the price. Now, the supply curve. We will use the points (15, 25) and (20, 60).

$$\begin{aligned}m &= \frac{60 - 25}{20 - 15} = 7 \\25 &= 7(15) + b \\b &= -80\end{aligned}$$

This gives the supply curve as $q = 7p - 80$, where s is again the price.

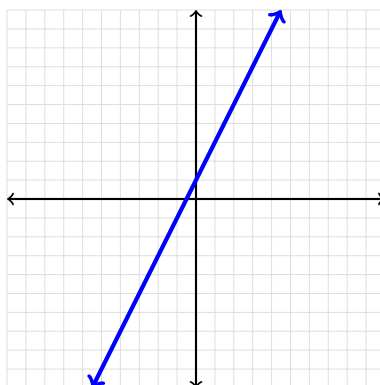
Now, to find equilibrium, we set these two equations equal to each other.

$$\begin{aligned}-4p + 110 &= 7p - 80 \\11p &= 190 \\p &= 17.27\end{aligned}$$

So, there will be no surplus or shortage if the shirts are sold for \$17.27 each.

3.5 Graphing Inequalities

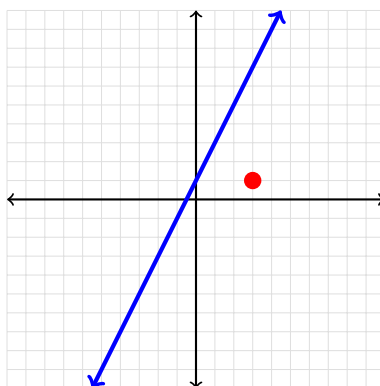
A **linear inequality** is similar to a linear function in that we plot the lines in the exact same manner. The difference is that instead of the line being the solution, the solution is the line and the region of the plane that satisfies the inequality. To see how this works, consider the inequality $y \leq 2x + 1$. We graph the line in the usual manner.



The difference comes in because of the inequality sign. We do not only want all points that satisfy $y = 2x + 1$, but also all points that are less and or equal to $2x + 1$. That is, if we choose a point, say $(3, 1)$, does the point satisfy the inequality? To see, we substitute into the inequality and see if it is true.

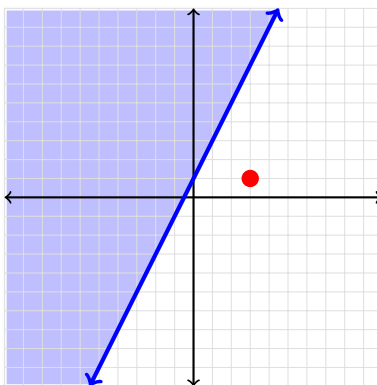
$$\begin{aligned} y &\leq 2x + 1 \\ 1 &\leq 2(3) + 1 \\ 1 &\leq 7 \end{aligned}$$

This is certainly true, so the point $(3, 1)$ is also part of the solution along with the line $y = 2x + 1$.



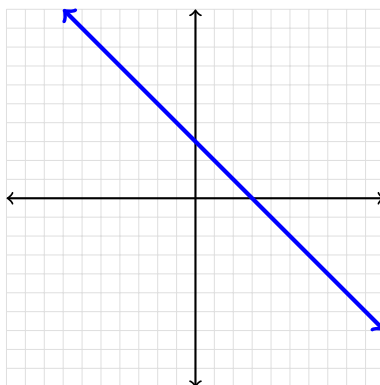
If we want all of the solution points, it would be impossible to do this for all possible points that would satisfy the equation. But all points that satisfy the inequality will be on the same side of the line as the point $(3, 1)$. Now, in many texts, you are instructed to shade the side of the line that is the solution of the inequality. Here, however, we will shade the side of the plane that is not the solution. The reason is that when we are looking for the solution set for a number of inequalities, it is sometimes too difficult to see what

the solution is because we would need to look at the intersection of all of the shadings. With the calculator, this can be produced without confusion but on paper it is extremely difficult. So, we will shade the opposite side, as in the graph below.



Example 3.5.1 Graph the inequality $y \geq -x + 3$.

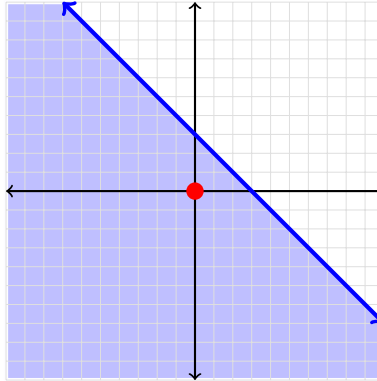
Solution We first plot the line $y = -x + 3$.



Then we want to shade the side that is not part of the solution. In order to determine which side of the line to shade, we can use a ‘test point’; that is, we can pick any point we want and see if it satisfies the inequality. If the point makes the inequality false then shade that side. If the point makes the inequality true then we know not to shade the side that the point is on. The only problem with this method is if the point we choose is on the line. The inequality will be true, but we still will not know which side to shade. An easy point to use to test is the origin $(0,0)$ as long as this is not the y -intercept. (If it is then select another point - I suggest $(0,1)$ or $(1,0)$).

$$\begin{aligned} y &\geq -x + 3 \\ 0 &\geq -(0) + 3 \\ 0 &\geq 3 \end{aligned}$$

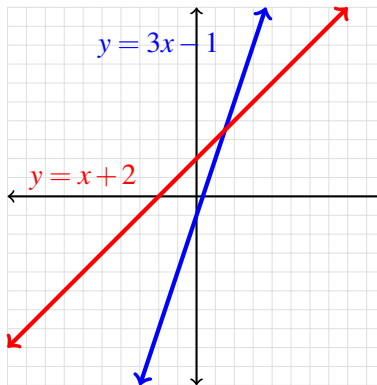
This is clearly not true, so we know the side to shade is the side containing the origin because we are shading the side that is not true for our purposes.



Example 3.5.2 Graph the system of linear inequalities.

$$\begin{cases} y \geq 3x - 1 \\ y \leq x + 2 \end{cases}$$

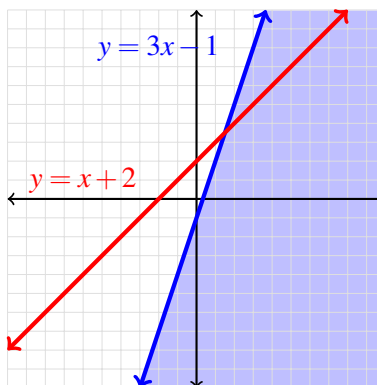
Solution First we graph the two lines as before.



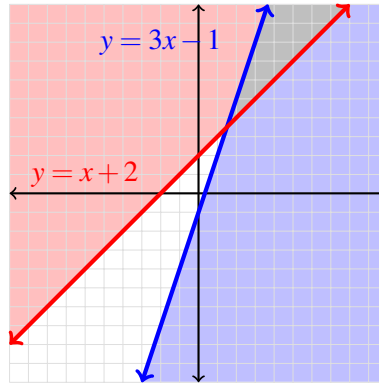
Now we need to decide which side to shade. Since the origin is not on either of the lines, we can use that as our test point for both of them.

$$\begin{aligned} y \leq x + 2 &\Rightarrow 0 \leq 0 + 2 \Rightarrow 0 \leq 2 \\ y \geq 3x - 1 &\Rightarrow 0 \geq 3(0) - 1 \Rightarrow 0 \geq -1 \end{aligned}$$

Both of these are true. So, in both cases, we will shade the region that does not contain the origin. In the case of $y \geq 3x - 1$, we shade below the line.



In the case of $y \leq x + 2$, we shade above the line.



The solution to the system of inequalities is therefore the region that is not shaded. This is the region that simultaneously satisfies both of the inequalities. We call this set of points that satisfies all of the inequalities at the same time the **feasible set**.

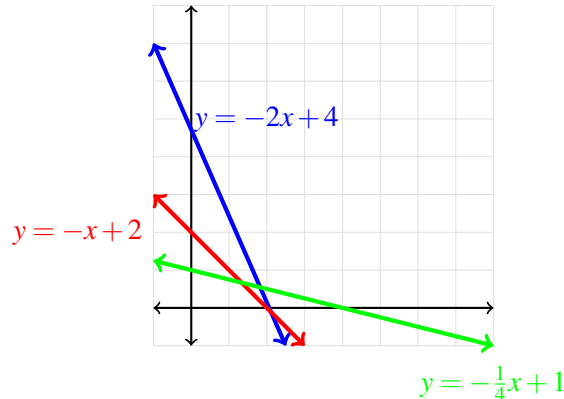
Example 3.5.3 Graph the system of linear inequalities.

$$\begin{cases} 2x + y \geq 4 \\ x + y \geq 2 \\ x + 4y \geq 4 \\ x \geq 0, y \geq 0 \end{cases}$$

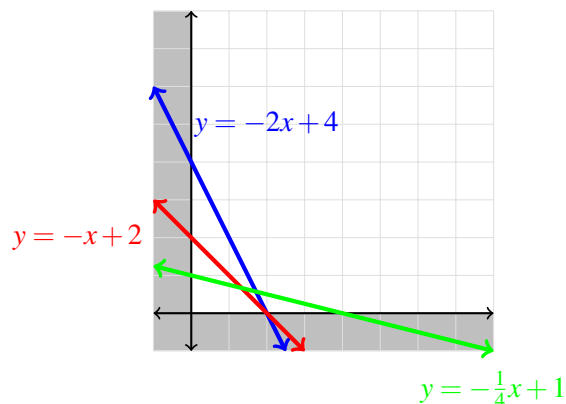
Solution Notice two things that are different in this example than before. First, the inequalities are in standard form. Before we can plot the lines, we need to put each of the first three into slope-intercept form. The second thing that is different is the last two constraints. In practical applications, it is not realistic to have negative quantities or prices, so we often restrict our x and y values to be nonnegative.

$$\begin{cases} 2x + y \geq 4 \\ x + y \geq 2 \\ x + 4y \geq 4 \\ x \geq 0, y \geq 0 \end{cases} \Rightarrow \begin{cases} y \geq -2x + 4 \\ y \geq -x + 2 \\ y \geq -\frac{1}{4}x + 1 \\ x \geq 0, y \geq 0 \end{cases}$$

Now we want to plot the three linear functions associated with the linear inequalities, as before.



First, note that we have constraints that x and y have to be nonnegative. This means that the feasible set must on;y include points in the first quadrant or on the x or y axes.



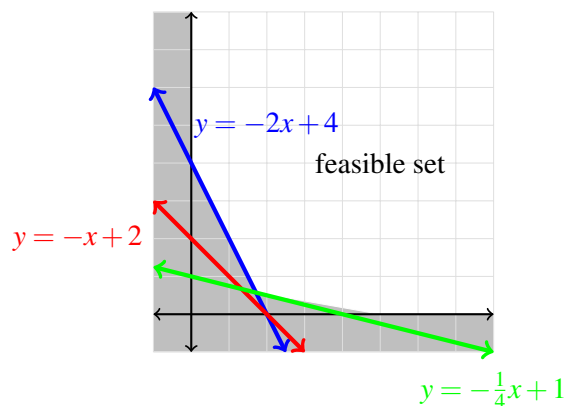
Now we use a test point to see which side of the lines to shade. Since the origin is not on any of the lines, we can use the point $(0,0)$ as the test point.

$$0 \geq -2(0) + 4 \Rightarrow 0 \geq 4$$

$$0 \geq -(0) + 2 \Rightarrow 0 \geq 2$$

$$0 \geq -\frac{1}{4}(0) + 1 \Rightarrow 0 \geq 1$$

Since these are all false, we shade the region that includes the point $(0,0)$.



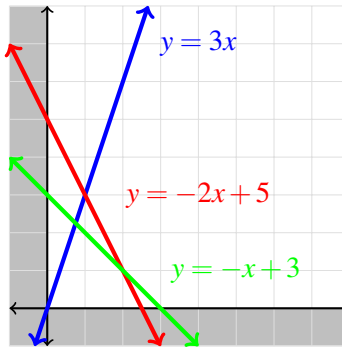
Example 3.5.4 Graph the system of linear inequalities.

$$\begin{cases} -3x + y \leq 0 \\ 2x + y \leq 5 \\ x + y \leq 3 \\ x \geq 0, y \geq 0 \end{cases}$$

Solution Just like before, we need to start by putting the inequalities in slope-intercept form.

$$\begin{cases} -3x + y \leq 0 \\ 2x + y \leq 5 \\ x + y \leq 3 \\ x \geq 0, y \geq 0 \end{cases} \Rightarrow \begin{cases} y \leq 3x \\ y \leq -2x + 5 \\ y \leq -x + 3 \\ x \geq 0, y \geq 0 \end{cases}$$

Now we plot the linear equations related to these linear inequalities. Notice again that we are restricted to the first quadrant since both x and y are nonnegative.

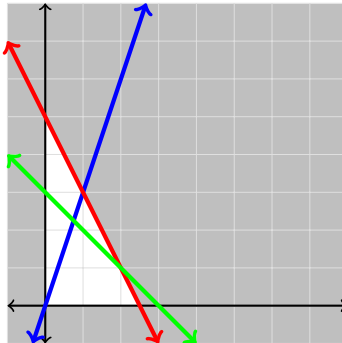


Now we use a test point to see which side of each line we need to shade. First we will check the second and third inequalities using the point $(0,0)$.

$$0 \leq -2(0) + 5 \Rightarrow 0 \leq 5$$

$$0 \leq -(0) + 3 \Rightarrow 0 \leq 3$$

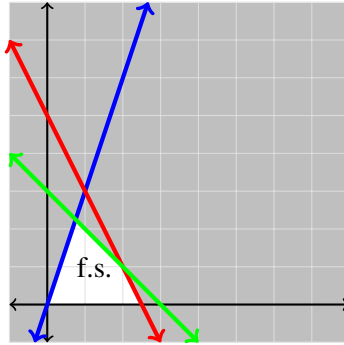
So, in both cases we shade the region that does not include the origin since both are true and we want to shade the part of the plane that does not satisfy the inequality.



We cannot use the origin to test the other inequality, however, because the line $y = 3x$ passes through the origin. We can use any point that is not on the line $y = 3x$. Since we want to use an 'easy' point, we will use $(1,0)$.

$$y \leq 3x \Rightarrow 0 \leq 3(1) \Rightarrow 0 \leq 3$$

To finish off the problem we shade the side of the line $y = 3x$ that does not contain the point $(1, 0)$.



3.5.1 Exercises

1. Graph $y \leq 3x + 2$.
2. Graph $\frac{1}{2}x - \frac{1}{5}y \leq 1$.
3. Graph $x \leq 2y$.
4. Graph $x \geq -3$.
5. Graph the feasible set for $y \leq 3x - 6$.

6. Graph the feasible set for

$$\begin{cases} x + 2y \geq 12 \\ 2x + y \geq 3 \end{cases}$$

7. Graph the feasible set for

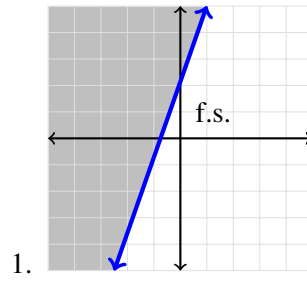
$$\begin{cases} x + y \geq 4 \\ x + 2y \geq 6 \\ x \geq 1 \end{cases}$$

8. Graph the feasible set for

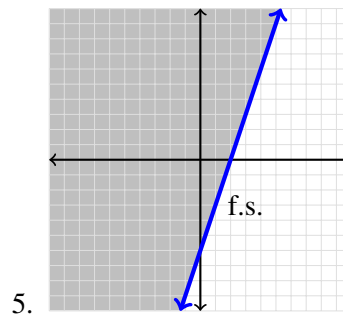
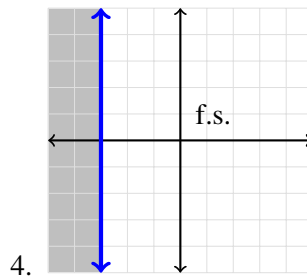
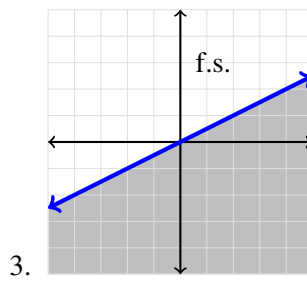
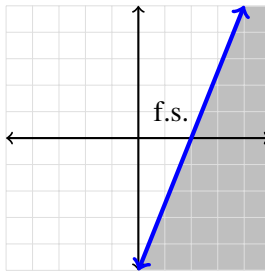
$$\begin{cases} x + 6y \leq 12 \\ x + y \leq 9 \\ x \geq 0, y \geq 0 \end{cases}$$

9. Determine if $(2, 4)$ is above or below $y = 3x + 7$.
10. Determine if $(7, 1)$ is above or below $x = 5y + 3$.
11. Determine if $(3, 2)$ is above or below $5x - 7y = 3$.
12. Determine if $(-2, 3)$ is above or below $10 + 3x + 4y = 1$.
13. What is meant by the term “feasible set”?

3.5.2 Solutions

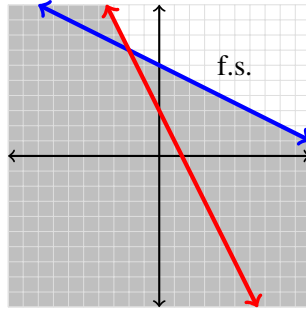


2. Rewriting, we get $y \geq \frac{5}{2}x - 5$.



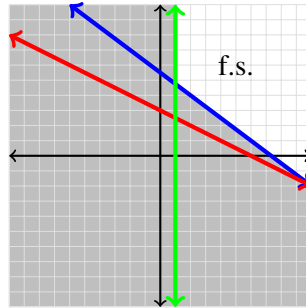
6.

$$\begin{cases} y \geq -\frac{1}{2}x + 6 \\ y \geq -2x + 3 \end{cases}$$



7.

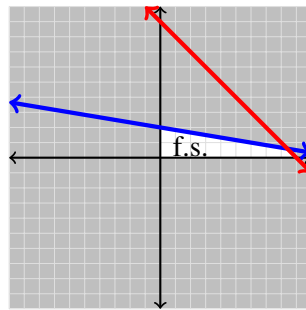
$$\begin{cases} y \geq -x + 4 \\ y \geq -\frac{1}{2}x + 3 \\ x \geq 1 \end{cases}$$



8. Graph the feasible set for

$$\begin{cases} x + 6y \leq 12 \\ x + y \leq 9 \\ x \geq 0, y \geq 0 \end{cases}$$

$$\begin{cases} y \leq -\frac{1}{6}x + 2 \\ y \leq -x + 9 \\ x \geq 0, y \geq 0 \end{cases}$$



9. Since $3(2) + 7 = 13 > 4$, the point $(2, 4)$ is below the line.

10. If we put this line in slope-intercept form, we get $y = \frac{1}{5}x - \frac{3}{5}$. Then, $\frac{1}{5}(7) - \frac{3}{5} = \frac{4}{5} < 1$ and so the point is above the line.

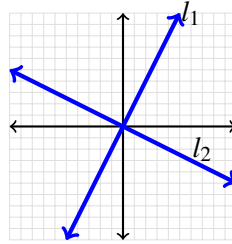
11. Putting the equation in slope-intercept form gives $y = \frac{5}{7}x - \frac{3}{7}$. Then, $\frac{5}{7}(3) - \frac{3}{7} = \frac{12}{7} < 2$. Therefore, the point is above the line.

12. Rewriting, we get $y = -\frac{3}{4}x - \frac{9}{4}$. Then, we have $-\frac{3}{4}(-2) - \frac{9}{4} = -\frac{3}{4} < 3$. So, the point is above the line.

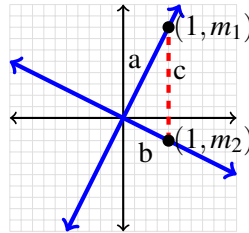
13. The feasible set is the set of all points that satisfy the system of inequalities.

3.6 Justification of Slopes of Perpendicular Lines

We know that the slopes of perpendicular lines are negative reciprocals. But why is this so? Suppose we have two perpendicular lines that intersect at the origin.



At the point where $x = 1$, each line has risen (or fallen) exactly the value of the slope. So, at $x = 1$, the y -coordinate of l_1 is m_1 and the y -coordinate of l_2 is m_2 , where m_1 and m_2 are the slopes of lines l_1 and l_2 , respectively.



These two lines are only perpendicular if the triangle above satisfies the Pythagorean theorem, which says that in any right triangle the sum of the squares of the legs is equal to the square of the hypotenuse. We generally express this as $a^2 + b^2 = c^2$.

In order to apply this, we need to find the distance between the vertices of the triangle and then substitute those values into the relation above.

$$\text{Distance between } (0,0) \text{ and } (1,m_1): a^2 = \left(\sqrt{(1-0)^2 + (m_1-0)^2}\right)^2 = 1 + m_1^2$$

$$\text{Distance between } (0,0) \text{ and } (1,m_2): b^2 = \left(\sqrt{(1-0)^2 + (m_2-0)^2}\right)^2 = 1 + m_2^2$$

$$\text{Distance between } (1,m_2) \text{ and } (1,m_1): c^2 = \left(\sqrt{(1-1)^2 + (m_2-m_1)^2}\right)^2 = (m_2 - m_1)^2$$

This gives the following:

$$\begin{aligned} a^2 + b^2 &= c^2 \\ 1 + m_1^2 + 1 + m_2^2 &= (m_2 - m_1)^2 \\ 2 + m_1^2 + m_2^2 &= m_2^2 - 2m_1m_2 + m_1^2 \\ 2 &= -2m_1m_2 \\ -\frac{1}{m_1} &= m_2 \end{aligned}$$

That is, the slopes are negative reciprocals of each other.

Chapter 4

Applications of Linear Functions

4.1 Linear Programming: The Set-up

4.1.1 What Is Linear Programming?

When we are faced with real life problems, we can use the method of linear programming to find the minimal cost or maximal output. Simply put, linear programming is a method of problem solving. We will solve word problems with two variables graphically using linear programming.

We will begin in this section by learning how to extract the information we need from the word problem and take the problem through to the graph. In the next section, we will look at how to take the problem from the beginning through the graph to the optimal solution.

4.1.2 What Are We Looking For?

When reading a problem, we are looking for information that gives us restrictions on such things as time, hours and the amount of raw materials we have at our disposal. We are also looking for the **object function**; that is, the expression that we want to optimize. The wording of the problem will tell us if we are looking to maximize or minimize the object function. Let's begin with a word problem. While doing this problem, we will number all of the steps we need to solve the problem. Remember, though, that we will just be setting up the problem here.

Example 4.1.1 *The most popular candy bar sold by the Lampes Candy Company combines peanuts and almonds. For a given month, at least 20 tons of nuts are needed to make enough candy bars to fill all of the orders. The company is tinkering with the formula, but they know that they want to have at least twice as many almonds as peanuts. Peanuts cost \$225 per ton and almonds cost \$325 per ton. How much of each should be used to minimize cost?*

Solution We will set this problem up in steps.

1. Make sure you know what you are looking for.

We are looking to minimize the cost.

2. Define the variables.

Let x = the amount of peanuts.

Let y = the amount of almonds.

3. Write out all constraints.

The amount of nuts we need is at least 20 tons, and we know we want at least twice as many almonds as peanuts (by weight), so we have

$$x + y \geq 20$$

$$y \geq 2x$$

We also have the constraints that we cannot have a negative amount of either type of nut, so we know that $x \geq 0$ and $y \geq 0$.

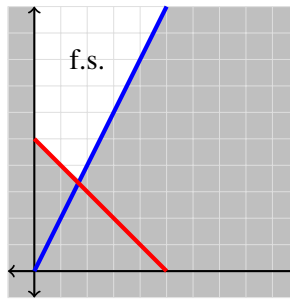
4. **Write out the object function; that is, the function to be optimized.**

We want to minimize cost and we are given the cost of the nuts per ton, so we have $C = 225x + 325y$. This equation is separate from the constraints, so it will not be graphed with the others. It might be a good idea to isolate it by boxing it or putting a star next to it.

5. **Put all of the constraint equations in slope-intercept form.**

$$\begin{cases} x + y \geq 20 \\ y \geq 2x \\ x \geq 0, y \geq 0 \end{cases} \implies \begin{cases} y \geq -x + 20 \\ y \geq 2x \\ x \geq 0, y \geq 0 \end{cases}$$

6. **Graph the system of inequalities.**



Example 4.1.2 *The Lampes Candy Company has two providers for the cocoa plant, one foreign and one domestic. In order to keep the factory in operation, at least 4 tons of cocoa plant must be processed per week. The foreign plants cost \$200 per ton to process and the domestic plants cost \$120 per ton to process. Costs must be kept below \$960 per week. The FDA requires that the amount of domestic plants processed cannot exceed twice the amount of foreign plants. If the foreign cocoa plants yield 25 lbs of chocolate per ton and the domestic plants yield 30 pounds of chocolate per ton, how many tons of cocoa plants from each source should be processed each week in order to maximize the chocolate produced?*

Solution We will follow the same steps as in the last example.

1. **Make sure you know what you are looking for.**

We are looking to maximize the amount of chocolate produced.

2. **Define the variables.**

Let x = the number of tons of foreign cocoa plants.

Let y = number of tons of domestic cocoa plants.

3. **Write out all constraints.**

The factory must process at least 4 tons of cocoa plants per week, so we have $x + y \geq 4$. The total cost must be kept below \$960, so $200x + 120y \leq 960$. We also have a restriction on the amounts of foreign and domestic plants in relation to each other. This gives us the constraint $y \leq 2x$. We also have the constraints that we cannot have a negative amount of either type of plant, so we know that $x \geq 0$ and $y \geq 0$.

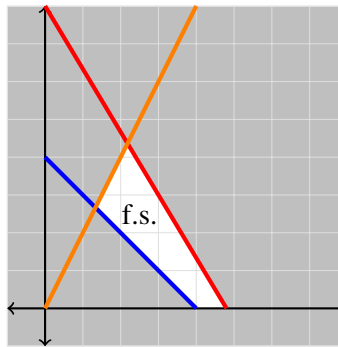
4. **Write out the object function; that is, the function to be optimized.**

We want to maximize the amount of chocolate produced, so the function we want to optimize is $P = 25x + 30y$.

5. **Put all of the constraint equations in slope-intercept form.**

$$\begin{cases} x + y \geq 4 \\ 200x + 120 \leq 960 \\ y \leq 2x \\ x \geq 0, y \geq 0 \end{cases} \implies \begin{cases} y \geq -x + 4 \\ y \leq -\frac{5}{3}x + 8 \\ y \leq 2x \\ x \geq 0, y \geq 0 \end{cases}$$

6. **Graph the system of inequalities.**



Example 4.1.3 A farmer has 12 acres on which to plant corn and wheat. He knows he wants to plant at least 8 acres of land but he has only \$2700 to spend and it costs \$200 per acre to plant corn and \$300 per acre to plant wheat. Plus, he only has 15 hours to get all of the planting done and it takes twice as long to plant corn as it does wheat. If he makes a profit of \$400 per acre of corn and \$250 per acre of wheat, how many acres of each should he plant in order to maximize profit?

Solution Once again, we will follow the same steps.

1. **Make sure you know what you are looking for.**

We are looking to maximize profit.

2. **Define the variables.**

Let x = the acres of corn.

Let y = the acres of wheat.

3. **Write out all constraints.**

The farmer has to plant at least 8 acres of land but he only has 12 acres on which to plant, so

$$\begin{aligned} x + y &\geq 8 \\ x + y &\leq 12 \end{aligned}$$

There is also a time constraint; he has only 15 hours to plant all of the crops and we know it takes twice as long to plant corn. This gives $2x + y \leq 15$. Finally, there is a cost associated with planting

each type of crop but there is a limit on the money available to plant. So, $200x + 300y \leq 2700$. We also have the constraints that we cannot have a negative amount of either type of crop, so we know that $x \geq 0$ and $y \geq 0$.

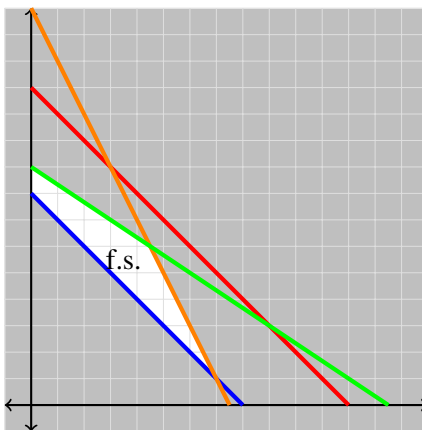
4. **Write out the object function; that is, the function to be optimized.**

We want to maximize profit, so the object function is $P = 400x + 250y$.

5. **Put all of the constraint equations in slope-intercept form.**

$$\begin{cases} x + y \geq 8 \\ x + y \leq 12 \\ 2x + y \leq 15 \\ 200x + 300y \leq 2700 \\ x \geq 0, y \geq 0 \end{cases} \implies \begin{cases} y \geq -x + 8 \\ y \leq -x + 12 \\ y \leq -2x + 15 \\ y \leq -\frac{2}{3}x + 9 \\ x \geq 0, y \geq 0 \end{cases}$$

6. **Graph the system of inequalities.**



4.1.3 Exercises

For problems 1-4, graph the system of inequalities to find the feasible set.

1.

$$\begin{cases} x + y \leq 4 \\ 4x + 2y \leq 12 \\ x \geq 0, y \geq 0 \end{cases}$$

2.

$$\begin{cases} \frac{3}{2}x + \frac{1}{2}y \geq 2 \\ 4x + 6y \geq 10 \\ x \geq 0, y \geq 0 \end{cases}$$

3.

$$\begin{cases} x + y \geq 1 \\ 2x + 3y \leq 6 \\ x \geq 0, y \geq 0 \end{cases}$$

4.

$$\begin{cases} 2x + 2y \leq 20 \\ 3x + 3y \geq 12 \\ 2x + 4y \leq 24 \\ x \geq 0, y \geq 0 \end{cases}$$

5. Explain why the given system has no solution.

$$\begin{cases} 2x + 4y \leq 24 \\ y \geq 7 \\ x \geq 0, y \geq 0 \end{cases}$$

6. The Lampes Shipping Company has two different sizes and weights for its shipping crates. Each crate of type 1 is 100 cubic feet and weighs 400 pounds, and each crate of type 2 is 20 cubic feet and weighs 720 pounds. The Lampes Shipping Company charges \$75 per crate of type 1 and \$100 per crate for type 2. The crates will be shipped by truck and each truck has a maximum load of 14,400 pounds and 2000 cubic feet.

(a) Fill in the accompanying chart.

	Type 1	Type 2	Truck Capacity
Volume			
Weight			
Charge			

(b) Let x be the number of crates of type 1 and let y be the number of crates of type 2 being shipped in one truck. Use the table to write the constraint equations and also any other constraints that there are.

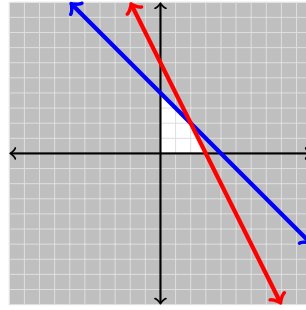
(c) Write the object function for this problem.

(d) Graph the feasible set for this problem.

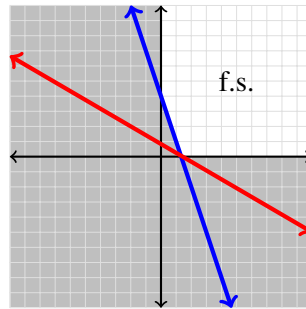
7. The Providence Grays Vintage Base Ball Club makes two kinds of bats: the Burlingame and the Hickory. They can turn up to 400 Burlingames and 500 Hickorys per year, but can only make 650 bats total per year. It takes 20 man hours to produce a Burlingame and 40 man hours to produce a Hickory and the team has 22,000 man hours per year to make bats. The profit on the bats is \$40 for a Burlingame and \$60 for a Hickory.
- (a) Write constraint equations for this problem, letting x represent the number of Burlingames and y represent the number of Hickorys.
 - (b) Write the profit function in terms of x and y .
 - (c) Graph the feasible set.

4.1.4 Solutions

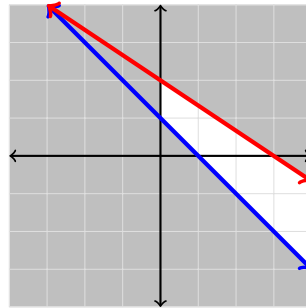
1.
$$\begin{cases} y \leq -x + 4 \\ y \leq -2x + 6 \\ x \geq 0, y \geq 0 \end{cases}$$



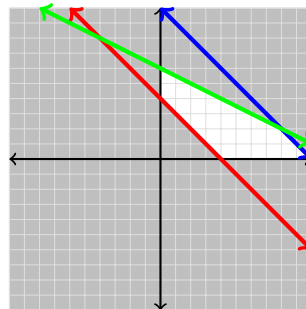
2.
$$\begin{cases} y \geq -3x + 4 \\ y \geq -\frac{2}{3}x + \frac{5}{3} \\ x \geq 0, y \geq 0 \end{cases}$$



3.
$$\begin{cases} y \geq -x + 1 \\ y \leq -\frac{2}{3}x + 2 \\ x \geq 0, y \geq 0 \end{cases}$$



4.
$$\begin{cases} y \leq -x + 10 \\ y \geq -x + 4 \\ y \leq -\frac{1}{2}x + 6 \\ x \geq 0, y \geq 0 \end{cases}$$



5. Explain why the given system has no solution.

$$\begin{cases} 2x + 4y \leq 24 \\ y \geq 7 \\ x \geq 0, y \geq 0 \end{cases}$$

When we rewrite the first equation in slope-intercept form, we get $y \leq -\frac{1}{2}x + 6$. We need to be less than this line for this inequality to be satisfied, and since $x \geq 0$, it must be that we use a non-negative

value for x . If we use $x = 0$, we get $y = 6$ and if we use $x > 0$, we get $y < 6$. But, we are also constrained that $y \geq 7$, so there are no points that can satisfy both of these inequalities at the same time while remaining in the first quadrant.

6. The Lampes Shipping Company has two different sizes and weights for its shipping crates. Each crate of type 1 is 100 cubic feet and weighs 400 pounds, and each crate of type 2 is 20 cubic feet and weighs 720 pounds. The Lampes Shipping Company charges \$75 per crate of type 1 and \$100 per crate for type 2. The crates will be shipped by truck and each truck has a maximum load of 14,400 pounds and 2000 cubic feet.

(a) Fill in the accompanying chart.

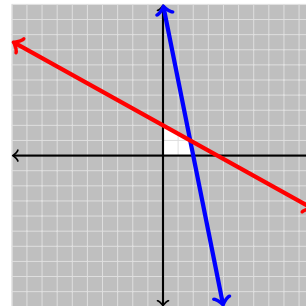
	Type 1	Type 2	Truck Capacity
Volume	100	20	2000
Weight	400	720	14400
Charge	\$75	\$100	

(b)

$$\begin{cases} 100x + 20y \leq 2000 \\ 400x + 720y \leq 14400 \\ x \geq 0, y \geq 0 \end{cases}$$

(c) $C = 75x + 100y$

(d)
$$\begin{cases} y \leq -5x + 100 \\ y \leq -\frac{5}{9}x + 20 \\ x \geq 0, y \geq 0 \end{cases}$$



7. The Providence Grays Vintage Base Ball Club makes two kinds of bats: the Burlingame and the Hickory. They can turn up to 400 Burlingames and 500 Hickorys per year, but can only make 650 bats total per year. It takes 20 man hours to produce a Burlingame and 40 man hours to produce a Hickory and the team has 22,000 man hours per year to make bats. The profit on the bats is \$40 for a Burlingame and \$60 for a Hickory.

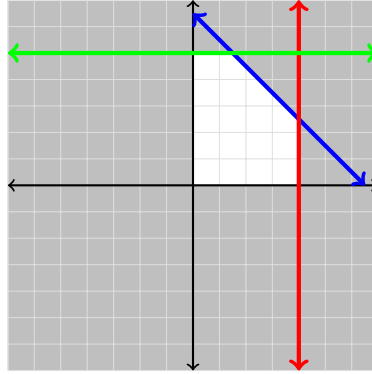
(a)

$$\begin{cases} x + y \leq 650 \\ 0 \leq x \leq 400 \\ 0 \leq y \leq 500 \end{cases}$$

(b) $P = 20x + 40y$

(c)

$$\begin{cases} y \leq -x + 650 \\ 0 \leq x \leq 400 \\ 0 \leq y \leq 500 \end{cases}$$

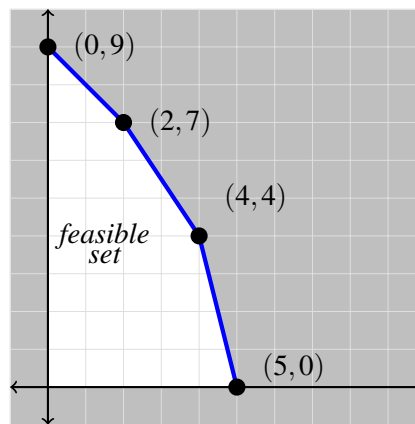


4.2 Linear Programming: Maximization and Minimization

We will approach this section by progressively adding more steps to the problem until we can start with a word problem and end with the quantities that optimize the situation. Last section we learned how to begin the problem but we stopped after the graphing of the feasible set. We next need to figure out what to do with the feasible set to get the x and y values to optimize the situation. We will first learn what to do with the points we obtain from the feasible set, then take a step back and learn how to find those points, and then finally we will put this section together with the last one to see how to solve a problem from the beginning.

4.2.1 Optimizing Given Lines and Points

Example 4.2.1 Use the given feasible set to maximize the object function $2x + 3y$.



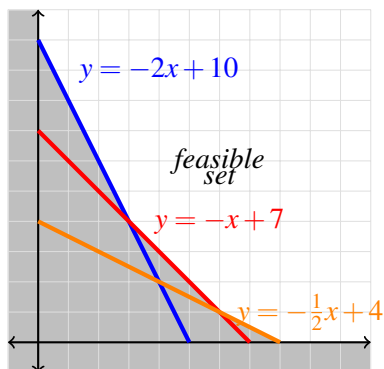
Solution The way we determine which x and y values maximize (or minimize) an object function is by substituting the values into the object function and see which values give us the largest or smallest value, as the case may be. Here we are looking to maximize $2x + 3y$, so we are looking for which ordered pair makes this expression the largest. One way we can organize this is to make a table.

Point	$2x + 3y$	Value
(0, 9)	$2(0) + 3(9)$	27
(2, 7)	$2(2) + 3(7)$	25
(4, 4)	$2(4) + 3(4)$	20
(5, 0)	$2(5) + 3(0)$	10

Since 27 is the largest value in the rightmost column, the ordered pair (0, 9) maximizes the expression $2x + 3y$.

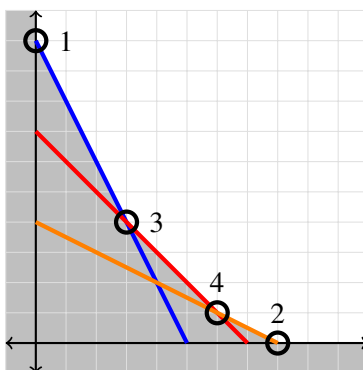
Where did these points come from? When we are dealing with a word problem, we are not going to be given the points. Rather, we will need to find them from the equations we will be working with. The next example shows us how to do this by using the concept of the intersection of a pair of lines that we learned about in chapter 3.

Example 4.2.2 Minimize the object function $5x + y$ using the information obtained from the given feasible set.



Solution We need to find the points at the corners of the feasible set so that we can substitute them into the object function as we did in Example 1. To do so, we find where the appropriate lines intersect; that is, we find where the lines intersect that make a corner of the feasible set. If the intersection does not occur in the feasible set, we do not care about that point because it cannot possibly provide the optimal point for the given situation. (The reason why the expression is optimized at a corner of the feasible set will be explained shortly.)

To find the values, we consider the equations associated with the inequalities above. For each pair of lines that intersect to make a corner of the feasible set, set the lines equal to each other to find the x -coordinate. Once we have the x -value, we can find the associated y -value by substituting the value we just found into either of the linear functions. We have four points to identify for this feasible set; there are two pairs of lines that intersect, one x -intercept and one y -intercept.



Point 1 is the y -intercept of $y = -2x + 10$. This point is $(0, 10)$.

Point 2 is the x -intercept of $y = -\frac{1}{2}x + 4$. This is the point where $y = 0$.

$$\begin{aligned}
 y &= -\frac{1}{2}x + 4 \\
 0 &= -\frac{1}{2}x + 4 \\
 \frac{1}{2}x &= 4 \\
 x &= 8
 \end{aligned}$$

So, this vertex of the feasible set is $(8, 0)$.

Point 3 is the intersection of the lines $y = -2x + 10$ and $y = -x + 7$.

$$\begin{aligned} -2x + 10 &= -x + 7 \\ -x + 10 &= 7 \\ -x &= -3 \\ x &= 3 \end{aligned}$$

By substituting this into either equation we get $y = 4$, giving the vertex of the feasible set as $(3, 4)$.

Point 4 is the intersection of the lines $y = -x + 7$ and $y = -\frac{1}{2}x + 4$.

$$\begin{aligned} -x + 7 &= -\frac{1}{2}x + 4 \\ -\frac{1}{2}x + 7 &= 4 \\ &= \frac{1}{2}x = -3 \\ x &= 6 \end{aligned}$$

By substituting this into either of these equations we get $y = 1$, giving the vertex of the feasible set as $(6, 1)$.

Now that we have the vertices of the feasible set, we can set up a chart to determine which point minimizes the object function.

Point	$5x + y$	Value
$(0, 10)$	$5(0) + 10$	10
$(8, 0)$	$5(8) + 0$	40
$(3, 4)$	$5(3) + 4$	19
$(6, 1)$	$5(6) + 1$	31

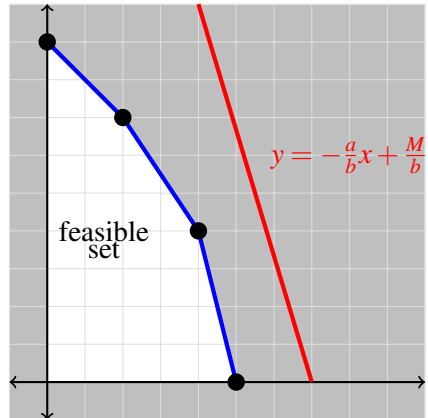
Since the minimum value is 10, the point that minimizes $5x + y$ is $(0, 10)$.

4.2.2 Why the Vertices Are the Only Points We Care About

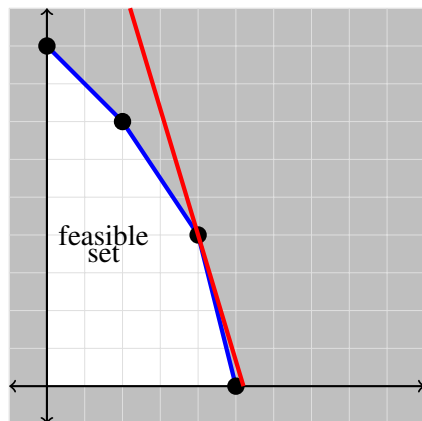
Let's say we want to optimize the expression $ax + by$ and the feasible set is given below. (Whether we are minimizing or maximizing, the explanation is the same.) Let us call the value that optimizes the expression M . If we put into slope intercept form, we get $M = ax + by$.

$$y = -\frac{a}{b}x + \frac{M}{b}$$

Notice that no matter what the optimizing value is, the slope of the object function is always the same.



The only thing that changes is the value of M because the slope is always the same. In order to find what value optimizes the situation, we ‘slide’ the object function to the feasible set until they intersect.



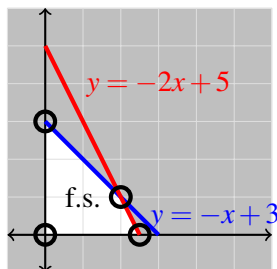
The value of M that gives the y -intercept of the line that intersects the feasible set is the optimizing value. The first place that the object line touches the feasible set is at a corner. Consequently, the only points we consider when optimizing are the corners of the feasible set.

4.2.3 More Involved Examples

Example 4.2.3 Maximize $3x + y$ subject to the constraints

$$\begin{cases} y \leq -x + 3 \\ y \leq -2x + 5 \\ x \geq 0, y \geq 0 \end{cases}$$

Solution Since all of the inequalities are in slope-intercept form, we can go right to the plot so that we can find the vertices of the feasible set.



There are three corners to this feasible set.

Point 1 is the y -intercept of $y = -x + 3$, which is $(0, 3)$.

Point 2 is the x -intercept of $y = -2x + 5$, which is $(\frac{5}{2}, 0)$.

Point 3 is the intersection of $y = -x + 3$ and $y = -2x + 5$.

$$\begin{aligned} -x + 3 &= -2x + 5 \\ x + 3 &= 5 \\ x &= 2 \end{aligned}$$

Point 4 is the origin, $(0, 0)$. Substituting this into either equation gives the point $(2, 1)$.

Testing these points gives

Point	$3x + y$	Value
$(0, 3)$	$3(0) + 3$	3
$(2, 1)$	$3(2) + 1$	7
$(\frac{5}{2}, 0)$	$3(\frac{5}{2}) + 0$	$\frac{15}{2}$
$(0, 0)$	$3(0) + 0$	0

The expression $3x + y$ attains its maximum value of $\frac{15}{2}$ at the point $(\frac{5}{2}, 0)$.

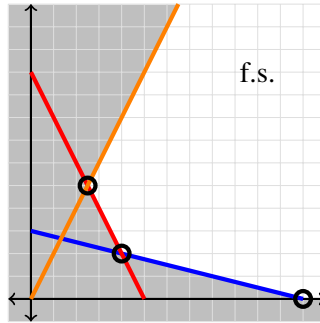
Example 4.2.4 Minimize $2x + 3y$ subject to the constraints

$$\begin{cases} x + 4y \geq 12 \\ 2x + y \geq 10 \\ y \leq 2x \\ x \geq 0, y \geq 0 \end{cases}$$

Solution We first need to put these inequalities in slope-intercept form so that we can plot the feasible set.

$$\begin{cases} x + 4y \geq 12 \\ 2x + y \geq 10 \\ y \leq 2x \\ x \geq 0, y \geq 0 \end{cases} \implies \begin{cases} y \geq -\frac{1}{4}x + 3 \\ y \geq -2x + 10 \\ y \leq 2x \\ x \geq 0, y \geq 0 \end{cases}$$

Now that we have the inequalities in the correct form, we can plot them to see what the feasible set looks like.



There are three corners for this feasible set; there are two intersections of lines and one x -intercept. Notice here that the y -intercepts of each of the three lines lie outside the feasible set and therefore are not considered.

Point 1 is the x -intercept of $y = -\frac{1}{4}x + 3$. This intercept is the point $(12, 0)$.

Point 2 is the intersection of $y = -\frac{1}{4}x + 3$ and $y = -2x + 10$.

$$\begin{aligned} -\frac{1}{4}x + 3 &= -2x + 10 \\ \frac{7}{4}x + 3 &= 10 \\ \frac{7}{4}x &= 7 \\ x &= 4 \end{aligned}$$

Substituting this value into either of the associated equations gives the point $(4, 2)$.

Point 3 is the intersection of the lines $y = 2x$ and $y = -2x + 10$.

$$\begin{aligned} 2x &= -2x + 10 \\ 4x &= 10 \\ x &= \frac{5}{2} \end{aligned}$$

Substituting this into either of the equations gives the point $(\frac{5}{2}, 5)$.

Now that we have the corners of the feasible set we need to test the points to see at which point we attain the minimum value.

Point	$2x + 3y$	Value
$(12, 0)$	$2(12) + 3(0)$	24
$(\frac{5}{2}, 5)$	$2(\frac{5}{2}) + 3(5)$	20
$(4, 2)$	$2(4) + 3(2)$	14

The minimum value of 14 is attained at the point $(4, 2)$.

4.2.4 Word Problems

Example 4.2.5 *The Providence Grays Vintage Base Ball Club makes two different types of baseballs. One ball has a figure eight cover and the other has a tulip stitch cover. The figure eight ball requires 1 hour to cut the leather, 2 hours to wind the ball and 2 hours to stitch on the cover. The tulip stitch ball requires 1 hour to cut the leather, 3 hours to wind the ball and 1 hour to stitch on the cover. The figure eight ball sells for \$20 each and the tulip stitch ball sells for \$18. The Grays have a maximum of 70 hours per month to spend on cutting leather, 180 hours to spend on winding and 120 hours to spend on covering the baseballs. If the cost of the materials is the same for each type of balls, how many of each type of ball should be made each month to maximize revenue?*

Solution In order to organize the data we have, it is sometimes helpful to create a table.

	Cutting	Winding	Covering	Price
Figure Eight	1	2	2	20
Tulip Stitch	1	3	1	18
Hours/Month	70	180	120	

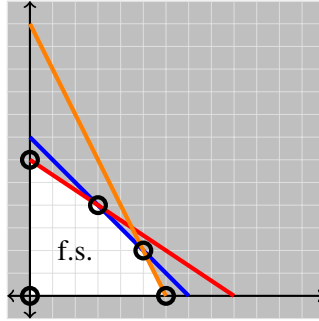
Now, we need to assign variables to each of the baseballs. Although we could use any letter we want to represent them, for simplicity's sake, let x represent the number of figure eight baseballs and let y represent the number of tulip stitch baseballs. From the table, we can now write out constraint inequalities and our object function.

Our task is to find the amount of each type of ball we want to produce to maximize revenue and we know how much each ball sells for. This gives us an object function of $P = 20x + 18y$. We have that it takes 1 hour to cut the leather for each type of ball and a maximum of 70 hours to spend on this task, so we get $x + y \leq 70$. It takes 2 hours to wind the figure eight ball and 3 to wind the tulip stitch ball and we have a maximum of 180 hours allotted for winding, so we get $2x + 3y \leq 180$. Finally, it takes 2 hours to cover the figure eight ball and only 1 hour to cover the tulip stitch ball and we have 120 hours for this task, so our third constraint inequality is $2x + y \leq 120$. Since we are talking about a real-life situation, it is not reasonable to consider a negative number of either type of ball being produced, so we additionally have the constraints $x \geq 0$ and $y \geq 0$.

We now need these inequalities to be written in slope-intercept form so that we can plot them to determine the feasible set.

$$\begin{cases} x + y \leq 70 \\ 2x + 3y \leq 180 \\ 2x + y \leq 120 \\ x \geq 0, y \geq 0 \end{cases} \implies \begin{cases} y \leq -x + 70 \\ y \leq -\frac{2}{3}x + 60 \\ y \leq -2x + 120 \\ x \geq 0, y \geq 0 \end{cases}$$

Plotting these and shading appropriately gives us the following feasible set.



There are four corners to this feasible set.

Point 1 is the y -intercept of $y = -\frac{2}{3}x + 60$, which is $(0, 60)$.

Point 2 is the x -intercept of $y = -2x + 120$, which is $(60, 0)$.

Point 3 is the intersection of $y = -\frac{2}{3}x + 60$ and $y = -x + 70$.

$$\begin{aligned} -\frac{2}{3}x + 60 &= -x + 70 \\ \frac{1}{3}x + 60 &= 70 \\ \frac{1}{3}x &= 10 \\ x &= 30 \end{aligned}$$

Substituting this into either equation yields the point $(30, 40)$.

Point 4 is the intersection of $y = -x + 70$ and $y = -2x + 120$.

$$\begin{aligned} -x + 70 &= -2x + 120 \\ x + 70 &= 120 \\ x &= 50 \end{aligned}$$

When we substitute this x -value into either equation, we get the ordered pair $(50, 20)$.

Point 5 is the origin, $(0, 0)$.

Now that we have all of the corners of the feasible set, we need to test the points to see which one maximizes the revenue.

Point	$20x + 18y$	Value
$(0, 60)$	$20(0) + 18(60)$	1080
$(60, 0)$	$20(60) + 18(0)$	1320
$(30, 40)$	$20(30) + 18(40)$	1360
$(50, 20)$	$20(50) + 18(20)$	1200
$(0, 0)$	$20(0) + 18(0)$	0

So, the maximum revenue of \$1360 is attained when 50 figure eight baseballs and 20 tulip stitch balls are produced and sold.

Example 4.2.6 The Grassy Knoll Lawnmower Company makes two kinds of lawnmowers, the ‘Mulcher’ and the ‘Bagger’, in two production plants. The Boston plant produces 32 Mulchers and 40 Baggers in one week, while the Newport plant produces 24 Mulchers and 40 Baggers per week. An order is received for 192 Mulchers and 280 Baggers. It costs \$2000 per week to operate the Boston factory and \$1600 per week to operate the Newport factory. How many weeks should the manufacturer operate each factory to fill the order at minimum cost?

Solution Like the last problem, we will begin with a chart to organize the data we are given in the problem.

	Boston	Newport	Order
Mulcher	32	24	192
Bagger	40	40	280
Cost/Week	\$2000	\$1600	

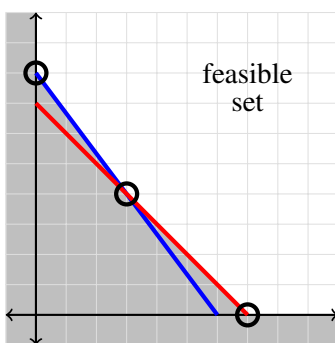
We will label the number of weeks that the Boston factory is open as x and use y for the number of weeks that the Newport factory will be open. This gives us the object function $C = 2000x + 1600y$. We have following constraints:

$$\begin{cases} 32x + 24y \geq 192 \\ 40x + 40y \geq 280 \\ x \geq 0, y \geq 0 \end{cases}$$

As before, we need to write these in slope-intercept form.

$$\begin{cases} 32x + 24y \geq 192 \\ 40x + 40y \geq 280 \\ x \geq 0, y \geq 0 \end{cases} \implies \begin{cases} y \geq -\frac{4}{3}x + 8 \\ y \geq -x + 7 \\ x \geq 0, y \geq 0 \end{cases}$$

Now we plot our constraints so that we can find the corners of our feasible set.



As we can see from the figure, there are three corners to the feasible set.

Point 1 is the y -intercept of the line $y = -\frac{4}{3}x + 8$, which is the point $(0, 8)$.

Point 2 is the x -intercept of the line $y = -x + 7$, which is the point $(7, 0)$.

Point 3 is the intersection of $y = -\frac{4}{3}x + 8$ and $y = -x + 7$.

$$-\frac{4}{3}x + 8 = -x + 7$$

$$8 = \frac{1}{3}x + 7$$

$$1 = \frac{1}{3}x$$

$$3 = x$$

Substituting this into either equation gives the point (3, 4).

Now we test these points to see which minimizes our cost function.

Point	$2000x + 1600y$	Value
(0, 8)	$2000(0) + 1600(8)$	12800
(3, 4)	$2000(3) + 1600(4)$	12400
(7, 0)	$2000(7) + 1600(0)$	14000

This gives a minimum cost of \$12,400 when the Boston factory is operated for 3 weeks and the Newport factory is operated for 4 weeks.

Example 4.2.7 *A potter is making cups and plates. It takes her 6 minutes to make a cup and 3 minutes to make a plate. Each cup uses 3/4 lb. of clay and each plate uses one lb. of clay. She has 20 hours available for making the cups and plates and has 250 lbs. of clay on hand. She makes a profit of \$2 on each cup and \$1.50 on each plate. How many cups and how many plates should she make in order to maximize her profit?*

Solution Note that time is given in hours and minutes, but we need to pick one to make it consistent. In order to avoid more fractions, we will use minutes.

First, we set up our table.

	cups	plates	limits
time	6	3	1200
clay	$\frac{3}{4}$	1	250
profit	2.00	1.50	

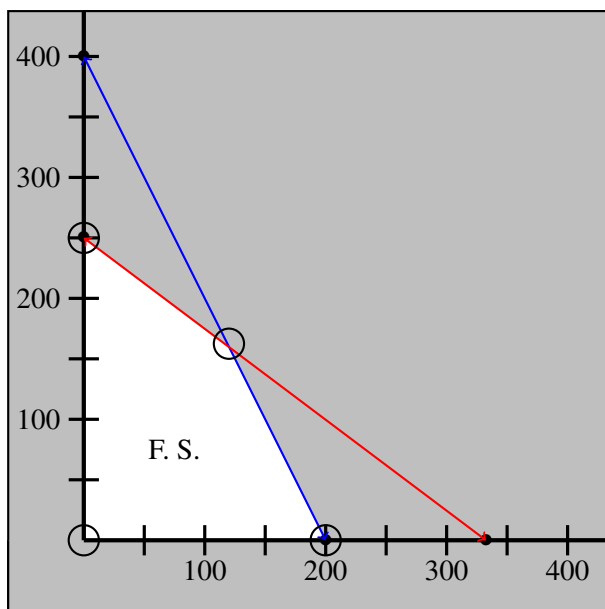
From the table, we get that we need to maximize $P = 2x + 1.5y$ subject to the constraints

$$\begin{cases} 6x + 3y \leq 1200 \\ \frac{3}{4}x + y \leq 250 \\ x \geq 0, y \geq 0 \end{cases}$$

and when we rewrite them in slope-intercept form we have

$$\begin{cases} y \leq -2x + 400 \\ y \leq -\frac{3}{4}x + 250 \\ x \geq 0, y \geq 0 \end{cases}$$

Graphing, we see



There are 4 corner points to be determined.

1. The origin $(0, 0)$
2. The y -intercept of $y = -\frac{3}{4}x + 250$, which is $(0, 250)$
3. The x -intercept of $y = -2x + 400$

$$\begin{aligned} 0 &= -2x + 400 \\ 2x &= 400 \\ x &= 200 \end{aligned}$$

So, we have the point $(200, 0)$

4. The intersection of $y = -\frac{3}{4}x + 250$ and $y = -2x + 400$.

$$\begin{aligned} -2x + 400 &= -\frac{3}{4}x + 250 \\ -8x + 1600 &= -3x + 1000 \\ 600 &= 5x \\ 120 &= x \end{aligned}$$

Substituting into either equation we get $y = 160$, giving the point of intersection as $(120, 160)$.

Now, we test the points.

Point	$P = 2x + 1.5y$	Value
$(0, 0)$	$2(0) + 1.5(0)$	0
$(0, 250)$	$2(0) + 1.5(250)$	375
$(200, 0)$	$2(200) + 1.5(0)$	400
$(120, 160)$	$2(120) + 1.5(160)$	480

The maximum profit of \$480 occurring when 120 cups and 160 plates are sold.

4.2.5 Exercises

For problems 1-3, graph the system of inequalities and find the corners of the feasible set. You are not given an object function here, so you are not trying to optimize here.

$$1. \quad \begin{cases} x+y \leq 10 \\ 2x+y \leq 8 \\ 2x+2y \geq 4 \\ x \geq 0, y \geq 0 \end{cases} \quad 2. \quad \begin{cases} x+y \leq 20 \\ 2x+y \leq 25 \\ x \leq 10 \\ x \geq 0, y \geq 0 \end{cases} \quad 3. \quad \begin{cases} 3x+2y \leq 72 \\ 2x+4y \leq 80 \\ x \leq 20 \\ x \geq 0, y \geq 0 \end{cases}$$

4. Minimize $2x + 5y$ subject to the constraints

$$\begin{cases} 4x+2y \leq 16 \\ \frac{1}{2}x + \frac{1}{2}y \leq 5 \\ 2x+2y \geq 8 \\ x \geq 0, y \geq 0 \end{cases}$$

5. Maximize $3x + 5y$ subject to the constraints

$$\begin{cases} 4x+y \leq 12 \\ 3x+3y \geq 6 \\ y \leq 8 \\ x \geq 0, y \geq 0 \end{cases}$$

6. Minimize $3x + 5y$ subject to the constraints

$$\begin{cases} 4x+y \leq 12 \\ 3x+3y \geq 6 \\ y \leq 8 \\ x \geq 0, y \geq 0 \end{cases}$$

7. Maximize $2x + 3y$ subject to the constraints

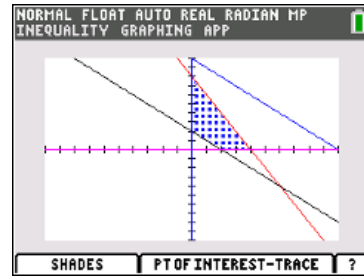
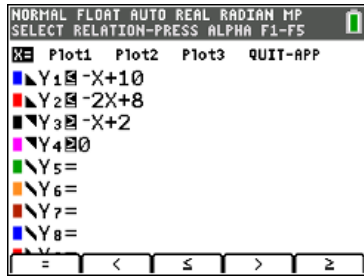
$$\begin{cases} x+y \geq 10 \\ 2x+4y \leq 24 \\ 4x+2y \leq 24 \\ x \geq 0, y \geq 0 \end{cases}$$

8. The Lampes Watch Company makes two stopwatches, one regular and one specifically for runners. There are 12 hours available per day on the assembly line and 18 hours available per day at the packaging center. Each regular watch takes 2 hours of assembly and 2 hours to package. Each runner's watch takes 1 hour to assembly but 3 hours to package. If the profit from each runner's watch is \$70 and is \$60 from each regular watch, how many of each should be made to maximize profit?

9. A new diet fad requires at least 90 units of protein, no more than 180 units of sodium and no more than 120 units of sugars. From each serving of the energy bar, the consumer gets 45 units of protein, 45 units of sodium and 15 units of sugars. From each serving of the energy drink, the consumer gets 30 units of protein, 30 units of sodium and 30 units of sugars. If the energy bar costs \$2.00 and the energy drink costs \$1.80, how many of each can you buy while keeping the cost at a minimum?

4.2.6 Solutions

1.



We can see that the blue line does not make any corners of the feasible set, but the x and y intercepts of the red and black lines do.

$$y = -x + 2$$

$$x - \text{int} : 0 = -x + 2 \Rightarrow x = 2 \Rightarrow (2, 0)$$

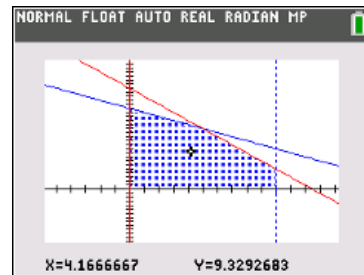
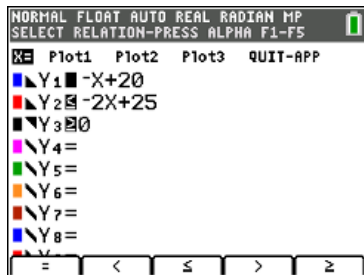
$$y - \text{int} : y = -(0) + 2 \Rightarrow y = 2 \Rightarrow (0, 2)$$

$$y = -2x + 8$$

$$x - \text{int} : 0 = -2x + 8 \Rightarrow x = 4 \Rightarrow (4, 0)$$

$$y - \text{int} : y = -2(0) + 8 \Rightarrow y = 8 \Rightarrow (0, 8)$$

2.



Here, we can see five corners: the origin $(0, 0)$, the x -intercept of $x \leq 10$, which is $(10, 0)$, the y -intercept of $y = -x + 20$, which is $(0, 20)$, where the red and blue lines intersect, and where the red line intersects with the vertical line.

$$\text{red} = \text{blue}$$

$$-x + 20 = -2x + 25$$

$$x = 5$$

$$y = -5 + 20 = 15$$

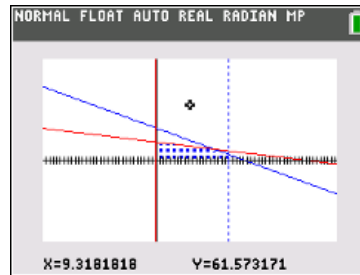
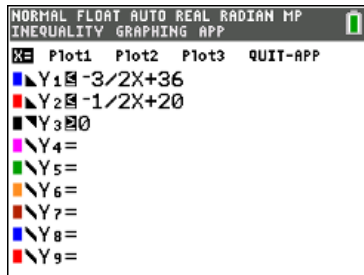
$$(5, 15)$$

$$\text{red} = \text{vertical}$$

$$y = -2(10) + 25 = 5$$

$$(10, 5)$$

3.

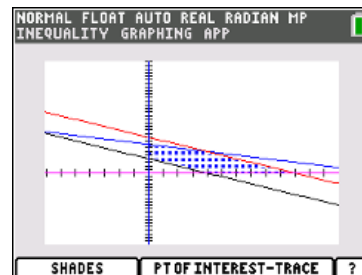
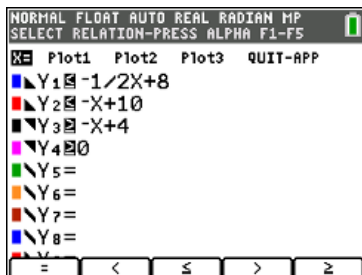


The corners here are: the origin (0,0), the y-intercept of the red line, (0,20), the x-intercept of the vertical line, (20,0), where the blue line intersects the vertical line, and where the red and blue lines intersect.

$$\begin{aligned} \text{blue} &= \text{vertical} \\ y &= -\frac{3}{2}(20) + 36 = 6 \\ &(20, 6) \end{aligned}$$

$$\begin{aligned} \text{red} &= \text{blue} \\ -\frac{3}{2}x + 36 &= -\frac{1}{2}x + 20 \\ -3x + 72 &= -x + 40 \\ x &= 16 \\ y &= -\frac{3}{2}(16) + 36 = 12(16, 12) \end{aligned}$$

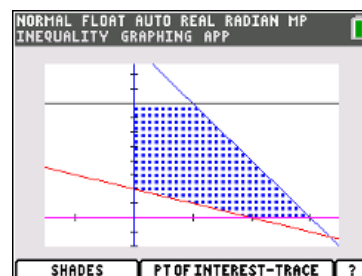
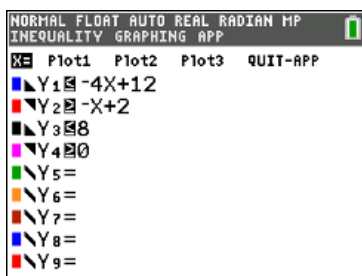
4.



Corner	Point	Value $m = 2x + 5y$
y-int black line	(0, 4)	20
y-int blue line	(0, 8)	40
red = blue	(4, 6)	38
x-int red	(10, 0)	20
x-int black	(4, 0)	8

So, the minimum of 8 occurs at (4,0).

5.

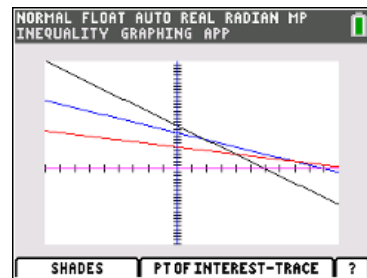
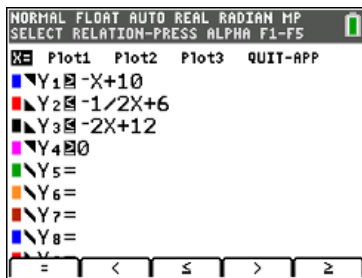


Corner	Point	Value $M = 3x + 5y$
x-int blue	(3,0)	9
blue = black	(1,8)	43
y-int black	(0,8)	40
y-int red	(0,2)	10
x-int red	(2,0)	6

The maximum value of 43 occurs at (1,8).

6. Notice that this is the same set of constraints and object function as the last problem. The difference is that here, we want to minimize. Inspecting the work we did already shows that the minimum value of 6 occurs at (2,0).

7.



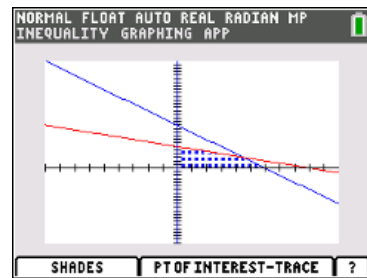
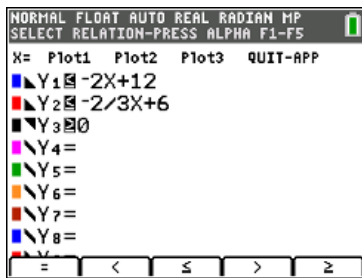
Notice that there are no regions that satisfy all of the inequalities at the same time. This means that there is no solution to this problem. It may be uninteresting, but it happens sometimes. We have to be careful, though, when it comes to applications that this does not happen. If it did, it would mean that in a business setting, we would be presented with an impossible production situation ...

8.

	Assembly	Packaging	Profit
Regular	2	2	60
Runner	1	3	70
Limits	12	18	

Maximize $P = 60x + 70y$ subject to

$$\begin{cases} 2x + y \leq 12 \\ 2x + 3y \leq 18 \\ x \geq 0, y \geq 0 \end{cases}$$



Corner	Point	Value $P = 60x + 70y$
red = blue	$(\frac{9}{2}, 3)$	\$480
x-int blue	(6,0)	\$360
origin	(0,0)	0
y-int red	(0,6)	\$420

The maximum occurs at $(\frac{9}{2}, 3)$, but this is impossible - we cannot make half a watch. We cannot make more watches either or we will exceed the limits we have on time. So, we would need to round down to 4 regular watches. If we do, this would give $4(60) + 3(70) = \$450$, which would give a larger profit than any of the corner points.

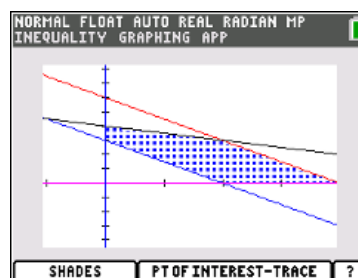
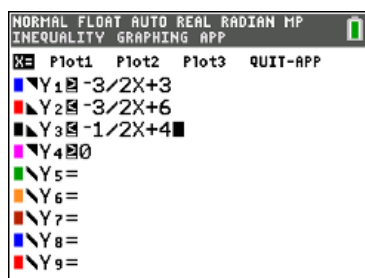
Note: Sometimes, in real-life situations, these fractional answers do come into play. Most of the problems in this text are contrived in that they only give answers that make sense in context if you solve correctly. But we wanted to point out that these situations may arise and that we have to adjust to have the answer make sense in the context of the problem.

9.

	protein	sodium	sugars	cost
bar	45	45	15	\$2.00
drink	30	30	30	\$1.80
limits	≥ 90	≤ 180	≤ 120	

Minimize $C = 2.00x + 1.80y$ subject to

$$\begin{cases} 45x + 30y \geq 90 \\ 45x + 30y \leq 180 \\ 15x + 30y \leq 120 \\ x \geq 0, y \geq 0 \end{cases}$$



Corner	Point	Value $C = 2x + 1.8y$
red = black	(2, 3)	\$9.40
x-int blue	(2, 0)	\$4.00
y-int blue	(0, 3)	\$5.40
y-int black	(0, 4)	\$7.20
x-int red	(4, 0)	\$8.00

We minimize cost by buying two energy bars and no energy drinks.

4.3 Linear Programming on the TI-84

If you have a TI-84, you can minimize or maximize using the Inequalz application. We still have to set the problem up in the same manner as before, but once we reach the point where we have the constraints in slope- intercept form and are ready to graph, we can do the rest on the calculator. The procedure will be the same for any of the problems we do. So, to show you that we will get the same solution using this method, we will illustrate this by solving an example from the previous section.

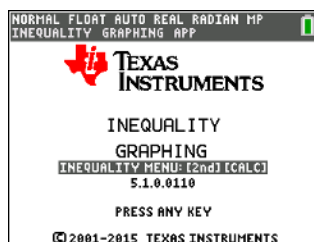
Example 4.3.1 Minimize $2x + 3y$ subject to the constraints

$$\begin{cases} x + 4y \geq 12 \\ 2x + y \geq 10 \\ y \leq 2x \\ x \geq 0, y \geq 0 \end{cases}$$

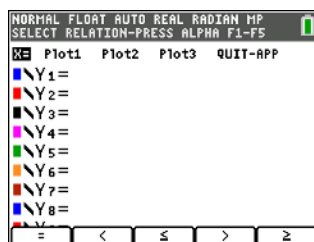
Solution We first need to put these inequalities in slope-intercept form so that we can plot the feasible set.

$$\begin{cases} x + 4y \geq 12 \\ 2x + y \geq 10 \\ y \leq 2x \\ x \geq 0, y \geq 0 \end{cases} \implies \begin{cases} y \geq -\frac{1}{4}x + 3 \\ y \geq -2x + 10 \\ y \leq 2x \\ x \geq 0, y \geq 0 \end{cases}$$

Now, press the **APPS** key, then scroll down to **Inequalz** and press **Enter**. The screen should look like this:

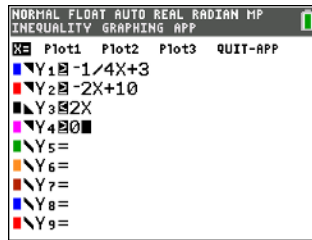


When you press a key, the screen looks similar to the way it does when you press **Y=**, but it gives us the option to also include x constraints. We can also select the equation and then using the **Alpha** key, we can select the appropriate inequality.

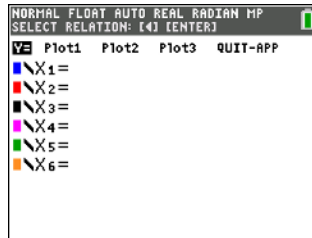


Now we can input the functions just as we would if we were graphing the equation, but we can include the inequality too.

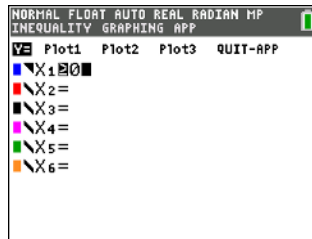
So, using this ability, we can change the equal signs to the appropriate inequalities and then input the functions themselves in the usual way. We can also change the signs by using the



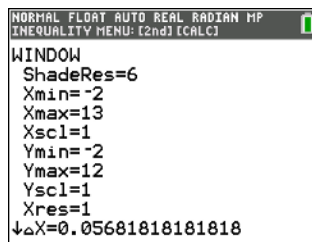
Notice now that there is a 'X=' in the upper right corner of the screen on which you input the functions. Move the cursor to here and press **ENTER**. The screen will look like this:



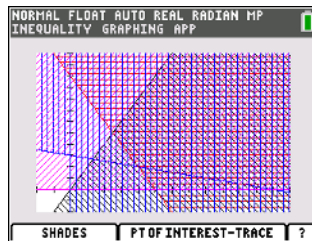
This is where we can put in our x constraints. We only have the one in this case, so when that is put in, the screen will look like



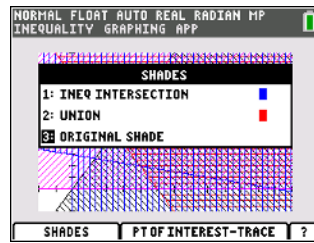
Once we have all of the constraints in, we now want to look at the plot of these inequalities. Before doing so, however, we need to make sure the window is of the correct restrictions to show what we want it to. For this example, try the following values (you can edit them by pressing the **WINDOW** key).



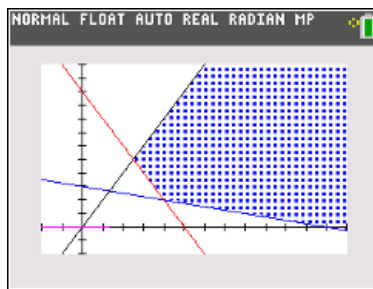
This will give the following plot:



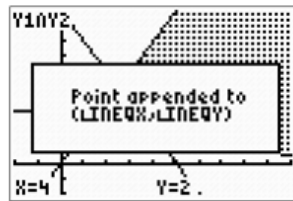
This doesn't appear to be all that helpful. However, by using the *Shades* command (on the screen) we can make this feasible set much easier to see. To use this, press **ALPHA** then **WINDOW**. This brings up a new menu.



The option we want here is the intersection of the inequalities, so press the **1** key. This will only give us where all of the inequalities coincide; that is, it will give us the feasible set.



If the only thing we could do with this was to get a clear picture of the feasible set, then it would be worthwhile, but there is more that we can do to help us solve the problem more quickly. We can use the calculator to find the corners of the feasible set and also to evaluate those points with respect to the object function. To find the vertices of the feasible set, we use the *PoI-Trace* function. To activate this, press **ALPHA** then **ZOOM**. One of the points of intersection (corners of the feasible set) will now have a blinking x and a little circle over it. Save this point to lists called *INEQX* and *INEQY* by pressing the **STO** key. The screen will look like this when it is confirmed:



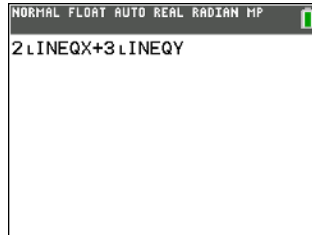
Repeat this process for the other two corners of the feasible set by using the arrows to maneuver to the other desired points and then press the **STO** key. Once done, press the **STAT** key and select *EDIT*. The following spreadsheet will appear:

INEQX	INEQY					?
4	2					
12	0					
2.5	5					

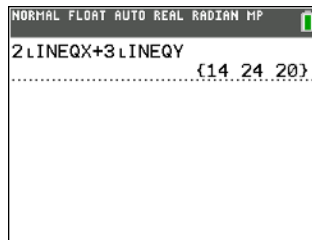
INEQX(1)= 4

In the two columns are the three corners from the feasible set that you collected. Now press the **2nd** then **MODE** key to clear the screen.

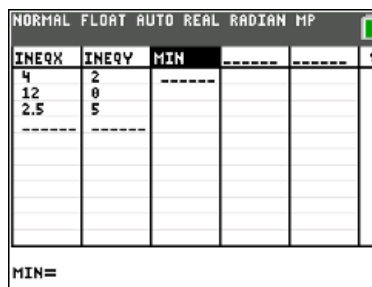
On this blank screen, input the object function (remember for this example, the object function was $2x + 3y$). When you input it, however, use *INEQX* instead of x and *INEQY* instead of y . To find these variables (*INEQX* and *INEQY*) press the keys **2nd** and **STAT**. These will be in the list. When done, the screen should look as:



When you press **ENTER**, you get the following to appear on the screen:



We can see that the minimum value is 14, but not which point corresponds to that value. We can also calculate these values in the spreadsheet so that we can see which point corresponds to which value. To do this, press **STAT** and **1** to get the spreadsheet on the screen. Then use the arrows to place the cursor in the label box in the column next to the *INEQY* column. In this place, name the column anything you want. I chose 'MIN' since we are looking for the minimum value here.



If you again use the arrows to scroll to this cell, it will have an equals sign after the 'MIN'. Here, you can input the object function just as you did on the blank screen earlier by using the *INEQX* and *INEQY* instead of x and y , respectively. Be sure to put the expression in quotation marks also, as this will make the list behave like spreadsheet. The keystrokes that will input the function are the following:



Note: the *INEQX* and *INEQY* are in the same location as before, so if they were not commands 7 and 8 in the *LIST* menu then use what you did in the prior step to get them.

When you press **ENTER**, the formula will automatically be copied to the cells below the 'MIN' heading and the values are automatically calculated.

INEQX	INEQY	MIN
4	2	14
12	0	24
2.5	5	20

MIN(1)=14

Here we can see that, as before, the minimum value of 14 is attained at the point (4,2).

Chapter 5

Exponential and Logarithmic Functions

5.1 Exponential Functions

5.1.1 What Makes a Function Exponential?

We saw earlier that the defining characteristic of a linear function is that there is a constant rate of change. This is not the case with exponential functions - the defining characteristic of an exponential function is that there is a constant ratio between consecutive y values. We can see this in the following table:

x	1	2	3	4	5	6
$f(x)$	4	5	6.25	7.8125	9.765625	12.20703125

If we calculate ratios of consecutive elements, we see ...

$$\frac{5}{4} = 1.25 \quad \frac{6.25}{5} = 1.25 \quad \frac{7.8125}{6.25} = 1.25$$

and this will continue in this manner. And, we also have

$$\frac{6.25}{4} = 1.5625 = 1.25^2 \quad \frac{7.8125}{5} = 1.5625 = 1.25^2$$

So where does this come from? How do exponential functions work? We will explore this through the next example.

Example 5.1.1 Suppose Sox tickets are \$45 and increase at 4% per year.

- Right now, after 0 years, the tickets cost \$45.
- After 1 year, the ticket costs $45(.04) + 45 = 45(1.04) = \46.80
- After 2 years, we have

$$\begin{aligned} 46.80 + 46.80(.04) &= 46.80(1.04) \\ &= (45(1.04))(1.04) \\ &= 45(1.04)^2 \\ &= \$48.67 \end{aligned}$$

- After 3 years, we have

$$\begin{aligned} 48.67 + 48.67(.04) &= 48.67(1.04) \\ &= (46.80(1.04))(1.04) \\ &= (45(1.04)(1.04))(1.04) \\ &= 45(1.04)^3 \\ &= \$50.62 \end{aligned}$$

We find the price for each successive year by adding 4% of the price to the current price, which is the same as finding 104% of the previous price (remembering that we need to express the percent as a decimal). So year two is 104% of the price of year one, which is 104% of year 0. So, after t years, we have that the price is $P(t) = 45(1.04)^t$.

5.1.2 Types of Exponential Functions

There are two types of exponential functions. The difference deals with how often we are multiplying by this common ratio.

Annual Growth

- $f(x) = a(1+r)^x$ or $f(x) = ab^x$
- a is the initial value
- r is the rate with $r > 0$ for growth, $r < 0$ for decay
- r must be in decimal form
- $b = 1 + r$ with $b > 0$

Continuous Growth

- $f(x) = ae^{kx}$
- a is the initial value
- k is the rate in decimal form
- $k > 0$ for growth, $k < 0$ for decay

The question is, when do we use each type? Commonly, we use annual growth when we work with situations like interest and other periodic changes. ‘Annual’ does not mean we can only use this form when we have increments taken once a year, but rather when we have periodic (yearly, monthly, daily, etc.) calculations.

Example 5.1.2 *You invest \$500 into a savings account at an interest rate of 2.3% per year. Write an exponential function giving the amount in the account after t years.*

Solution Since we are compounding yearly, we want to use an annual growth function. Since we are increasing, we know the rate is positive, and after we rewrite the percent, we see that $r = .023$. This gives that our situation is governed by the equation

$$P(t) = 500(1.023)^t$$

The other type of exponential growth function is for continuous growth. We use this type in situations where it wouldn’t make sense to only calculate at discrete times. Real-life applications involving population growth or the time required for caffeine to leave the blood stream would work like this.

Example 5.1.3 *Carbon-14 decays at a rate of 11.4% per 1000 years. Write a formula for the amount of carbon-14 left in an artifact after t thousand years if the artifact had 35 grams of carbon-14 when it was made.*

Solution Even though the problem says that the rate is per 1000 years, it wouldn’t make sense to say that there is some amount of carbon-14 left for 1000 years but then at the beginning of the 1001 year, there is suddenly 11.4% less carbon-14. So this is a candidate for continuous growth, where the rate is $-.114$ because it is a decay function. We do not know what the initial quantity is, so we will leave that as generic here.

$$Q(t) = ae^{-.114t}$$

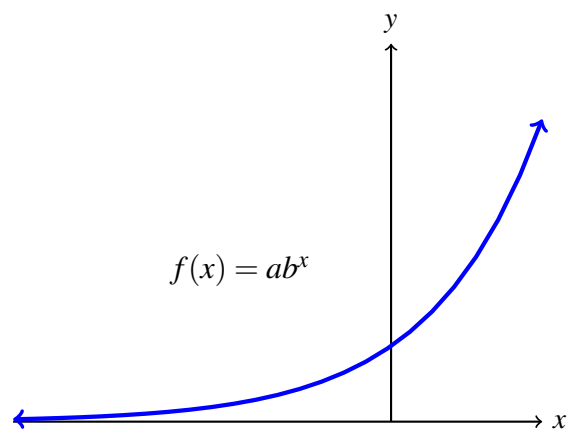
Now, we do have to be careful here because the rate is per 1000 years and not per year. This is taken care of when we express the number of years. For example, if we wanted to talk about how much carbon-14 an artifact had left after 2500 years in this situation, we would use $t = 2.5$ and this would leave us with

$$Q(2.5) = ae^{-.114(2.5)} \approx .752a$$

which shows that after 2500 years, we would have approximately 75.2% of the original carbon-14 remaining.

Note: we often use P_0 or Q_0 for the initial quantity instead of a . We do this so that we are explicitly representing the initial quantity as the quantity at time $t = 0$.

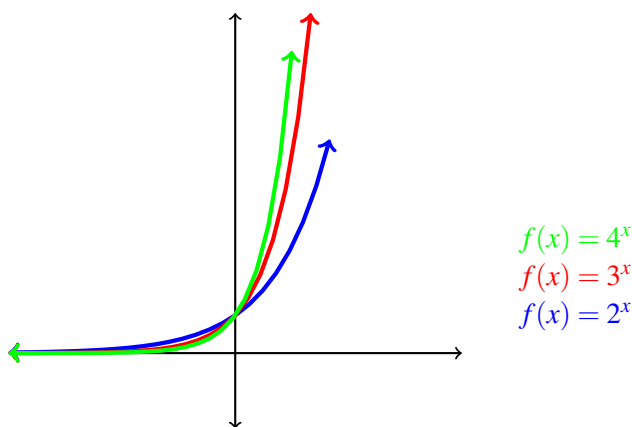
5.1.3 The Graph of Exponential Functions



Properties

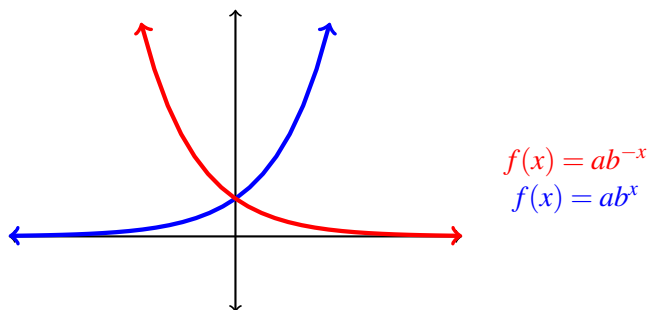
- Unless shifted, the function is never 0 or negative
- As $x \rightarrow \infty$, $f(x) \rightarrow \infty$
- As $x \rightarrow -\infty$, $f(x) \rightarrow 0$
- y-intercept is $(0, a)$

Changing the value of b doesn't change the shape of the graph, but changes how quickly the graph grows.

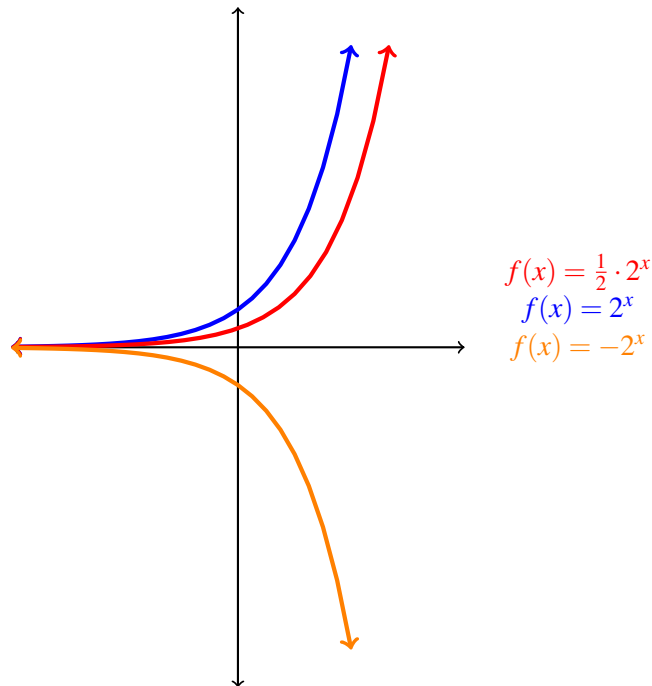


Notice that when $x < 0$, the curve with the larger base is closer to the x -axis.

When we look at the exponent, multiplying by a negative reflects the graph across the y -axis.



We also have the parameter a to worry about.

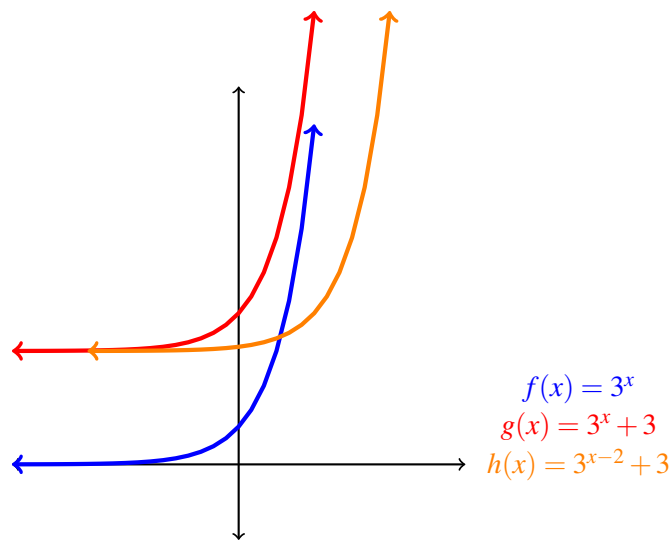


We can see that multiplying a by a negative reflects the graph across the x -axis. And, the larger the magnitude of a , the faster the curve grows.

We also can perform translations on exponential curves in the usual way. The general form of an exponential function representing annual growth is

$$f(x) = ab^{x-h} + k$$

where the value of k dictates a vertical shift by k units (where negative indicates a downward shift) and h represents a horizontal shift of h units (where a shift right is represented by $x - h$ and a shift left is represented by $x + h = x - (-h)$).



5.1.4 Examples

Example 5.1.4 Find the equation of the exponential function that passes through $(1, 3)$ and $(5, 27)$.

Solution Since we are not given enough information to determine which type of exponential function we need, we will assume we have annual growth. We first need to set up a system of equations using the data so that we can solve for a and b .

$$\begin{aligned}3 &= ab^1 \\27 &= ab^5 \\ \frac{27}{3} &= \frac{ab^5}{ab^1} \\9 &= b^4 \\b &= 1.732 = \sqrt[4]{9}\end{aligned}$$

So, $P(t) = a(\sqrt[4]{9})^t$. Now, we need to find a .

$$\begin{aligned}P(t) &= a(\sqrt[4]{9})^t \\3 &= a(\sqrt[4]{9})^1 \\a &= \frac{3}{\sqrt[4]{9}}\end{aligned}$$

Therefore, $P(t) = \frac{3}{\sqrt[4]{9}}(\sqrt[4]{9})^t$, which can be simplified to $P(t) = 3(\sqrt[4]{9})^{t-1}$ or $P(t) = (\sqrt[4]{9})^{t+1}$.

Example 5.1.5 A census is taken every two years, beginning in 1990. In 1992, there were 321,000 people in a certain city. In 1994, there were 345,000 people in that same city. If the population grows exponentially, find the formula for the population after t years. Express the population in thousands of people.

Solution Since we know this is an exponential model, we need only two points to find the formula. If we were not given the type of function, however, we would need to check ratios and rates to see what type of function we had.

We will let x represent the number of years since 1990 and let y represent the number of thousands of people x years since 1990. So, the data we are given, when writing them as ordered pairs, are $(2, 321)$ and $(4, 345)$.

Next decision is about which type of exponential function we need. Since we are only given census data in 2 year increments, we would use the annual growth form. That is, we are looking at something of the form $P(t) = ab^t$.

Next, we use the points to set up a system of equations.

$$\begin{aligned}321 &= ab^2 \\345 &= ab^4 \\ \frac{345}{321} &= \frac{b^4}{b^2} \\ \frac{345}{321} &= b^2 \\b &= 1.0367\end{aligned}$$

At this point, we know that $P(t) = a(1.0367)^t$. But we still need to solve for a , which we can do using one of the points and the equation we have so far.

$$\begin{aligned}P(t) &= a(1.0367)^t \\321 &= a(1.0367)^2 \\a &= \frac{321}{1.0747} \\a &= 298.678\end{aligned}$$

So, the function that gives the population after t years is

$$P(t) = 298.678(1.0367)^t$$

Example 5.1.6 *You take cold medicine at 100 grams per dose. The medicine leaves the blood stream at a rate of 22% per hour. How much is left in your blood stream after 4 hours?*

In order to solve, we need to determine which type of exponential function we need to work with and then determine what information we are given.

Since the medicine is continually leaving the blood stream, we want to find the equation of an exponential function of the form

$$Q(t) = Q_0e^{-kt}$$

We are given that the rate of decay is 22%, so we have $k = .22$ and we are given that the initial quantity $Q_0 = 100$ grams. Substituting these values into the equation, we have

$$Q(t) = 100e^{-.22t}$$

We want to find out how much is left in the blood stream after 4 hours has passed.

$$Q(4) = 100e^{-.88}$$

Using our calculators, we see $Q(4) = 41.48$ grams.

Example 5.1.7 *A cup of coffee has 95 mg of caffeine, which leaves the bloodstream at a rate of 17.33% per hour. You wake up and have a cup of coffee to get your day going and then have a second cup two hours later. How much caffeine is in your bloodstream after 3 hours if you had none in your system when you woke up?*

Solution This is another candidate for a continuous decay function. So, we need to write an equation of the form

$$Q(t) = Q_0e^{-kt}$$

We know $k = .1733$ and that $Q_0 = 95$ and that t is in hours since we drank the coffee. This gives

$$Q(t) = 95e^{-.1733t}$$

After two hours, the amount of caffeine in the bloodstream would be

$$Q(2) = 95e^{-1.733(2)} \approx 67.17 \text{ mg}$$

At this point, we drank another cup of coffee and added another 95 mg of caffeine, so we have a total of $95 + 67.17 = 162.17$ mg present. An hour later, which would be 3 hours since we started the first cup, we would have

$$Q(t) = 162.17e^{-.1733t}$$
$$Q(1) = 162.17e^{-.1733(1)} \approx 136.37 \text{ mg}$$

These examples all involve solving for a quantity at a given time given the rate of growth or decay. But what if we wanted to find the time required to reach a certain quantity or determine the rate of growth or decay? To solve problems like that, we need logarithmic functions.

5.1.5 Exercises

1. Determine which of the following tables could represent an exponential function. If one is exponential, find the formula.

x	1	2	3	4	5
$f(x)$	2	4	8	16	32
$g(x)$	2	5	8	11	14
$h(x)$	9	3	1	$\frac{1}{3}$	$\frac{1}{9}$

2. Find the formula for the exponential function passing through the points $(1, 5)$ and $(3, 20)$.
3. An investment of \$2,500 is made into an account with an annual interest rate of 3%. Write an appropriate exponential function that describes this situation and determine how much money will have been made after 10 years.
4. The amount of carbonation in soda decreases exponentially at a rate of 4% per minute. Write an appropriate function describing this situation and determine what percent of carbonation is left after one hour.
5. Suppose a bank account has \$1000 initially and after 3 years, the account has \$2,200. Find the exponential function that describes the balance in the account after t years.
6. The number of people who subscribe to a cable company has decreased over the past few years due to the lower cost of alternative providers like Netflix and Hulu. If the number of subscribers to cable decreases at an exponential rate of 15% per year and there are initially 100,000 subscribers, find the number of cable subscribers there will be in 10 years.

5.1.6 Solutions

1. $f(x)$ is an exponential function. The ratio of consecutive values is always 2.

$$\begin{aligned}f(x) &= a \cdot 2^x \\2 &= a \cdot 2^1 \\a &= 1\end{aligned}$$

This gives that $f(x) = 2^x$.

$g(x)$ is not an exponential function. Notice that $\frac{5}{2} = 2.5$ but $\frac{8}{5} = 1.6$. Since we do not get the same ratio, this function cannot be exponential.

$h(x)$ is an exponential function. The ratio of consecutive values is always $\frac{1}{3}$.

$$\begin{aligned}h(x) &= a \left(\frac{1}{3}\right)^x \\9 &= a \left(\frac{1}{3}\right)^1 \\a &= 27\end{aligned}$$

This gives that $h(x) = 27 \left(\frac{1}{3}\right)^x$.

$$\begin{array}{ll}2. & \begin{aligned}20 &= ab^3 \\5 &= ab^1 \\4 &= b^2 \\b &= 2\end{aligned} & \begin{aligned}y &= a \cdot 2^x \\20 &= a \cdot 2^3 \\a &= \frac{5}{2}\end{aligned}\end{array}$$

So, $f(x) = \frac{5}{2} \cdot 2^x = 5 \cdot 2^{x-1}$

- 3.

$$\begin{aligned}P(t) &= 2500(1.03)^t \\P(10) &= 2500(1.03)^{10} = \$3359.79\end{aligned}$$

- 4.

$$\begin{aligned}Q(t) &= Q_0(.96)^t \\Q(60) &= Q_0(.96)^{60} = .086Q_0\end{aligned}$$

So, there is approximately 8.6% of the carbonation left after one hour.

5. We will use the points (0, 1000) and (3, 2200).

$$\begin{array}{ll}2200 &= ab^3 \\1000 &= ab^0 \\2.2 &= b^3 \\b &= 1.3 & \begin{aligned}P(t) &= a(1.3)^t \\1000 &= a(1.3)^0 \\a &= 1000\end{aligned}\end{array}$$

So, $P(t) = 1000(1.3)^t$. Note that we didn't need to solve for a because we were given that \$1000 was the initial quantity.

6.

$$P(t) = 100000(.85)^t$$
$$P(10) = 100000(.85)^{10} = 19687.44$$

So, there will be roughly 19687 people still subscribing to cable at this rate after 10 years.

5.2 Logarithmic Functions

Why do we need logarithmic functions? If we want to solve something like $2^x = 16$, we need to figure out what exponent we need so that when we multiply 2 times itself that many times, we get 16. Here, $x = 4$ and we can reason through this relatively quickly. But what if we wanted to solve $2^x = 15$? It won't be an integer value and cannot be reasoned much more than that we can say the value of x is somewhere between 3 and 4 and probably much closer to 4. This is why we need logarithms.

5.2.1 What is a Logarithm?

The best way to think of a logarithm is as a different form of an exponential equation. Suppose we have

$$x = b^y$$

Logarithms tell us the power we need to raise b to in order to get x for the answer.

$$x = b^y \Leftrightarrow y = \log_b x$$

There are two main types of logarithms we are concerned with, and the reason is because these are the two that we can find values for in the calculator.

Common log: $y = \log x \Leftrightarrow 10^y = x$

Natural log: $y = \log_e x = \ln x \Leftrightarrow e^y = x$

When solving with logarithms, which one we use doesn't matter and we often choose which to use based on convenience.

5.2.2 Properties of Logarithms

Let's look at some of the most useful properties associated with logarithms. We will look at these with any generic base, but these are all true with any base. We can apply these rules to common logs and natural logs.

Rule 1: The Power Rule

$$\log_a x^n = n \log_a x$$

Proof: Let $m = \log_a x$. Then, using the definition of logarithms, we can rewrite this as

$$m = \log_a x \Rightarrow x = a^m$$

Now,

$$\begin{aligned} x &= a^m \\ x^n &= (a^m)^n \end{aligned}$$

Writing back in logarithmic form and substituting, we have

$$\begin{aligned} \log_a x^n &= nm \\ \log_a x^n &= n \log_a x \end{aligned}$$

Rule 2: The Product Rule

$$\log_a xy = \log_a x + \log_a y$$

Proof: Let $m = \log_a x$ and $n = \log_a y$. Then, using the definition of logarithms, we can write

$$m = \log_a x \Rightarrow x = a^m$$

$$n = \log_a y \Rightarrow y = a^n$$

Then,

$$x \cdot y = a^m \cdot a^n = a^{m+n}$$

Now, we take the logarithm, base a of both sides and simplify.

$$xy = a^{m+n}$$

$$\log_a xy = \log_a a^{m+n}$$

$$\log_a xy = (m+n)\log_a a$$

$$\log_a xy = m+n$$

$$\log_a xy = \log_a x + \log_a y$$

Rule 3: The Quotient Rule

$$\log_a \frac{x}{y} = \log_a x - \log_a y$$

Proof: We will proceed in the exact same manner as the product rule. We first let $m = \log_a x$ and $n = \log_a y$. Then, we have

$$\begin{aligned} m = \log_a x &\Rightarrow x = a^m \\ n = \log_a y &\Rightarrow y = a^n \end{aligned}$$

Now the algebra ...

$$\frac{x}{y} = \frac{a^m}{a^n} = a^{m-n}$$

Taking the logarithms, base a , of both sides and simplifying gives

$$\begin{aligned} \frac{x}{y} &= a^{m-n} \\ \log_a \frac{x}{y} &= \log_a a^{m-n} \\ \log_a \frac{x}{y} &= (m-n) \log_a a \\ \log_a \frac{x}{y} &= m-n \\ \log_a \frac{x}{y} &= \log_a x - \log_a y \end{aligned}$$

Rule 4: $\log_a 1 = 0$

Think ‘What power do we raise a to in order to get 1 for an answer?’

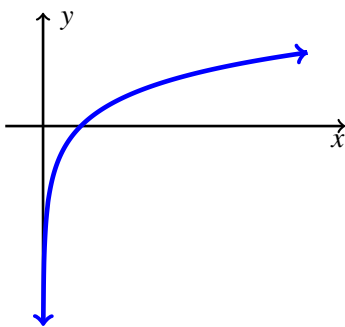
Rule 5: $\log_a a = 1$

Think ‘What power do we raise a to in order to get a for an answer?’

Rule 6: $a^{\log_a x} = x$

This is because a^x and $\log_a x$ are inverses.

5.2.3 Graphs of Logarithmic Functions

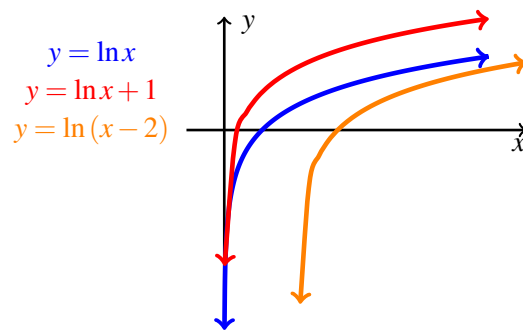


No matter what base we have for our logarithm, the graphs will all have the same shape and satisfy the following properties, provided we didn't shift the graph or multiply by a scalar that will change the rate of growth. The properties of the graph are:

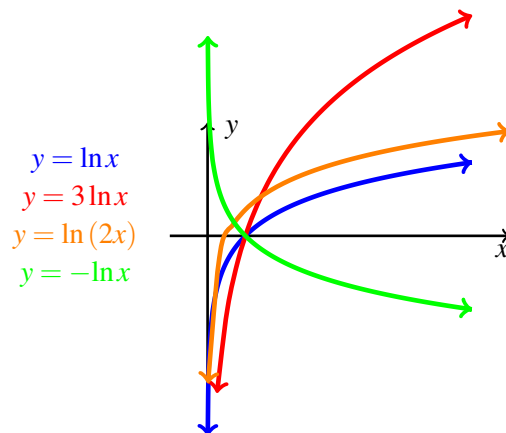
- The domain of $\log_a x$ is $(0, \infty)$.

- As $x \rightarrow 0$ from the right, $\log_a x \rightarrow -\infty$.
- As $x \rightarrow \infty$, $\log_a x \rightarrow \infty$.
- The graph is increasing and concave down.
- The graph never crosses the y-axis.
- The x-intercept is $(1, 0)$.

We can shift the graph around the plane in the usual manner.



We can also manipulate the shape of the graph by multiplying by scalars.



5.2.4 Using Logarithms

Like we said before, which logarithm we use doesn't change what answer we will get but a good choice may cut out some of the work.

Example 5.2.1 Solve $2^x = 15$

Option 1: Common or natural log

$$\begin{aligned}2^x &= 15 \\ \log 2^x &= \log 15 \\ x \log 2 &= \log 15 \\ x &= \frac{\log 15}{\log 2} \\ x &\approx 3.9069\end{aligned}$$

Option 2: Change of base

$$\begin{aligned}2^x &= 15 \\ \log_2 2^x &= \log_2 15 \\ x \log_2 2 &= \log_2 15 \\ x &= \log_2 15 \\ x &= \frac{\log 15}{\log 2} \\ x &\approx 3.9069\end{aligned}$$

The change of base formula can be used to change any base logarithm into another but the practical usage for us is to change into common log or into natural log so that we can use technology to get the numerical approximation. Notice in the second option, we use \log_2 , which we cannot find a numerical value for using the calculator. But using the change of base formula, we converted to common log and then could get a value.

The change of base formula works as follows:

$$\log_a x = \frac{\log_b x}{\log_b a}$$

Whereas we can do this for any base using this formula, there are really two that will be convenient for us.

$$\begin{aligned}\log_a b &= \frac{\log b}{\log a} \\ \log_a b &= \frac{\ln b}{\ln a}\end{aligned}$$

Example 5.2.2 Find a numerical approximation for $\log_3 2$.

Solution Since we are given base 3 and we cannot get a numerical value using the calculator, we will convert to one of the two convenient based. Here, we will use natural log.

$$\begin{aligned}\log_3 2 &= \frac{\ln 2}{\ln 3} \\ &\approx \frac{.6931}{1.0986} \\ &\approx .6309\end{aligned}$$

Example 5.2.3 Solve $e^{2x} = 7$.

Solution Since we are given an exponential function with e as the base, natural log is the most convenient choice here. By using that, we do not have to use the change of base formula to find the numerical value.

$$\begin{aligned}
e^{2x} &= 7 \\
\ln e^{2x} &= \ln 7 \\
2x \ln e &= \ln 7 \\
2x &= \ln 7 \\
x &= \frac{\ln 7}{2} \\
x &\approx .973
\end{aligned}$$

Example 5.2.4 Solve $\ln x = 3$.

Solution Here, since we are given a natural log function, we will solve this using the base-exponent property.

The base-exponent property states that if we have two equal quantities, then using those as exponents for a common base will not destroy equality. That is, $x = y \Leftrightarrow a^x = a^y$.

$$\begin{aligned}
\ln x &= 3 \\
e^{\ln x} &= e^3 \\
x &= e^3 \\
x &\approx 20.08
\end{aligned}$$

Example 5.2.5 Solve $2^{x+1} = 4^x$.

Solution We will use logarithms here as well. Since we do not have the same base on each side of the equation, there is no easier choice for which to use, so without loss of generality, natural log will be used here.

$$\begin{aligned}
2^{x+1} &= 4^x \\
\ln 2^{x+1} &= \ln 4^x \\
(x+1) \ln 2 &= x \ln 4 \\
x \ln 2 + \ln 2 &= x \ln 4 \\
\ln 2 &= x \ln 4 - x \ln 2 \\
\ln 2 &= x(\ln 4 - \ln 2) \\
\ln 2 &= x \ln \left(\frac{4}{2} \right) \\
\ln 2 &= x \ln 2 \\
1 &= x
\end{aligned}$$

Notice here that we could have saved ourselves a lot of work if we would have applied properties. If we can write both sides of the equation in terms of the same base, we could simply apply the base-exponent property in the other direction to solve.

$$\begin{aligned} 2^{x+1} &= 4^x \\ 2^{x+1} &= 2^{2x} \Rightarrow x+1 = 2x \\ 1 &= x \end{aligned}$$

Example 5.2.6 Solve $\log(x^2 + 3x) = 1$

Solution Here, we use the base-exponent property to simplify. Since we are given an equation involving the common log, we will use 10 as the base.

$$\begin{aligned} \log(x^2 + 3x) &= 1 \\ 10^{\log(x^2 + 3x)} &= 10^1 \\ x^2 + 3x &= 10 \\ x^2 + 3x - 10 &= 0 \\ (x+5)(x-2) &= 0 \\ x &= 2 \end{aligned}$$

The quadratic equation has two roots like it should; they are $x = 2$ and $x = -5$. so why is there only one solution? We have to remember the domain ... the only input values allowed are in the interval $(0, \infty)$. If we rewrote the original problem and applied properties, we would see

$$\begin{aligned} \log(x^2 + 3x) &= \log x(x+3) \\ &= \log x + \log(x+3) \end{aligned}$$

and if we let $x = -5$, neither of these is defined. We always have to consider these situations when solving problems of this type.

5.2.5 Exercises

1. Write $y = 3^x$ as a logarithm.
2. Write $y = e^{2x}$ as a logarithm.
3. Write $\log_3 x = y$ as an exponential function.
4. Write $\ln y = x$ as an exponential function.
5. Use properties to find the value of $\log_a bc^2$ if $\log_a b = 3$ and $\log_a c = 4$.
6. Use properties to find the value of $\log_a \frac{b^2}{c^3}$ if $\log_a b = 3$ and $\log_a c = 4$.
7. Find the numerical approximation for $\log_4 5$.
8. Solve $\log x = 4$.
9. Solve $\log_3 x = 10$.
10. Solve $\ln 2x = 4$.
11. Solve $\log(x^2 + 3x) = 1$.

5.2.6 Solutions

1. $\log_3 y = x$

2. $\ln y = 2x$

3. $3^y = x$

4. $e^x = y$

5.

$$\begin{aligned}\log_a bc^2 &= \log_a b + \log_a c^2 \\ &= \log_a b + 2\log_a c \\ &= 3 + 2(4) \\ &= 11\end{aligned}$$

6.

$$\begin{aligned}\log_a \frac{b^2}{c^3} &= \log_a b^2 - \log_a c^3 \\ &= 2\log_a b - 3\log_a c \\ &= 2(3) - 3(4) \\ &= -6\end{aligned}$$

7. $\log_4 5^{\frac{\log 5}{\log 4}} \approx 1.16$

8.

$$\begin{aligned}\log x &= 4 \\ 10^{\log x} &= 10^4 \\ x &= 10000\end{aligned}$$

9.

$$\begin{aligned}\log_3 x &= 10 \\ 3^{\log_3 x} &= 3^{10} \\ x &= 3^{10}\end{aligned}$$

10.

$$\begin{aligned}\ln 2x &= 4 \\ e^{\ln 2x} &= e^4 \\ 2x &= e^4 \\ x &= \frac{1}{2}e^4\end{aligned}$$

11.

$$\begin{aligned}\log(x^2 + 3x) &= 1 \\ 10^{\log(x^2 + 3x)} &= 10^1 \\ x^2 + 3x &= 10 \\ x^2 + 3x - 10 &= 0 \\ (x + 5)(x - 2) &= 0 \\ x &= 2, -5\end{aligned}$$

But ...

$$\begin{aligned}\log(x^2 + 3x) &= \log x(x + 3) \\ &= \log x + \log(x + 3)\end{aligned}$$

If $x = -5$, then the above would give us the logarithm of a negative number, which lies outside the domain of the function. Therefore, the only solution to this equation is $x = 2$.

5.3 Interest

One of the most common applications of exponential functions in the world of finance is that of interest calculations. There are two main types of interest, simple and compound. But before we get to formulas, there are a couple of terms we will use throughout.

Definition 5.3.1 • The interest on a loan is the amount of extra money that will be paid for taking out the loan

- The principle, P , is the amount borrowed
- r is the interest rate, expressed in decimal form
- t is time

5.3.1 Simple Interest

The idea here is that simple interest is that the amount of interest paid is the same every year. There is no factoring in of how much of the loan has been paid off over time.

Definition 5.3.2 Simple interest is interest calculated on only the principle.

Formula 5.3.3 To calculate the simple interest on a loan with principle P and interest rate r over t years, we use

$$I = Prt$$

Example 5.3.4 Suppose you borrowed \$2000 to buy a used car and the the loan will be over 5 years at 4% interest per year. What is the interest that will be paid?

Solution If we assume simple interest here, the interest will be

$$I = 2000(.04)(5) = \$400$$

We certainly don't have to only calculate simple interest on an annual basis. If we wanted to use this concept for another time period, we could calculate the interest for one time period and then multiply by how many time periods there are.

Example 5.3.5 Suppose you borrowed \$100 from a 'friend' who was charging you 10% interest per week. How much would you owe after 8 weeks?

Solution Here, we see that 10% of \$100 is \$10. After 8 weeks, we would therefore owe $8(\$10) = \80 in interest.

A couple of places we could see simple interest in real-life situations are car loans and short-term loans. But, most loans are calculated to include a more complicated way that benefits the lender because the interest rate is compounded often.

5.3.2 Compound Interest

Definition 5.3.6 *Compound interest is interest that is calculated on both the principle and on the interest accumulated over previous time periods of the loan.*

We have to be careful when talking about compound interest to be sure we are clear on not only the length of the loan/investment but also on how often the interest is compounded. The more often it is compounded, the more beneficial it is to the party receiving the payment.

Formula 5.3.7 *To calculate the compound interest where interest is compounded once per time period, we use*

$$I = P(1 + r)^t - P$$

where P is the principle, r is the nominal annual rate and t is the number of compounding periods.

Example 5.3.8 *Find the interest to be paid on a loan of \$10,000 with an interest rate of 5% per year if the loan is to be paid off over 10 years.*

Solution From what we had seen before, if this was a simple interest calculation, we would be paying back $.05(10000) = \$500$ and so for the life of the loan, we would pay \$5,000. But if the interest is compounded annually, we have

$$I = 10000(1 + .05)^{10} - 10000 = 10000(1.05)^{10} - 10000 = \$6288.95$$

As we can see in the last example, we have to pay much more back because we are paying interest on the interest that has yet to be paid off. This is exactly why banks and credit cards want you to pay minimum due.

Example 5.3.9 *Suppose you invest \$100 in a savings account that pays a nominal interest rate of 2.5% per year. If you never withdraw any money from this account and you never make another deposit, how much will you have earned in 25 years?*

Solution We have $P = 100$, $r = .025$ and $t = 25$. The interest is therefore

$$I = 100(1 + .025)^{25} - 100 = 100(1.025)^{25} - 100 = \$85.39$$

Example 5.3.10 *Suppose you invested \$1000 in a small company and they lost 10% per year for the first 5 years. You stuck by them and did not cut your losses but you did not invest more either. After the 5 years, the company turned a profit and you made 10% per year on what remained of your original investment. How much money do you have at the end of the 10 years?*

Solution We first need to determine what we had left from our original investment at the end of 5 years. We will do so by calculating the *future value* of the investment. This is essentially the same as what we have done already with the difference being that we do not subtract the principle so that we are not isolating the interest. Notice here that we are losing money at first, so $r = -.1$.

$$FV = 1000(1 - .1)^5 = 1000(.9)^5 = \$590.49$$

Notice that this actually worked out better for us than if the loss was calculated simply with the same loss per year.

Now, we need to calculate how much we made over the next 5 years. Since we are earning now, $r = .1$ and our principle at this point is \$590.49, the amount of money we had at the end of the 5 year period - we do not revert back to the original \$1000 to find the gain.

$$FV = 590.49(1 + .1)^5 = 590.49(1.1)^5 = \$950.99$$

So, since we were compounding our gain on a smaller principle than the one used to calculate our loss, we ended up losing money in the long run. If the company would have maintained this earning trend, however, we would have finally made a profit.

$$FV = 950.99(1.1)^1 = \$1046.09$$

In all of these examples, we assumed that we were compounding interest once per time period. What if we weren't? What if we were given the annual nominal interest rate but wanted to compound more often? We need to take this into account when calculating interest by dividing the rate by how many compounding periods per year and then multiplying the number of years by how many annual compounding periods.

Formula 5.3.11 *If we want to calculate the interest on a principle investment/loan in the amount of P dollars at a annual nominal rate of $r\%$ (expressed in decimal form) for t years with n compounding periods per year, we use*

$$I = P \left(1 + \frac{r}{n} \right)^{nt} - P$$

and if we want the future value of this investment/loan, we use

$$FV = P \left(1 + \frac{r}{n} \right)^{nt}$$

Example 5.3.12 *Suppose you invest \$100 in a savings account that pays a nominal interest rate of 2.5% per year. If you never withdraw any money from this account and you never make another deposit, how much will you have earned in 25 years if the interest was compounded on the first of every month?*

Solution This is similar to 5.3.9 but now we want to compound $n = 12$ times per year.

$$FV = 100 \left(1 + \frac{.025}{12} \right)^{12(25)} = \$186.70$$

We didn't make a great deal more money since this was a small investment with a small interest rate, but we can see the power of compounding more often.

Example 5.3.13 *Using the same information as last example, calculate the future value of the investment if we were to compound the interest daily.*

Solution We use the same formula, but this time with $n = 365$.

$$FV = 100 \left(1 + \frac{.025}{365} \right)^{12(365)} = \$186.82$$

Example 5.3.14 Suppose you take out a loan for \$10,000 at a nominal interest rate of 3% with the intent to pay back the loan in 10 years where interest is compounded weekly. Rather than pay it back in instalments, you decide to pay it off in one lump sum at the end of the 10 years. How much will you owe?

Solution Since we are paying off the entire balance at the end, we can use the same formula we have been working with.

$$FV = 10000 \left(1 + \frac{.03}{52} \right)^{10(52)} = \$13,497.42$$

Each example so far in this section has assumed that no other payments or deposits were made. If they are, we have to calculate interest on different amounts each period and we cannot use the simple or compound interest formulas discussed earlier.

5.3.3 Exercises

1. Find the simple interest on a loan of \$10,000 at an annual interest rate of 6 % for 3 years.
2. Find the simple interest on a loan of \$20,000 at an annual rate of 4 % for 18 months.
3. Find the simple interest on a loan of \$3,000 at an annual rate of 5 % for 75 days.
4. Find the interest when a loan of \$25,000 at an annual rate of 3.5% for 5 years where the interest is compounded annually.
5. Find the interest when a loan of \$25,000 at an annual rate of 3.5% for 5 years where the interest is compounded monthly.
6. Find the future value when a loan of \$25,000 at an annual rate of 3.5% for 5 years where the interest is compounded monthly.
7. You are given the opportunity to choose the type of investment you will make on a \$1000 initial deposit. You can choose, for a 10 year investment, from the following options.
 - Option 1: 6 % nominal interest rate with simple interest
 - Option 2: 5.5 % nominal interest rate, compounded annually
 - Option 3: 5.25 % nominal interest rate, compounded weekly
 - Option 4: 5 % nominal interest rate, compounded continuously

Which is the best investment?

8. You invest \$25,000 at a nominal rate of 3 % for 10 years compounded monthly. After the first 5 years, however, the investment was changed without your knowledge to a 3.25 % nominal interest rate but compounded annually. How much different will this investment be at the end of the 10 years? Did you make out in the change or did you lose potential money?

5.3.4 Solutions

1. $I = 10000(.06)(3) = \$1800$
2. $I = 20000(.04)(1.5) = \$1200$
3. $I = 3000(.05) \left(\frac{75}{365}\right) = \30.82
4. $I = 25000(1.035)^5 - 25000 = \4692.16
5. $I = 25000 \left(1 + \frac{.035}{12}\right)^{60} - 25000 = \4773.57
6. $FV = 25000 \left(1 + \frac{.035}{12}\right)^{60} = \29773.57
7.
 - $FV = 1000 + 1000(.06)(10) = \1600
 - $FV = 1000(1.055)^{10} = \1708.14
 - $FV = 1000 \left(1 + \frac{.0525}{52}\right)^{520} = \1690.01
 - $FV = 1000e^{.05(10)} = \$1648.72$

The best investment is option 2.

8. If the investment stayed the same for 10 years, we would have

$$FV = 25000 \left(1 + \frac{.03}{12}\right)^{120} = \$33733.84$$

But, for the changed investment, we would have

$$FV = 25000 \left(1 + \frac{.03}{12}\right)^{60} = 29040.42$$
$$FV = 29040.42(1.0325)^5 = 34076.36$$

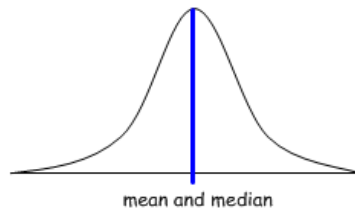
So, we actually made more money because of the change in the terms of the investment.

Chapter 6

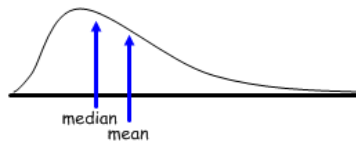
Introduction to Statistics

Which is better to use in what situation? The mean is an arithmetic average, so if there are any values that are not consistent with the rest of the data, they can skew the mean. In our first example, we had a mean near 27 and all of the data is between 21 and 35. If we added one more value to the set that was not in that same range, like 125, the mean would be 39.625. This mean is affected by what we call an **outlier**. An **outlier** is a value that is not consistent with the rest of the data. (There is a quantitative way to determine if elements are outliers and this will be discussed shortly.) The median, on the other hand would be exactly the same if we had used 125 for the 8th element instead 37. Median is not affected by outliers like the mean is. Because of this we say that the median is **resistant**.

We also need to consider the general pattern of the data when deciding which is better to use. We call this general pattern the **distribution** of the data. In real life, many data sets are roughly symmetric, which would look roughly like the picture below.



Notice that the data looks to be spread out evenly on both sides of the middle. This is a characteristic of a roughly symmetric distribution. In a roughly symmetric distribution, the median and mean will be approximately equal. Alternately, we could have skewed data or data that contains strong outliers. An example of a distribution of this type would be



Notice that the tail extends to the right. We would say that this distribution is **skewed to the right**. Also notice that when the data is skewed, the median will be closer to the peak than the mean. If we were extended the other way, we would say that the distribution is **skewed to the left**.

When we consider these factors, the mean is best used when the data is roughly symmetric and has no strong outliers. The median is better to use when we do not have symmetric data but it could be used in either situation.

6.2 Standard Deviation

When working with the mean, it is often useful to consider the **standard deviation** of the sample. The **standard deviation** is a measure of the average distance of the observations from the mean of the sample. This gives us an idea of how spread out the data is the larger the standard deviation, the further from the mean the average observation is. That is, the standard deviation gives us insight into the dispersion of the data.

It should be noted that the square of the standard deviation, denoted s^2 , is called the **variance**. It is another way to measure the spread of the data. We will not worry about the variance here.

To calculate the standard deviation, we first need to find the mean of the observations in the sample. Then we need to find the square of the difference between each observation and the mean and add these values together. Then we divide by the number of elements (minus 1) and then take the square root. In formula form, this is

$$s = \sqrt{\frac{\sum(\bar{x} - x_k)^2}{n - 1}}$$

where x_k is the value of the k^{th} observation. Here, you have seen all of the symbols besides maybe Σ . This is the capital Greek letter sigma and means summation. What it is telling us to do is add all of the squares. We normally write lower and upper limits with sigma notation, like

$$\sum_{k=1}^n$$

where n is the number of observations and k is the index, but it would have looked too smashed together in the above formula. This is really is not as hard as it seems. Before we get to an example, however, we need to look at why we use squares and square roots.

If we wanted to find the total distance that all of the observations have from the mean, we could just add together the differences between the observations and the mean. Problem here is that some of them will be positive and some will be negative. When the sum of these differences is taken, we would get 0. This is not very useful.

Another way we could go is to use absolute values. There actually is a measure called **absolute deviation** which uses the absolute value instead of the square of the differences. The reason we more commonly use standard deviation, however, is because absolute values are unwieldy to work with when working theoretically. So, in order to guarantee that the distances will all be positive, we take the square of the difference in each case. Then we divide to take the mean of these distances. Finally, the square root is to ‘undo’ the squaring of the distances.

Example 6.2.1 Find the standard deviation of the following set of data.

21, 24, 25, 26, 29, 32, 35

Solution First, we need the mean of the set. We calculated this in the first example and found it to be $\bar{x} \approx 27.43$. Next we will set up a table to organize the many parts of the calculation of the standard deviation.

x_k	$x_k - \bar{x}$	$(\bar{x} - x_k)^2$
21	$21 - 27.43 = -6.43$	41.3449
24	$24 - 27.43 = -3.43$	11.7649
25	$25 - 27.43 = -2.43$	5.9049
26	$26 - 27.43 = -1.43$	2.0449
29	$29 - 27.43 = 1.57$	2.4649
32	$32 - 27.43 = 4.57$	20.8849
35	$35 - 27.43 = 7.57$	57.3049
	Sum	141.7143

Now that we have the sum, we divide by $n - 1 = 6$ and then take the square root to get the standard deviation.

$$s = \sqrt{\frac{141.7143}{6}} \approx 4.8599$$

So, the average distance from the mean of the set of data is a little under 4.9 units.

Example 6.2.2 Find the standard deviation of the daily caloric intake for a person over the course of a week.

$$\{1792, 1666, 1362, 1614, 1460, 1867, 1439\}$$

First we find the mean.

$$\bar{x} = \frac{11200}{7} = 1600$$

Then, we need to find the difference between each of these values and the mean, then square that differences and then sum them.

x_i	$(x_i - \bar{x})^2$	square	contribution
1792	$(1792 - 1600)^2$	192^2	36864
1666	$(1666 - 1600)^2$	66^2	4356
1362	$(1362 - 1600)^2$	$(-238)^2$	56644
1614	$(1614 - 1600)^2$	14^2	196
1460	$(1460 - 1600)^2$	$(-140)^2$	19600
1867	$(1867 - 1600)^2$	267^2	71289
1439	$(1439 - 1600)^2$	$(-161)^2$	25921

Then, we need to find the difference between each of these values and the mean, then square that differences and then sum them.

x_i	$(x_i - \bar{x})^2$	square	contribution
1792	$(1792 - 1600)^2$	192^2	36864
1666	$(1666 - 1600)^2$	66^2	4356
1362	$(1362 - 1600)^2$	$(-238)^2$	56644
1614	$(1614 - 1600)^2$	14^2	196
1460	$(1460 - 1600)^2$	$(-140)^2$	19600
1867	$(1867 - 1600)^2$	267^2	71289
1439	$(1439 - 1600)^2$	$(-161)^2$	25921
		sum	214870

Next, we divide by $n - 1$, which here is 6.

$$s^2 = \frac{214870}{6} \approx 35811.67$$

Now we take the square root.

$$s = \sqrt{35811.67} \approx 189.24$$

So, the average value of the caloric intake is approximately 189 calories from the mean. Notice that we only care about magnitude and not whether we are above or below the mean.

This is certainly a tedious process, especially considering that most situations will have far more than 7 data elements, so it would be nice if we could have another method for finding the standard deviation. Fortunately we can find this easily on the calculator and we will discuss how later in this chapter.

What can we learn from the relationship between the mean and standard deviation?

- We can use them to relate individuals within our data set to the distribution of the sample.
- This is related to probability.
- The total area underneath a distribution curve is always 1, so the area under the curve is the same as the percent of observations falling in the region.
- We will see this better when we get to Normal distributions.

6.3 The 5-Number Summary

When dealing with the median, we often calculate a group of values to help with the analysis. These values are known as the **5-number summary**, which consists of the median, maximum, minimum, first quartile and third quartile. We know what the median is and the maximum and minimum are self explanatory. The quartiles are the median of half of the data; that is, the **first quartile** is the median of the elements that are smaller than the median and the **third quartile** is the median of the elements that are larger than the median. We will do two examples to show how to find the 5-number summary, one each with an even and odd number of elements.

Example 6.3.1 Find the 5-number summary for the following set of data.

7, 15, 34, 24, 20, 25, 22, 28, 23

Solution The first thing we need to do is put the elements in numerical order. This gives us

7, 15, 20, 22, 23, 24, 25, 28, 34

Since there is 9 elements, the median is the 5th element in the list. So, $M = 23$. Since the elements are in order, it is clear that the minimum is 7 and the maximum is 34. We have left only to find the quartiles.

For the first quartile, denoted Q_1 , we consider the values that are smaller than the median only, so we are looking for the median of 7, 15, 20, 22. Since we have an even number of elements here, the median will be the mean of the two central elements, 15 and 20. Their mean is 17.5 and so we say that $Q_1 = 17.5$. Similarly, we consider only values greater than the median to find Q_3 , so we are looking at 24, 25, 28, 34. There again are an even number of elements (there will always be the same number of elements to consider when finding the quartiles) so will be the mean of the two central elements of this set, 25 and 28. This gives that $Q_3 = 26.5$.

Example 6.3.2 Find the 5-number summary for the following set of data.

1, 4, 8, 9, 27, 42

Solution These elements are already in order, so we can obtain the maximum of 42 and minimum of 1 right away. Since there is an even number of elements, the median of the set will be the mean of the two central elements. In this case we are looking for the mean of 8 and 9, which is 8.5. For the first quartile, we are

looking at the median of the elements smaller than the median of 8.5, which are 1,4,8. Since there is an odd number of elements to be considered here, the median is the central element 4, so $Q_1 = 4$. Similarly, the third quartile is the median of the three elements larger than the median, so we need the median of 9,27,42. The value we need is $Q_3 = 27$.

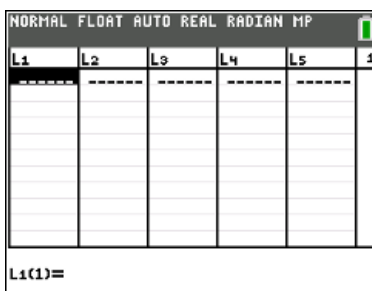
6.4 Statistics Using the TI-Series Calculator

We can also find all of these values using the calculator. The TI-83 and TI-84 has a statistical package that allows us to quickly calculate values like the mean and median and also create statistical plots. We will first look at how to find the values in the calculator and then how to create a couple of visual representations of the data.

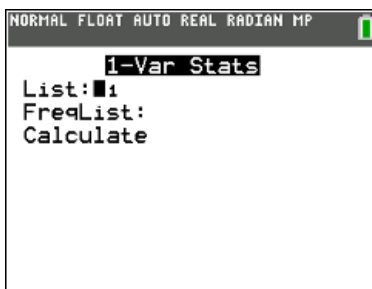
Begin by pressing the **STAT** key. The screen will look like



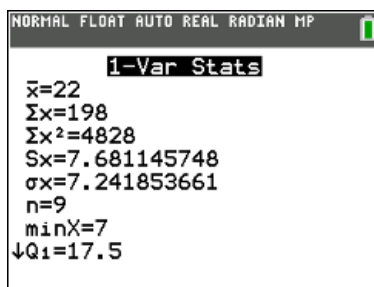
We want to go into the *EDIT* menu. When we do we have



Enter the set of data under L_1 . We will use the data from an earlier example to illustrate this. That data set again is 7,15,34,24,20,25,22, 28,23. You do not have to order the data first, however. The calculator will take care of that. After you finish, press *2nd* and **QUIT** to get back to the blank screen. If we don't, the calculator will think we are trying to enter data into a cell. Next, press **STAT** again and scroll to *CALC*. What we want is 1-Var Stats, which stands for statistics for one variable. Press **ENTER** on this or press the **1** key (because it is the first option) and the screen will look like

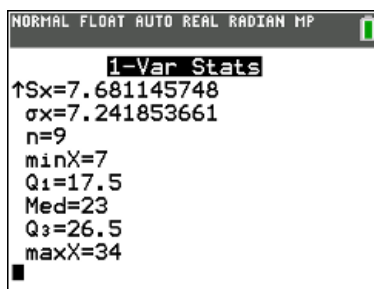


If you entered your data under L_1 then you just need to scroll to ‘Calculate’ **ENTER** again. If you did not put the data under L_1 then you need to specify which list you want the statistics calculated for by pressing *2nd* and then one of the keys 1-6 (the one that corresponds to the list you used). When you eventually press **ENTER** the screen have some statistics that will look familiar and some that will not. The mean, \bar{x} , is the same as what we would have obtained (if we would have found it for this set of data) and also two different standard deviations. The one that we are interested in is the sample standard deviation, denoted as S_x . The difference between the two is that the sample standard deviation has a denominator of $n - 1$ and the population standard deviation has a denominator of n .



In order to understand why there are two standard deviations, we need to understand the difference between a sample and a population. A **sample** is a collection of data for which we can calculate our basic statistics. A **population** would be the entire collection of observations that cannot be used for calculations. For example, if we were doing a study on the class average for students taking Statistics, our class could be a sample for the population and all students across the country taking a Statistics class would be the population. We can use the statistics from our class to project what would happen if we considered all Statistics classes (with varying degrees of accuracy) but it would not be reasonable to calculate the average for every Statistics student in every class in every college and university in the entire country. When we have all of the possible observations, we can simply use n in the calculation of the standard deviation. But when we are using a sample to make projections for the whole population, we divide by a smaller number to account for the error.

If we use the arrows to scroll on the calculator screen, we see another set of numbers. This set is the 5-number summary that we were calculating earlier.



Notice that these are the same values we found by hand.

6.5 Percentiles

We can measure the position in terms of the percent of elements that are below the value in question. This is the use of percentiles. There is one problem, however, which is that there is no universal definition of

percentile.

Definition 6.5.1 *The x^{th} percentile is the lowest score that is greater than $x\%$ of the scores.*

For example, the 80th percentile out of 100 scores would be the 81st score.

Definition 6.5.2 *The x^{th} percentile is the lowest score that is greater than or equal to $x\%$ of the scores.*

The 80th percentile here would be the score 80.

Note: We often talk of being *in the x^{th} percentile* because we generally have many values in a small window when looking at enough data. We think of all of the scores greater than the 80th percentile but less than the 81st percentile as being in the 80th percentile.

The third definition is more complicated, more accurate and will seem easier to follow in terms of an example.

Example 6.5.3 *Suppose we wanted to find the 25th percentile of the scores given here:*

<i>Number</i>	3	5	7	8	9	11	13	15
<i>Rank</i>	1	2	3	4	5	6	7	8

We next need to calculate the rank (R) of the 25th percentile.

Formula 6.5.4 *Rank R is calculated using the following formula:*

$$R = \frac{P}{100} \times (N + 1)$$

where N is the number of data elements and P is the desired percentile.

Here, we have

$$R = \frac{25}{100} \times (8 + 1) = \frac{9}{4} = 2.25$$

If R is an integer, the P^{th} percentile is the number with rank R .

If R is not an integer, then ...

- Define IR as the integer portion of R . Here, $IR = 2$.
- Define FR as the fractional portion of R . Here, we have $FR = \frac{1}{4} = .25$
- Find the scores with rank I_R and $I_R + 1$. Here, since $R = 2$, we want the second and third scores, so 5 and 7.
- Now, multiply the difference between the scores by FR and add the lower of the two scores. This will give the desired percentile. So, we have $.25(7 - 5) + 5 = 5.5$

So, here, the smallest score greater than 5.5 would be 7. This is a different result than we would have gotten from the second definition, but is a much more appropriate way to calculate the percentile from a large data set.

Example 6.5.5 *Colleges and universities consider class rank as one of the factors in determining whether or not to admit a student and being in the 90% percentile is desirable. Suppose a graduating class from a local high school had 220 students. What is the cut-off for the 90th percentile?*

Solution First, we begin with the rank.

$$R = \frac{90}{100} \times 221 = 198.9$$

Since this is not an integer, we have $IR = 198$ and $FR = .9$. We would then find the 198th and 199th highest grade point averages from the class. Let's say that the GPA's here are 3.625 and 3.638. Then, we need to find their difference and use that to scale the lower score.

$$.9(3.638 - 3.625) + 3.625 = 3.6367$$

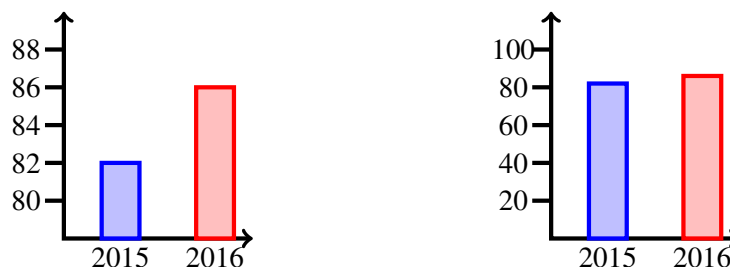
So, the lowest GPA greater than 3.6367 would be the cut-off for the 90th percentile.

Using the other definitions, we could have much more quickly identified the student with either the 198th or 199th highest GPA as the cut-off score, but this is a much more formal way to calculate and gives us a better justification than 'well, it was the closest to ...'

6.6 Misuses of Statistics

We have to be careful when using statistics as the perception and the reality are not always the same. We can use the statistics to our advantage if we know how to manipulate them and this is routinely done by the media. The following are some examples of statistics being manipulated.

Days with Temperatures Above 90° F Days with Temperatures Above 90° F



Both of these graphs represent the same data in the same situation. But, the first graph makes it appear that there is a much greater difference between the days of temperatures exceeding 90°. This is because of the vertical scale - the one on the right has the gaps between hash marks at the same width, but on the left, the gap between the origin and 80 days is the same size as that between each of the other totals even though this first width is supposed to represent 80 days and the others represent 2. This dramatically changes the perception for the casual observer.

The graph below has a different kind of distortion. Notice that Item C represents the smallest wedge of the pie chart, but your eye is drawn to this section because of the way the graphic is drawn. Color and effects like this draw the attention of the reader instead of focusing on the actual data that is being represented.

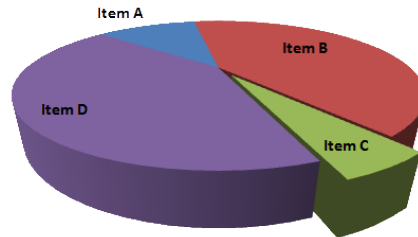


Image from <https://www.boundless.com/statistics/>

This next graph is also a pie chart but it was poorly constructed. The data represented is supposed to point out which political candidate is being backed, but the percents do not add up to 100%. It makes it so that we have no idea what is the actual percentage of the population that supports each candidate.

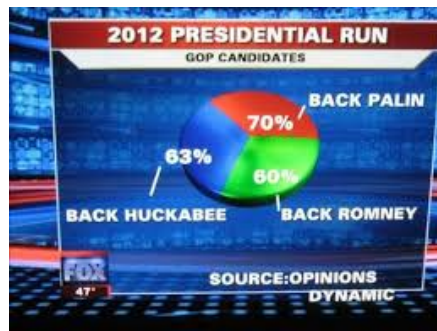


Image from <http://flowingdata.com/2009/11/26/>

These last two graphics epitomize the issues with pie charts. They are very easy to manipulate and they don't really give us any information more than would a numeric summary of the data. But because they are simplistic and colorful (in most cases), they are popular in presentations and with the media.

The next graph is a line graph that has multiple distortions. First, if we found the best-fit curve for this data, it would not be linear, yet a linear function was used to express the growth in the job loss statistic. Also, the height of the curve is not accurate in relation to the value represented; the height of the line at 15 million is not a little more than twice as tall as the point representing 7 million, but rather it is closer to 5 times larger. But with no scale given on the vertical axis, we have no way to gauge the accuracy of heights.



Image from <http://www.kdnuggets.com/2012/12/taking-misleading-statistics-to-a-new-level.html>

The next graphic also represents a skewing of the data because of the scale. The bars should be starting at 0 and the heights should be proportional based on pitch speed they represent. But, they do not do this; instead a difference of 2 MPH has the look of twice the speed.



Image from <http://i.imgur.com/arWY8Hb.jpg>

This final graph distorts the data in two ways as well. They are clearly trying to emphasize that 2012 was significantly warmer than the other years with high temperatures. They do not start the vertical axis at 0, thereby making the heights of the bars not accurate. They also make only the last bar in a different color - and red at that - which draws your eye to the dramatic difference in height rather than that the temperature difference is less than one degree.

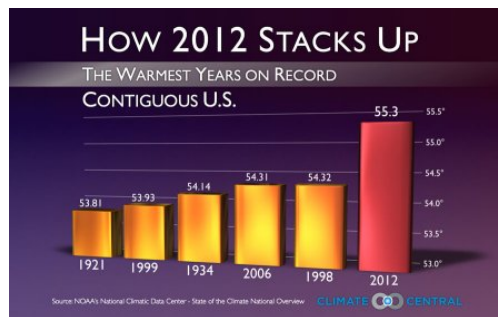


Image from <http://themonkeycage.org/2013/01/how-2012-stacks-up-the-worst-graph-on-record/>

Statistics can be manipulated in other ways besides graphics. Consider the following table for the grade distribution for a particular section of MAT 128:

Class Average	Relative Frequency
A	17%
B	21%
C	24%
D	18%
F	12%
W	8%

We could use different aspects of the data to discuss what we wanted to emphasize, depending on the audience. For example, in looking at this data with colleagues, it might be pointed out that 38% got a D or worse in the class and we could talk about ways to lower this percentage. On the other hand, when talking to administration, we could talk up the fact that 38% of students received at least a B in the class. Same data, very different points of discussion. We have to be careful and responsible when discussing data and statistics so that we are not being intentionally (or unintentionally) misleading.

6.7 Exercises

1. Carl Yastrzemski played for the Boston Red Sox from 1961-1983 and is a Hall of Famer. His home run totals from his 23 year career are as follows¹:

{11, 19, 14, 15, 20, 16, 44, 23, 40, 40, 15, 12, 19, 15, 14, 21, 28, 17, 21, 15, 7, 16, 10}

Using this data set, find the following:

- (a) The mean \bar{x}
- (b) The sample standard deviation s
- (c) The median M
- (d) The 5-number summary

Be sure to properly label your answers.

2. The **gross national income** per capita, denoted GNI, is the dollar value of a country's final income in a given year divided by the size of its population. Given below is the GNI for the United States, since 1960, where dollar amounts are in trillions of dollars.²

.5464	.5668	.6092	.6431	.6907	.7490
.8201	.8671	.9486	1.0260	1.0769	1.1659
1.2839	1.4351	1.5569	1.6887	1.8740	2.0870
2.3550	2.6193	2.8528	3.2072	3.3747	3.6211
4.0383	4.3208	4.5304	4.8472	5.2758	5.6183
5.9229	6.1173	6.4595	6.7584	7.1958	7.6023
8.0754	8.6203	9.1676	9.7253	10.4213	10.7886
11.0989	11.5914	12.3726	13.2218	14.1408	14.5857
14.7912	14.4945	15.1211	15.8029	16.5961	17.0912
17.5861	18.1383				

Using this data set, find the following:

- (a) The mean \bar{x}
- (b) The sample standard deviation s
- (c) The median M
- (d) The 5-number summary

Be sure to properly label your answers.

3. The average price of gas (in dollars per gallon) for the United States as of September 2016 are as follows:³

1.965	1.973	1.976	1.987	2.000	2.003	2.009	2.050	2.076	2.084
2.085	2.096	2.106	2.127	2.137	2.137	2.140	2.143	2.152	2.155
2.171	2.182	2.185	2.192	2.194	2.197	2.208	2.213	2.222	2.222
2.225	2.228	2.246	2.248	2.254	2.258	2.264	2.279	2.311	2.323
2.335	2.364	2.375	2.421	2.491	2.525	2.595	2.711	2.766	2.824

Using this data set, find the following:

- (a) The mean \bar{x}
- (b) The sample standard deviation s
- (c) The median M
- (d) The 5-number summary

Be sure to properly label your answers.

¹Data obtained from <http://www.baseball-reference.com/players/y/yastrca01.shtml>

²Data obtained from data.worldbank.org

³Data obtained from <https://www.gasbuddy.com/USA>

6.8 Solutions

1. (a) $\bar{x} = 19.65$ home runs
(b) $s = 9.72$
(c) $M = 16$ home runs
(d) $\min = 7, Q_1 = 14, M = 16, Q_3 = 21, \max = 44$
2. (a) $\bar{x} = \$6.603$
(b) $s = \$5.561$
(c) $M = \$5.0615$
(d) $\min = \$5.5464, Q_1 = \$1.496, M = \$5.0615, Q_3 = \$10.94245, \max = \$18.1383$
3. (a) $\bar{x} = \$2.2289$
(b) $s = \$1.1926$
(c) $M = \$2.1955$
(d) $\min = \$1.965, Q_1 = \$2.106, M = \$2.1955, Q_3 = \$2.279, \max = \$2.824$

Chapter 7

Visual Representations of Data

There are two main types of variables: categorical and quantitative. The difference between them involves the type of data we can collect and how we can visually represent that data.

Definition 7.0.1 A categorical variable takes on values that could be selected from a finite list. It does not make sense to perform arithmetic operations on a categorical variable.

Definition 7.0.2 A quantitative variable taken on values that are numeric in nature and it makes sense to perform arithmetic operations, like means and totals.

7.1 Pie Charts

Pie Charts, or circle graphs, represent categorical variables. We will discuss later the advantages and disadvantages of using this type of graph, but first let's look at how to create them.

Creating Pie Charts

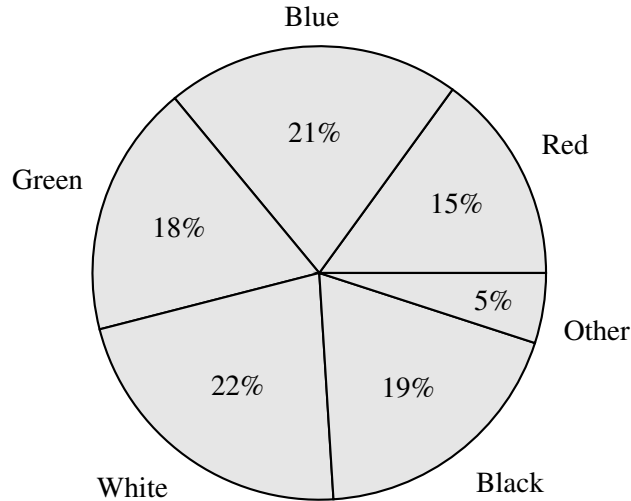
Example 7.1.1 You sit on an overpass and record the color of the first 100 cars you see. The results are as follows:

<i>color</i>	<i>frequency</i>
<i>red</i>	15
<i>blue</i>	21
<i>green</i>	18
<i>white</i>	22
<i>black</i>	19
<i>other</i>	5

Construct a pie chart to illustrate the relationship between the colors of these cars.

We have to make sure that the size of each slice is correct in relation to the other slices. To do this, we make sure the central angle is the correct size. Since we have 100 observations, the number of observations is the percent of the circle we need for that wedge. For example, for the red cars, we saw 15 of them, so we would need a central angle of $.15(360^\circ) = 54^\circ$.

Colors of Cars



Example 7.1.2 *The following is a breakdown of the solid waste that made up America's garbage in 2000. Values given represent millions of tons.*

<i>Material</i>	<i>Weight</i>
<i>Food</i>	<i>25.9</i>
<i>Glass</i>	<i>12.8</i>
<i>Metal</i>	<i>18.0</i>
<i>Paper</i>	<i>86.7</i>
<i>Plastics</i>	<i>24.7</i>
<i>Rubber</i>	<i>15.8</i>
<i>Wood</i>	<i>12.7</i>
<i>Yard Trimmings</i>	<i>27.7</i>
<i>Other</i>	<i>7.5</i>

Create a pie chart to represent this data.

We can't make a pie chart with this data; at least not yet. We need to first find the relative frequencies because we need our quantities to represent 100%. To find the relative frequency, we need to divide the weight in each category by the total weight and then multiply by 100 to get the percent representation. For example, for food, we have

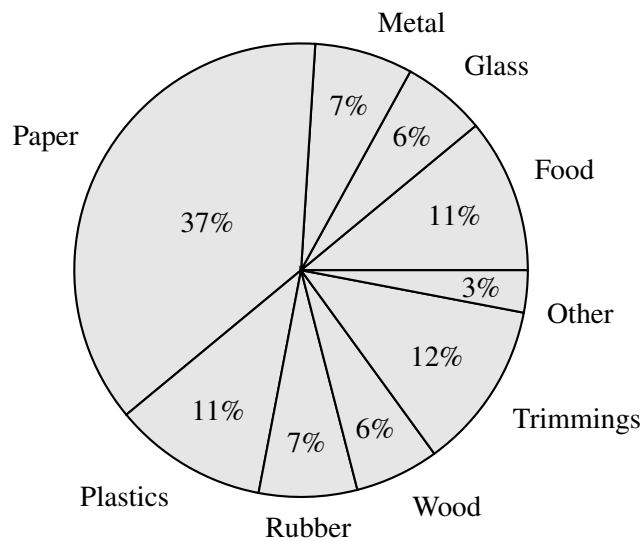
$$\frac{25.9}{231.9} = .112 \times 100 = 11.2\%$$

We use this method to find all of the relative frequencies.

Material	Weight	Relative Frequency
Food	25.9	11.2 %
Glass	12.8	5.5%
Metal	18.0	7.8%
Paper	86.7	37.4%
Plastics	24.7	10.7%
Rubber	15.8	6.8%
Wood	12.7	5.5%
Yard Trimmings	27.7	11.9%
Other	7.5	3.2%
Total	231.9	

Now we can find the central angles and create our pie chart. Using the same approach as before (rel freq in decimal form $\times 360^\circ$), we get

Material	Weight	Relative Frequency	Central Angle
Food	25.9	11.2%	40.3°
Glass	12.8	5.5%	19.8°
Metal	18.0	7.8%	28.1°
Paper	86.7	37.4%	134.6°
Plastics	24.7	10.7%	38.5°
Rubber	15.8	6.8%	24.5°
Wood	12.7	5.5%	19.8°
Yard Trimmings	27.7	11.9%	42.8°
Other	7.5	3.2%	11.5°



7.1.1 Drawbacks to Pie Charts

There are some issues with pie charts.

1. We must use relative frequencies

2. It is just as easy to read the frequency table as the pie chart
3. Only good for categorical variables
4. Not easy to compare two variables
5. Easy to manipulate
6. Be careful that all percentages are calculated the same way (i.e. the same denominator)

Pie charts are used to compare different categories in relation to each other. It is only good for one type of variable; we cannot do comparisons between different types of observations with one pie chart.

7.2 Bar Graphs

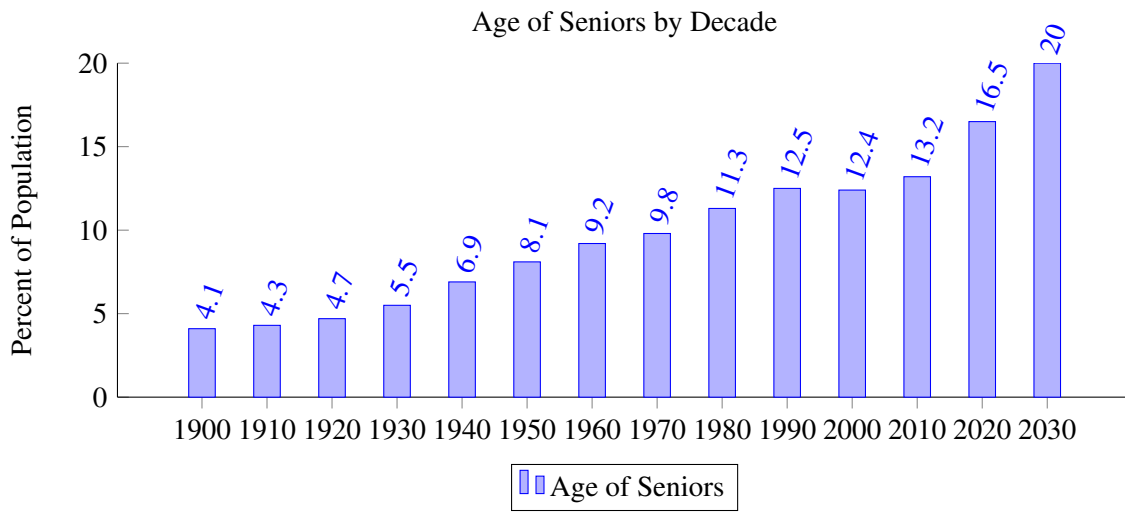
Bar graphs basically give us the same information as a pie chart, with a couple advantages. We will get to those after the construction.

7.2.1 Creating Bar Graphs

The most important part of a graph like this is the width of the bars; what determines which category has more is the area of the rectangle representing the class. If the bars all have the same width, we merely need to see which bar is taller to see which category is largest.

Example 7.2.1 *The growth of the US population age 65 and over is given in the table. Create a bar graph to represent this data.*

1900	4.1	1970	9.8
1910	4.3	1980	11.3
1920	4.7	1990	12.5
1930	5.5	2000	12.4
1940	6.9	2010	13.2
1950	8.1	2020	16.5
1960	9.2	2030	20.0

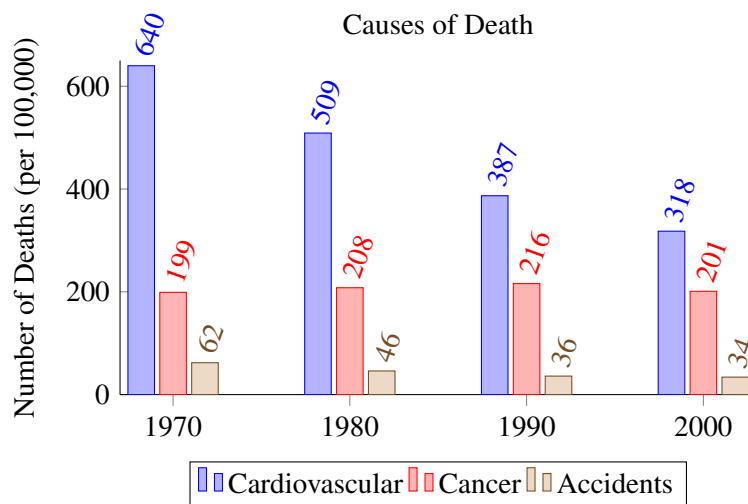


Notice that we can't do much analysis here other than see which class has the most. We don't even have to put the bars in any kind of order; if we did by size, we'd have a pareto graph. But since order does not matter, we cannot talk about the distribution the same way we will be able to for quantitative variables.

Here is how we can use bar graphs for comparison purposes.

Example 7.2.2 Create a bar graph for the given causes of death and analyze the results. Values given are the number per 100,000 people.

Cause of Death	1970	1980	1990	2000
Cardiovascular	640	509	387	318
Cancer	199	208	216	201
Accidents	62	46	36	34



- Cancer and accidents are roughly the same in each decade
- Cardiovascular disease decreases each decade and is approaching level of cancer deaths

7.2.2 Usage of Bar Graphs

There are some advantages to using bar graphs over pie charts.

- We can use raw frequencies as all that matters is the size of the rectangle
- We can compare multiple variables
- Bars can be vertical or horizontal

But, there are a couple of drawbacks as well.

- We generally need to represent categorical variables with bar graphs
- We cannot analyze distribution because the order of the classes is not necessarily in numerical order

7.3 Distributions

Before we discuss graphs that are appropriate to create using quantitative data, we need to make sure we understand both what is a distribution and also how to analyze the distribution of a graph.

Definition 7.3.1 A *distribution* is a representation of data vs. frequency. It shows all possible values and how often they occur.

Now we want to concern ourselves with the analysis of the graphs. We can analyze these in a much more constructive way that we could with the graphs of categorical variables. Here we are analyzing the distribution represented by the graph.

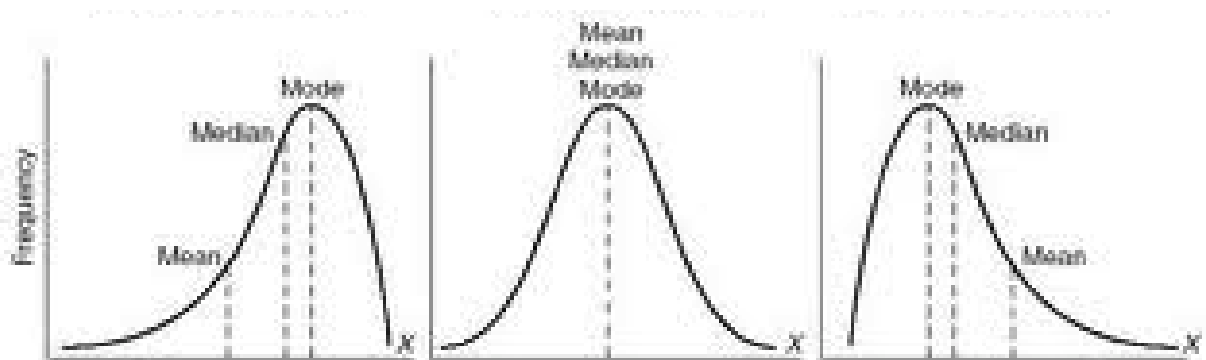
1. Center

The center can be one specific element or could be the class that contains the middle element. It depends on the way the data is represented, as we will see in the analysis of each graph.

2. Shape: Number of peaks, skewness

The following picture shows the relationship between the mean, median and mode in relation to the skewness of the graph.

- If a graph is roughly symmetric, the mean, median and mode are all roughly the same. There is approximately the same amount of data on either side of these statistics
- If the graph is skewed left, the tail stretches to the left and the peak is on the right
- If the graph is skewed right, the tail stretches to the right and the peak is on the left



We also include in the discussion of shape how many peaks the graph has.

- If the graph has one peak, we say it is unimodal

- If the graph has two peaks, we say it is bimodal
- For more than two distinct peaks, we say the graph is multimodal

We do have to be careful when defining a peak. We are looking for pronounced regions of the graph where there is data values or frequencies much larger than surrounding values when discussing the presence of multiple peaks.

3. Spread

Generally, when we are looking at the spread, we are asking about the range of the data. This is expressed as

$$\text{range} = \text{max value} - \text{min value}$$

Now that we have our means of analysis, let's look at the graphs suited for quantitative variables.

7.4 Dot Plots

Definition 7.4.1 A dot plot is the representation of a set of data over a number line. The number of dots over a number represents the relative quantity of the value.

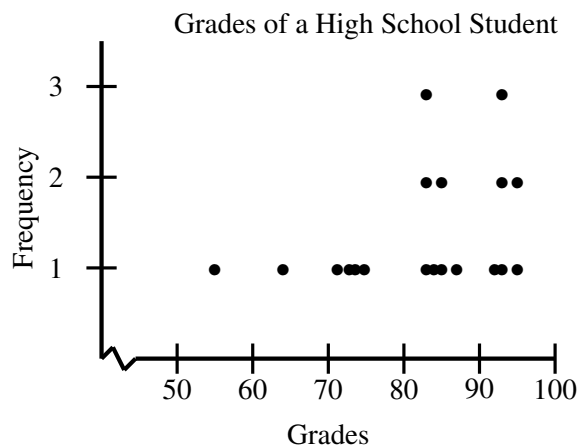
7.4.1 Creating Dot Plots

Let's begin with discussing how to create these graphs and then we will look at usage, as we did before.

Example 7.4.2 The following are the test scores for a particular high school student in their math class over the course of an academic year.

64	73	85	74	83
71	56	83	76	85
83	87	92	84	95
92	95	92	91	

To produce a dot plot, we are essentially plotting points where the height of a column is based on the frequency of the element in the data set.



- Range: Highest value - lowest value
Here, the range would be $95 - 56 = 39$.
- Center: The central value(s) is the center. It could be a value or a class, depending on the type of graph.
Here, the center is the 10th value, since there are 19 data points in the set. The value we seek is 84.
- Shape: How many peaks are there? Is it roughly in the middle or to one side?
Here we have one peak, so we would say the distribution is unimodal. That peak is to the right, so the tail stretches out to the left. We would say this graph is left skewed.

7.4.2 The Pros and Cons

There are a couple of good points about using dot plots.

- Gives a good idea of distribution
- Preserves all of the data points

But, there are a few issues with using this type of graph as well.

- Tedious to plot
- Can be hard to read
- Not practical for large data sets

7.5 Stem-and-Leaf Plots

Stem-and-leaf plots are similar to dot plots in that we preserve the data by representing each data point. But this is exactly one of its drawbacks ...

7.5.1 Creating Stem-and-Leaf Plots

Example 7.5.1 *Using the same data set as we did for the dot plot, construct a stem-and-leaf plot.*

First thing we need to do is order the data elements to make it easier for us to plot the data.

56	64	71	73	74
76	83	83	83	84
85	85	87	91	92
92	92	95	95	

Now, we have to create the ‘stem’. We do this by putting all but the last digit on the left side of the vertical bar. We need to represent all possible values from the minimum to the maximum and we cannot skip any values.

Grades for a High School Student

9	
8	
7	
6	
5	

Now, we list all of the last digits in increasing order from the stem on the right side. We do not use commas to separate the values and if a value appears more than once in the data set, we list it once for each occurrence.

Grades for a High School Student

9	1	2	2	2	5	5	
8	3	3	3	4	5	5	7
7	1	3	4	6			
6	4						
5	6						

Here we get the exact same answer for the range and the center, although we only give the class in which the center lies, so we would say that the center is in the 80's. We get that the shape is again unimodal and skewed left. It may look different, but since it represents the same distribution, we expect similar answers.

Notice that the values on the right are essentially in columns - this is what allows us to quickly see which classes have more elements.

What if we had a 3 digit number? Suppose the student got a 100 on the next exam? We would need to add a row.

Grades for a High School Student

10	0						
9	1	2	2	2	5	5	
8	3	3	3	4	5	5	7
7	1	3	4	6			
6	4						
5	6						

One advantage of this type of graph is that it can be used for comparison purposes, provided the variable being compared measure the same type of data.

Example 7.5.2 *Suppose we wanted to compare the careers of Babe Ruth and Mark McGwire in terms of their yearly home run totals to determine which player was the more consistent long ball hitter. Make a back-to-back stem-and-leaf plot to make the is determination.*¹

<i>Ruth</i>	54	59	35	41	46	25	47	60	54	46	49	46	41	34	22
<i>McGwire</i>	49	32	33	39	22	42	9	9	39	52	58	70	65	32	29

¹Data obtained from baseballreference.com

7.6.1 Creating Histograms

Example 7.6.1 For our test scores example, construct a histogram and analyze the distribution.

It is easier if the values are in order as we will be grouping them into classes.

56	64	71	73	74
76	83	83	83	84
85	85	87	91	92
92	92	95	95	

We first want to create a frequency table. This is a collection of non-overlapping classes and the frequency of observation in each of those classes. We need to determine the following in this order:

Number of classes

The rule of thumb with the number of classes is to use the square root of the number of observations in the data set.

$$\sqrt{19} \approx 4.36$$

So, we can use 4 or 5 classes. We tend to go up to the next integer to be sure I have enough classes. So we will use 5 for our graph.

Size of each class

We want them to be the same width so that the taller classes will be known to have the most elements. If not then we have to find the area of each rectangle to determine relative size.

To find the size, we divide the 'range' by the number of classes.

$$size = \frac{95 - 56 + 1}{5} = \frac{38}{5} = 7.6$$

We could use 7.6 for the class width or we can go to the next largest integer. Where we may have extra if we round up, it is better than not having enough of a range in the classes to cover all of the data. For the sake of simplicity, we will use 8.

Endpoints of each class

We start the smallest class with a left endpoint of 56, since that was our minimum. Then, to find the next left endpoint, add 8 to 56. Continue in this manner until we have 5 classes.

Grade Range	Frequency
56-	
64-	
72-	
80-	
88-	

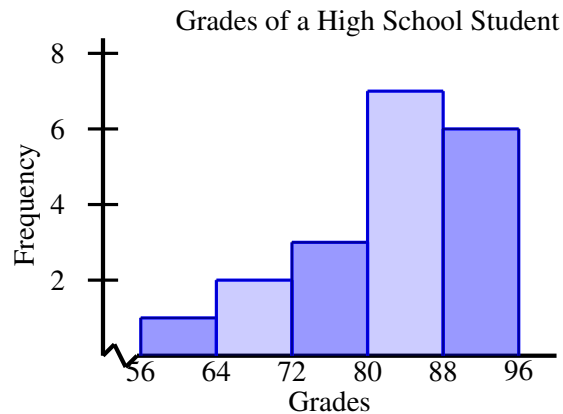
Then, we subtract 1 from each left endpoint to find the right endpoint of the previous class.

Grade Range	Frequency
56-63	
64-71	
72-79	
80-87	
88-95	

Finally, we count how many elements go in each class.

Grade Range	Frequency
56-63	1
64-71	2
72-79	3
80-87	7
88-95	6

Now, let's look at the visual.



We see the same range and shape. Here, we'd have no choice but to give the class only for the center as we would lose the ability to see individual values.

Example 7.6.2 For the following set of data, create a histogram.

2, 3, 3, 4, 4, 4, 5, 6, 7, 7, 8, 8, 8, 8, 9

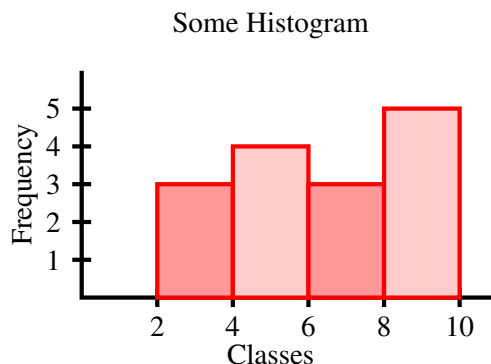
Solution There are 15 data elements, so the number of classes for our histogram will be $\sqrt{15} \approx 4$. Next, we need to know how big each class should be. $9 - 2 + 1 = 8$ and $\frac{8}{4} = 2$, so each class will be of size 2. Now we create the frequency table.

Class	Frequency
2-3.9	
4-5.9	
6-7.9	
8-9.9	

Now, we count.

Class	Frequency
2-3.9	3
4-5.9	4
6-7.9	3
8-9.9	5

And finally, we graph.

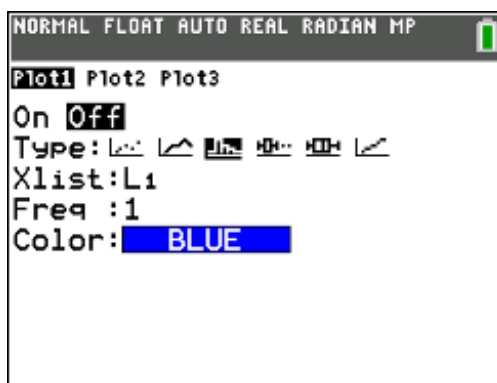


We see here a unimodal, skewed left graph with a range of 2-10 and a center in the class from 6-7.9.

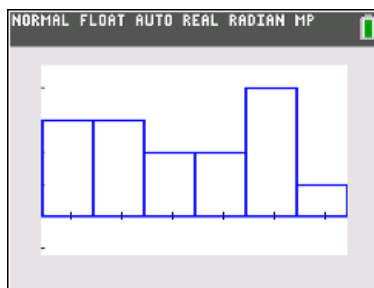
7.6.2 Histograms on the TI-Series Calculator

The calculator can create histograms as easily as it can create other graphs. We don't have to worry about class sizes, number of classes, etc. as this will all be done for us. We need only worry about making sure we correctly input the data and select the graph we want.

To create a histogram, open the *STATPLOT* menu again by pressing **2nd** and **Y=**. Note that if you are going to use a different plot than the last time you made a plot, you need to shut off the other plot. If you have more than one on, you will get an error. (On that note, if any of the *STATPLOTS* are on and you try to make a regular graph, you will also get an error.) After you select whichever plot you want to use, the histogram is the third graph.



Once you change the options to be what you want, you can again press **ZOOM** and **9**. The diagram will give a rectangle to represent each class. In order to see the frequency in each class or the class boundaries, press the **TRACE** key. The graph will appear as



This indicates that the first class ranges from 7 to just under 13.75 and there is one element in the class. As you scroll through the histogram, you will notice that all of the classes are the same width. Again, this is a necessity for a histogram because the relationship between the classes is compared by the area of the rectangles. The only way to make a direct comparison is if the rectangles have the same width and varying heights. The taller the rectangle, the higher the frequency of the class. The calculator also automatically determines the standard size for the classes.

7.7 Box Plots

A **box plot** (also known as a box-and-whisker plot) is a visual representation of the 5-number summary.

Example 7.7.1 *Ted Williams yearly RBI totals:*

145, 113, 120, 137, 123, 114, 127, 159, 97, 126, 3, 34, 89, 83, 82, 87, 85, 43, 72

Find the 5-number summary for this set of data,

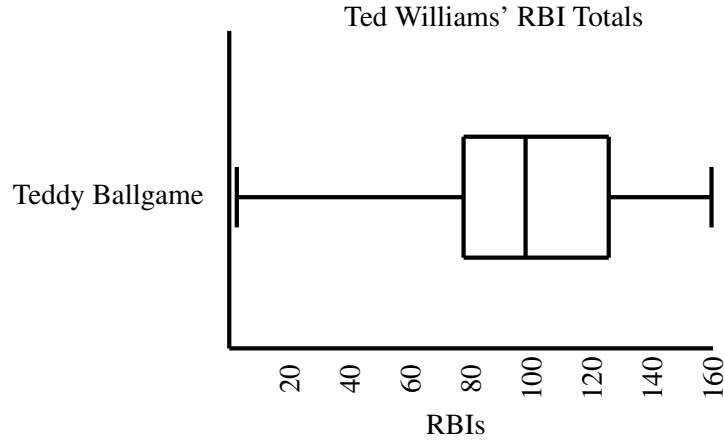
First, we put the values in order.

3, 34, 43, 72, 82, 83, 85, 87, 89, 97, 113, 114, 120, 123, 126, 127, 137, 145, 159

Then, we find the 5-number summary as we did before.

Minimum	3
Q_1	82
Median	97
Q_3	126
Maximum	159

To construct the box-and-whisker plot, we put vertical lines at each of the values from the 5-number summary, with the middle 3 lines being a bit longer (see below). Then, we put horizontal lines connecting Q_1 and Q_3 . This represents the middle 50% of the data and is where the ‘box’ part of the name comes from. We then add the ‘whiskers’, which are horizontal lines connecting the minimum and maximum to the box. Each of these represent 25% of the data.

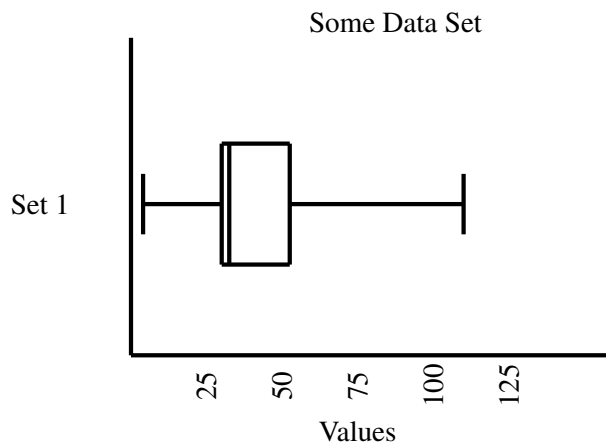


In analyzing a graph like this, we are a bit limited. We can see that the middle 50% is close to symmetric, but the left tail is much longer than the right tail. So, we can see that this is a skewed left distribution and we can see the range. But we cannot tell how many peaks there are. We can see, however, where the center is, as we have our middle vertical line as the median.

Example 7.7.2 Construct a box-and-whisker plot given 5-number summary.

<i>minimum</i>	4
Q_1	30
<i>Median</i>	33.5
Q_3	52.5
<i>maximum</i>	110

Solution



Our range here is 4-110. We know that the distribution is skewed right as well. We can also see that the right endpoint seems to be pretty far away, so we may think it is an outlier. But how do we determine if it is analytically? We need the IQR Criterion.

Definition 7.7.3 The *IQR Criterion* is an analytic way for us to determine if data points are outliers based on a 5-number summary. To determine outliers, we use

$$Q_1 - 1.5IQR$$

and

$$Q_3 + 1.5IQR$$

to give us endpoints, known as fences, of the acceptable data range, where *IQR* is the Interquartile Range and

$$IQR = Q_3 - Q_1$$

So, basically what we are doing is saying that any values no further away from the middle 50% than 1.5 times the range of the middle 50% are acceptable. Anything outside that range is an outlier.

To determine if there are any outliers, we first find the IQR.

$$Q_3 - Q_1 = 52.5 - 30 = 22.5$$

Now, we find the fences.

$$Q_1 - 1.5IQR = 30 - 1.5(22.5) = 30 - 33.75 = -3.75$$

$$Q_3 + 1.5IQR = 52.5 + 1.5(22.5) = 52.5 + 33.75 = 86.25$$

Since 110 is larger than this upper threshold, we would say it is an outlier.

Notice that we built this box-and-whisker plot only from the 5-number summary. We know nothing about the other data elements, including quantity, so this severely limits what we can do with box-and-whisker plots without having more information than just this summary.

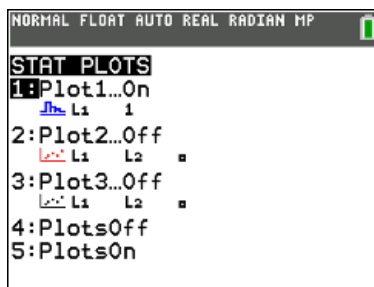
Example 7.7.4 Determine if there are any outliers in the set of data below.

7, 15, 34, 24, 20, 25, 22, 28, 23

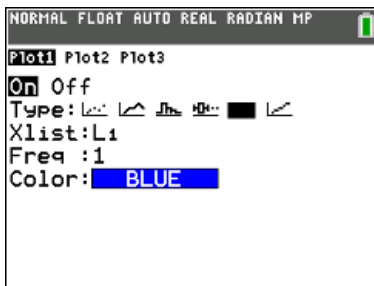
Solution We see that $Q_1 = 17.5$ and $Q_3 = 26.5$. This gives an inter-quartile range (IQR) of $26.5 - 17.5 = 9$ and $1.5(IQR) = 1.5(9) = 13.5$. Since Q_1 is 17.5, any values below $17.5 - 13.5 = 4$ would be outliers and since Q_3 is 26.5, any values above $26.5 + 13.5 = 40$ would be outliers. There are no values in these ranges, so this set of data contains no outliers.

7.7.1 Box-and-Whisker Plots on the TI-Series Calculator

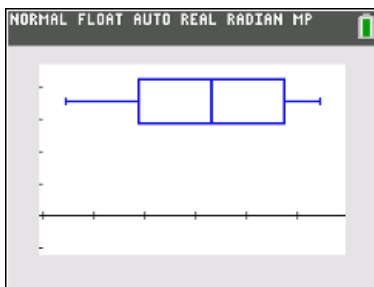
Once the data is in the calculator, we can produce a box plot using the statistical plotting capabilities of the calculator. To get to the *STATPLOT* menu, press **2nd** and **Y=**. What you should see is



Now open whichever plot you want. When you do the screen will have a number of choices. You need to use the arrows to scroll to **On** and press **ENTER**. Then select which plot you want. There are two box plot options, one that indicates outliers and one that does not (we will determine outliers numerically soon). The Xlist needs to be whichever list the data in question is in.

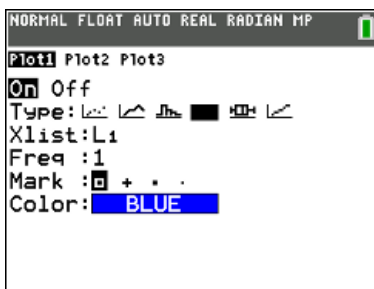


Once you have edited these options in whatever way you see fit (scroll to whichever box plot you want and press **ENTER**), press **ZOOM** and then **9** (Zoom Stat).



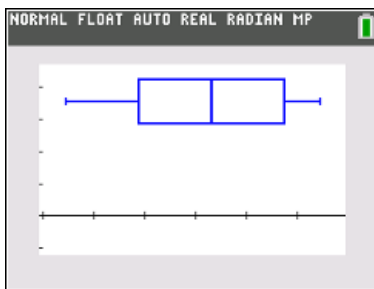
This plot gives us an idea of how the data is distributed. It shows us where half of the data lies (between the first and third quartiles is the middle 50% of the data) and it gives an idea of how far away from the median the extreme values are. Plots of this type are more useful in comparing the ranges or medians of two sets of data. We can get the exact values of the 5-number summary from the plot, though, by pressing the **TRACE** key and then using the arrows to scroll throughout the diagram.

We can also determine if there are any outliers using the calculator. The other option for box plots indicates outliers in the plot.



When we choose this box plot option, we are given a choice along the bottom for the type of mark we want to indicate outliers. The general shape of the box plot will be the same, but the lines that extend from the central rectangle will be shorter if there are outliers and the data elements that are outliers will be indicated by one of these symbols. If there are no outliers then the box plot will look exactly the same regardless of

which plot is chosen. In this case, they will be the same, as can be seen below, because we calculated that there are no outliers in 7.7.4.



If there are outliers, however, we can use the **TRACE** command and the arrows to scroll to the outliers to determine the values.

7.7.2 Comparisons with Box-and-Whisker Plots

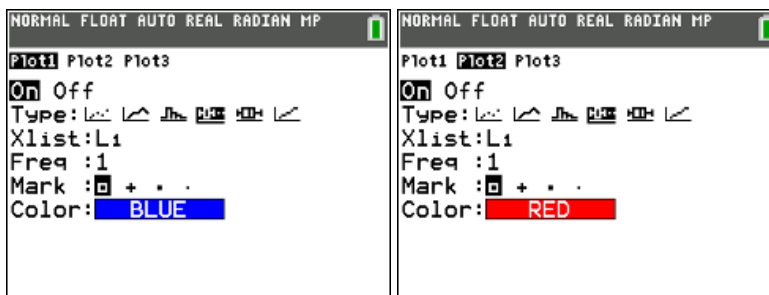
Box plots are best used for comparisons of sets of data. We can produce multiple box plots on the calculator, as can be seen in the next example.

Example 7.7.5 Create box plots on the same set of axes for the following sets of data.

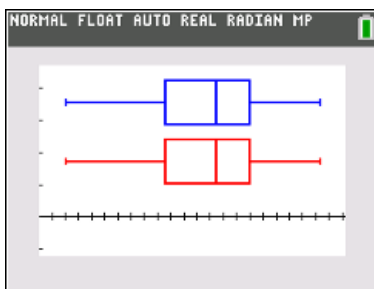
$$S = \{7, 15, 34, 24, 20, 25, 22, 28, 23\}$$

$$T = \{3, 6, 1, 8, 23, 21, 36, 23, 22\}$$

Solution Put both sets of data in the calculator the set S under L_1 and the set T under L_2 . Then enter the *STATPLOT* menu as before. Under *PLOT 1*, put in the specifics you'd like for the box plot of your choosing. Be sure that this plot is using L_1 as the *Xlist*. Then use the arrows to scroll over to *PLOT 2*. Put in the specifics for the box plot as you'd like (ideally the same as what you did for the first box plot) except the *Xlist* needs to be changed to L_2 . The screens should look something like this:



Then, as before, press **ZOOM** and **9** to get the plot.



7.7.3 Exercises

For each of the data sets in problems 1-3, do the following:

- Find the mean and standard deviation.
- Find the 5-number summary.
- Determine if there are any outliers using the IQR criterion.
- Produce a box plot.
- Produce a histogram.

Discuss the distribution in terms of center, shape and spread.

- $S = \{3, 5, 7, 9, 13, 22, 27, 31, 33, 38, 41, 42, 42, 45\}$
- $T = \{.5, .7, .2, .12, .42, .85, .22, .01, 0, 1\}$
- $U = \{1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, 233, 377\}$
- Suppose you waited at the entrance of the library for several hours and asked everyone what their major was when they entered. The data is summarized in the table below.

Major	Frequency
Education	43
English	34
History	25
Math	14
Psychology	39
Other	21

Create a pie chart and a bar graph for this data set. Which one gives you a better sense of the majors on campus?

- The following are the numbers of home runs per season hit by Manny Ramirez and by Alex Rodriguez between 1995 and 2006¹. Produce a box plot for each and use them to compare the two sluggers. Who would you rather have on your team as a power hitter? (Do not include personality in this decision because that makes the decision too easy).

<i>Manny</i>	31	33	26	45	44	38	41	33	37	43	45	35
<i>Arod</i>	5	36	23	42	42	41	52	57	47	36	48	35

- The following high temperatures were recorded in Salem during the month of August.

79	82	83	83	81	77	91	92	92	92
91	89	88	83	84	87	89	90	93	95
94	92	88	86	85	84	84	80	78	79
83									

Construct a dot plot for this data and analyze the distribution in terms of center, shape and spread.

7. Carl Yastrzemski played for the Boston Red Sox from 1961-1983 and is a Hall of Famer. His home run totals from his 23 year career are as follows³:

{11, 19, 14, 15, 20, 16, 44, 23, 40, 40, 15, 12, 19, 15, 14, 21, 28, 17, 21, 15, 7, 16, 10}

Construct a stem-and-leaf plot for Yaz's home run totals and analyze the distribution in terms of center, shape and spread.

8. Using the data from the Ramirez-Rodriguez example, create a back-to-back stem-and-leaf plot. Discuss each of the distributions in terms of center, shape and spread. Then indicate whether or not this representation of the data provides the same conclusion or a different one than using box plots.
9. The average price of gas (in dollars per gallon) for the United States as of September 2016 are as follows:²

1.965	1.973	1.976	1.987	2.000	2.003	2.009	2.050	2.076	2.084
2.085	2.096	2.106	2.127	2.137	2.137	2.140	2.143	2.152	2.155
2.171	2.182	2.185	2.192	2.194	2.197	2.208	2.213	2.222	2.222
2.225	2.228	2.246	2.248	2.254	2.258	2.264	2.279	2.311	2.323
2.335	2.364	2.375	2.421	2.491	2.525	2.595	2.711	2.766	2.824

Construct a histogram for this data set and discuss the center, shape and spread of the distribution.

¹Data obtained from www.baseballreference.com

²Data obtained from <https://www.gasbuddy.com/USA>

³Data obtained from <http://www.baseball-reference.com>

7.7.4 Solutions

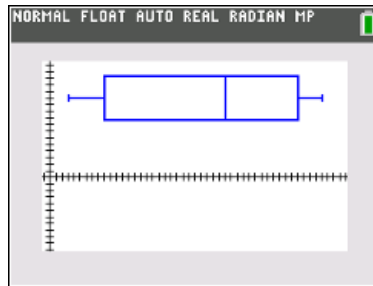
1. a. $\bar{x} = 25.57, s = 15.49$
- b. $\min = 3, Q_1 = 9, M = 29, Q_3 = 41, \max = 45$
- c.

$$1.5\text{IQR} = 1.5(41 - 9) = 48$$

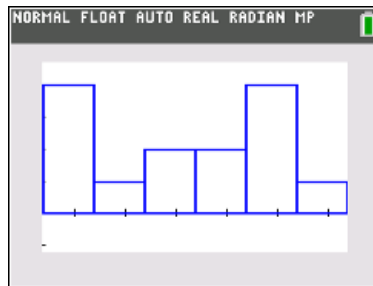
$$\text{Fences : } -37, 99$$

There are no outliers.

d.



e.



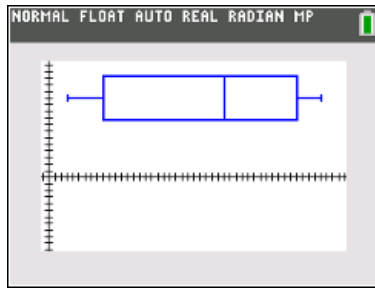
- f. The center is the boundary between the 3rd and 4th classes. The distribution is bimodal. The spread of the data is from 9 to 45, giving a range of 37.
2. a. $\bar{x} = .398, s = .357$
 - b. $\min = 0, Q_1 = .1, M = .31, Q_3 = .7, \max = 1$
 - c.

$$1.5\text{IQR} = 1.5(.7 - .1) = .9$$

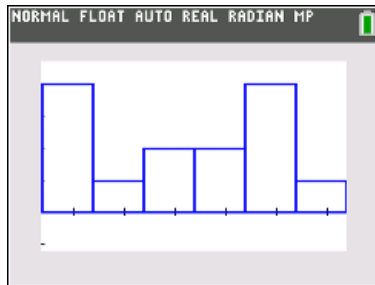
$$\text{Fences : } -.8, 1.6$$

There are no outliers.

d.



e.



f. The center is the boundary between the first two classes. The spread is from 0 to 1, giving a range of 1.01. The shape of the distribution is unimodal and it is skewed right.

3. a. $\bar{x} = 72.93, s = 115.37$

b. $\min = 1, Q_1 = 4, M = 13, Q_3 = 116.5, \max = 337$

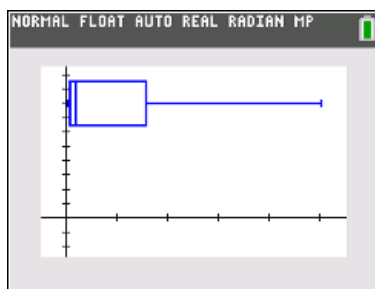
c.

$$1.5IQR = 1.5(116.5 - 4) = 168.75$$

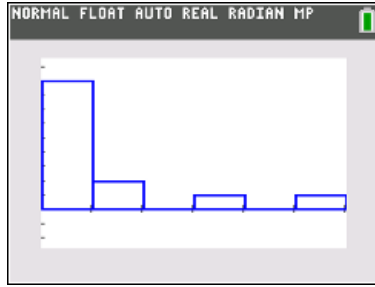
$$\text{Fences : } -164.75, 285.25$$

377 is an outlier.

d.

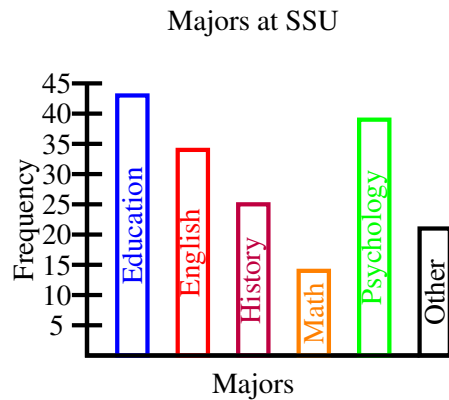
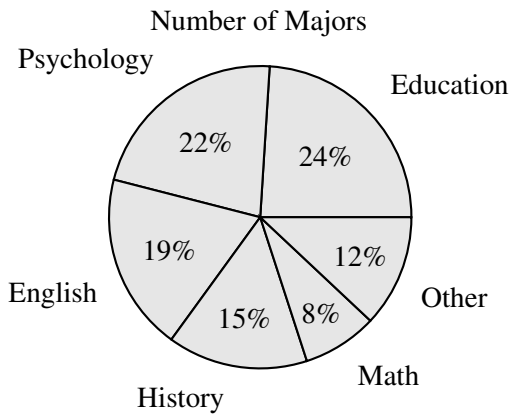


e.



f. The center of this distribution is in the first class, which is from 1 to 76.2. The distribution is unimodal and skewed right. The spread of the data is from 1 to 377, so the range is 377.

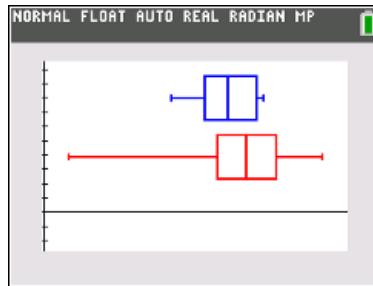
4.



It

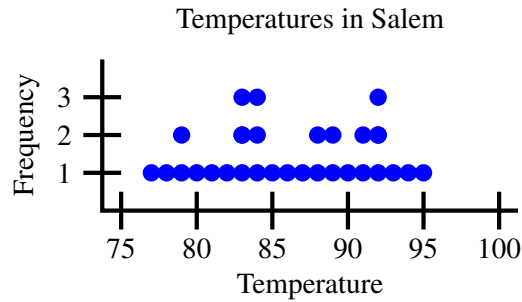
is subjective as to which a person feels they get a better sense of the situation from.

5.



Manny Ramirez was much more consistent during that time period, but Alex Rodriguez had a higher middle 50% of his data as well as a higher maximum. If the numbers were presented without knowing who the players were, one would probably choose the second set. But seeing who the players are makes the decision to choose Manny one based solely on who they are.

6.



The center of the distribution is 87. The distribution is bimodal. The spread of the data is from 77 to 95, giving a range of 19.

7.

0	7
1	0 1 2 4 4 5 5 5 5 6 6 7 9 9
2	0 1 1 3 8
3	
4	0 0 4

The distribution is unimodal and skewed right. The center is in the teens. The spread is from 7 to 44, giving a range of 38.

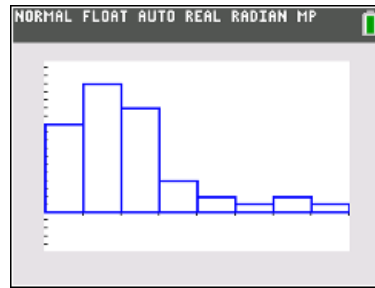
8.

					5	0
						1
					3	2 6
		6 6 5			3	1 3 3 5 7 8
8 7 2 2 1					4	1 3 4 5 5
		7 2			5	

<i>Manny</i>	31	33	26	45	44	38	41	33	37	43	45	35
<i>Arod</i>	5	36	23	42	42	41	52	57	47	36	48	35

for Manny, we have a unimodal distribution that is skewed right. The center is in the 30's and the data ranges from 26 to 45, giving a spread of 20 home runs. For Rodriguez, the data ranges from 5 to 57, giving a spread of 53 home runs. The distribution is unimodal and skewed left, with the center in the 40's. The general conclusions are the same as with the box-and-whisker plots in that Rodriguez has a higher center and multiple totals larger than the highest for Ramirez.

9.



The distribution is unimodal and skewed right. The data ranges from 1.965 to 2.824, giving a spread of .86. The center of the data is in the 2nd class, which ranges from 2.087 to 2.210.

Chapter 8

Scatter plots and Regression

8.1 Scatter plots

In this section, we will look at what happens when we have two quantitative variables that are measured on the same individuals. We express these two variables as ordered pairs and can plot them in the usual way as points on the graph. When we graph data of this type as a collection of points, we call the graph a scatter plot. Let us begin with an example where we look at the relationship by hand and then we will look at how to use technology to graph and find the relevant statistics.

Example 8.1.1 *Does the number of hours you watch TV per week impact your average grade in a class?*

Hours	12	10	5	3	15	16	8
Grade	70	85	82	88	65	75	68

To see if there is a relationship, we will create a scatter plot and analyze it.

Definition 8.1.2 A scatter plot is a geographical representation between two quantitative variables. They may be from the same individual (i.e. education v. income, height v. weight) or from paired individuals (i.e. age of partners in a relationship).

When working with scatter plots, there are two variables. They may be two different types.

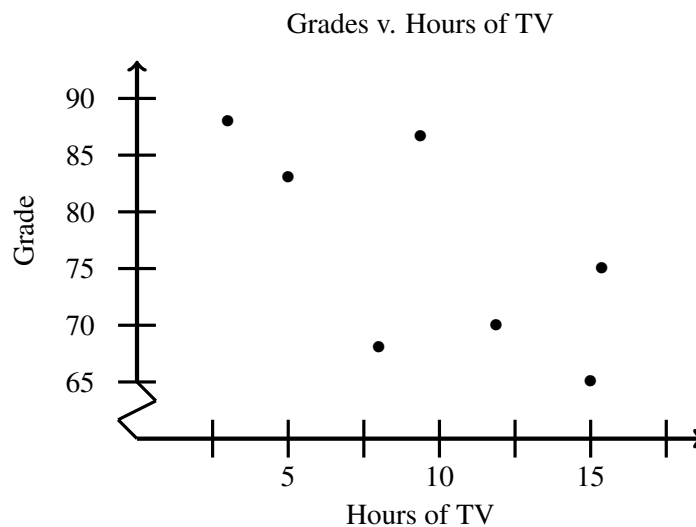
Definition 8.1.3 A response variable measures the outcome of a study.

Definition 8.1.4 An explanatory variable may explain or influence changes in a response variable.

Explanatory variables are often called independent and are on the x -axis. Response variables are often called dependent and are on the y -axis.

In our example, the explanatory variable is the number of hours of TV watched. The response variable is there for the average grade. So the question we are trying to answer is ‘Does watching TV influence the average grade in a class?’

Let’s plot the data and see what relationship appears.



It looks like the more hours of TV that are watched, the lower the average grade. But how good is the relationship? We can measure this in different ways. One is direction (+, -) and another is by ranking the strength. These are both accomplished by looking at the correlation coefficient.

Definition 8.1.5 *The correlation coefficient, r , is a number that quantifies the relationship between random variables.*

The following are properties of the correlation coefficient.

1. $-1 \leq r \leq 1$. The least correlation is 0 and the best correlation is ± 1 . Whether r is positive or negative only tells us which direction the relationship goes - whether y increases as x increases or if y decreases as x increases. Being negative is not 'bad'.
2. Correlation makes no distinction between x and y , that is, between the choice of explanatory and response variables. We need to make sure we are careful, though, as the next part (regression line) depends heavily on the correct choice.
3. Correlation measures only the linear relationship.
4. Correlation is not resistant.
5. Correlation has no units.

The formula for the correlation coefficient is not the easiest one we will see. That's why, after we look at how it works, we will rely on technology to find this value.

Formula 8.1.6 *Correlation Coefficient*

$$r = \frac{1}{n-1} \sum \left(\frac{x_i - \bar{x}}{S_x} \right) \left(\frac{y_i - \bar{y}}{S_y} \right) = \frac{1}{n-1} \sum z_x z_y$$

Let's find the correlation coefficient for our example. First, we need a few values, \bar{x} , \bar{y} , S_x , S_y .

$$\begin{array}{l|l} \bar{x} = 9.857 & \bar{y} = 76.143 \\ S_x = 4.880 & S_y = 8.971 \end{array}$$

The formula tell us that, for each pair, find the z -score for each value. Then multiply them together. After summing, divide by $n - 1$.

i	z_x	z_y	product
1	.4391	-.6848	-.3007
2	.0293	.9873	.0289
3	-.9953	.6529	-.6498
4	-1.4050	1.3217	-1.8570
5	1.0539	-1.2421	-1.3090
6	1.2588	-.1274	-.1604
7	-.3805	-.9077	.3454
			-3.9026

This gives

$$r = \frac{1}{6}(-3.9026) = -.6504$$

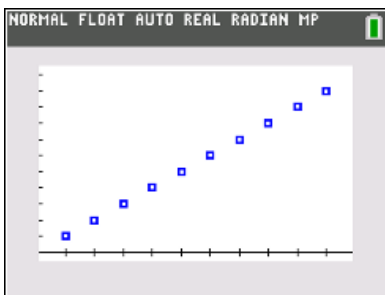
Our interpretation here is that we have a moderate, negative correlation. The negative indicates that the trend of the data suggests that as x increases, y decreases. The moderate designation is because of the magnitude of the value of r ; it is in the middle range between 0 and 1.

When describing the correlation coefficient, we want to think about how closely correlated the data appears to be. The number we find for r tells us what we need to know, but we want to have a visual in mind of what this really means.

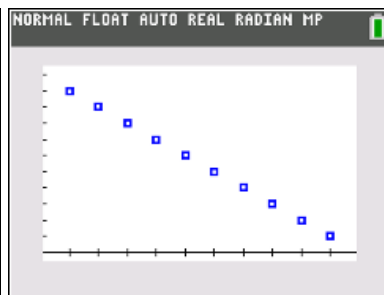
- A correlation coefficient of $r = 1$ indicates that there is a perfect positive relationship.
- A correlation coefficient of $r = .7$ indicates that there is a strong positive relationship.
- A correlation coefficient of $r = .3$ indicates that there is a weak positive relationship.
- A correlation coefficient of $r = 0$ indicates that there is no correlation.
- A correlation coefficient of $r = -.3$ indicates that there is a weak negative relationship.
- A correlation coefficient of $r = -.7$ indicates that there is a strong negative relationship.
- A correlation coefficient of $r = -1$ indicates that there is a perfect negative relationship.

Positive and negative correlation coefficients do not have anything to do with which is a ‘better’ relationship; rather, it indicates whether there is a direct or inverse relationship. If there is a direct relationship then the y -variables are increasing as the x -variables are increasing. If there is an inverse relationship then the y -variables are decreasing as the x -variables are increasing.

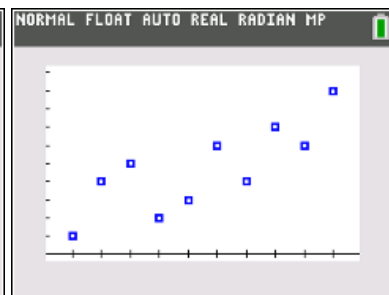
The following graphs give some possible scatter plots and the associated correlation coefficient.



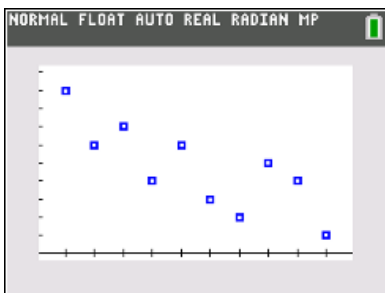
$r = 1$ perfect positive correlation



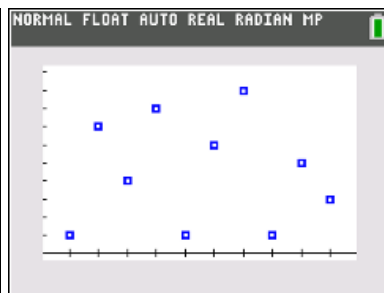
$r = -1$ perfect negative correlation



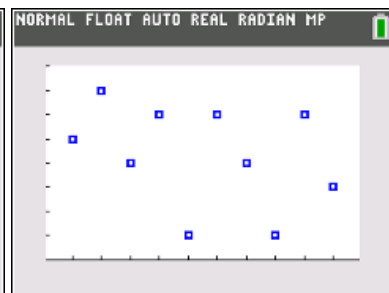
$r = .8$ strong positive correlation



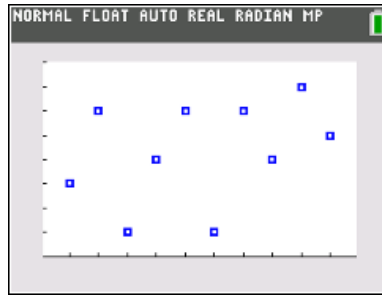
$r = -.8$ strong negative correlation



$r = -.108$ almost no relationship



$r = -.36$ weak negative correlation



$r = .36$ weak positive correlation

So, can we say that there is a direct relationship between the number of hours of TV watched and the average grade? Not so fast ... An important phrase to remember: **Correlation does not necessarily imply causation**. What this means is that just because it looks the part does not mean we have evidence that there is a relationship. We have to consider a couple of other things. One is **lurking variables**. These are variables that may be present but we are not actually considering them within the data.

We also need to test for significance to see what is going on.

- If $|r|\sqrt{n} > 3$, the correlation is significant
- Otherwise it is not significant

In our example, we see that we have less than 10 data pairs. Since $\sqrt{9} = 3$, we would need perfect correlation in order to have significance with this small a data set. But since $r = -.6504$, there is no way $|r|\sqrt{n} > 3$. So we may not have significance simply because there is not enough data. There also could just be a genuine lack of correlation causing this threshold to not be met. Remember, we shouldn't know if there is a correlation or not before we start, otherwise there is no point in doing the study in the first place. Therefore, we cannot consider watching TV to be a direct detriment on grades.

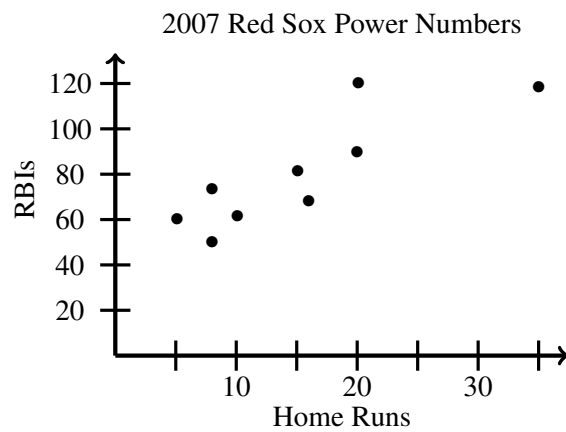
Now, a few assumptions and conditions associated with correlation coefficients.

- Make sure the variables are quantitative. We cannot use this process for categorical variables.
- We need to rely on both the statistics but also our gut reaction. We need to be able to look at the scatter plot and feel like the data is straight enough for us to proceed.
- We need to make sure there are no outliers that will dramatically affect the calculation of the regression line. Since we are basing our work on both the mean and the standard deviation, they are not resistant and outliers can have an averse effect.
- We need to make sure that the spread of the data points - the distance from where we feel the regression line will be - is approximately the same for all x values.

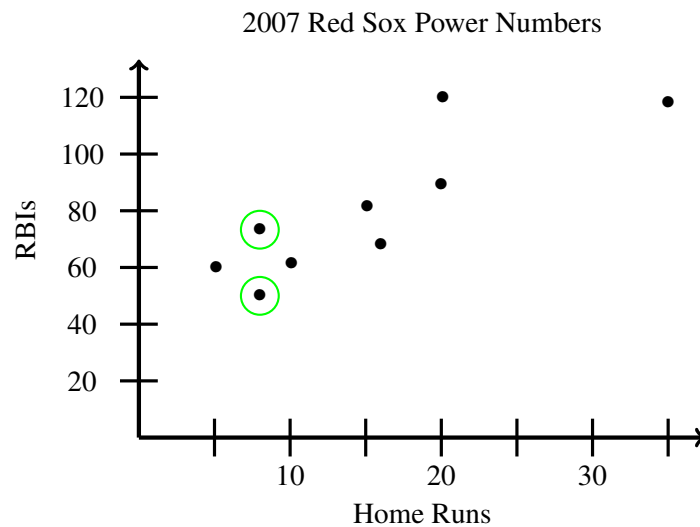
Example 8.1.7 *The following gives the power numbers for the starting 9 for the 2007 Boston Red Sox. Is there relationship between the number of home runs and the number of RBIs? Does the number of home runs affect the number of RBIs? Produce a scatter plot and discuss the correlation.*¹

<i>Player</i>	<i>Home Runs</i>	<i>RBI</i> s
<i>Varitek</i>	17	68
<i>Youkilis</i>	16	83
<i>Pedroia</i>	8	50
<i>Lowell</i>	21	120
<i>Lugo</i>	8	73
<i>Ramirez</i>	20	88
<i>Crisp</i>	6	60
<i>Drew</i>	11	64
<i>Ortiz</i>	35	117

Since we are asking if HR affects RBIs, HR would be the explanatory variable and therefore x . So RBIs is the response variable and therefore y .



Before we go on, notice that we have two values with the same x - coordinate.



¹data obtained from www.baseballreference.com

This is not a problem. Earlier in the course, we had to be working with functions, but no such condition exists here.

Now let's find the correlation coefficient using technology.

- Input data in usual way, with explanatory variable under L_1 and response variable under L_2
- Press **STAT** and scroll to **TESTS**
- Select **LinRegTTest**
- Make sure the **XList** and **YList** are the lists where the data for the explanatory and response variables are located, respectively
- Press **Calculate** and scroll to find r and r^2

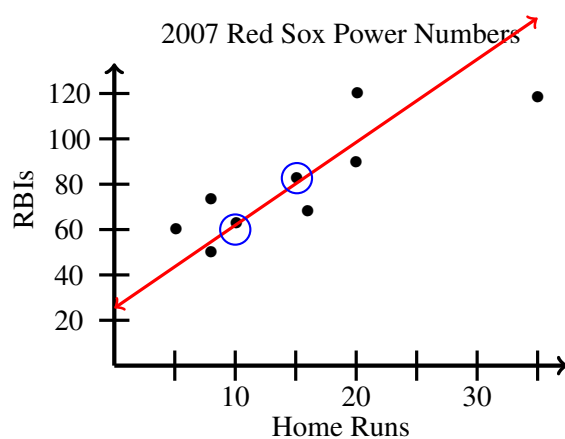
In a later example, we will look at using technology again in more detail.

For our example, we have

$$\begin{array}{c|c} r & r^2 \\ \hline .8463 & .7162 \end{array}$$

So the correlation coefficient indicates that there appears to be a strong positive correlation.

r^2 tells us how much better our predictions will be if we go through the trouble to find the regression line rather than just make our predictions with the means. Ours is pretty good here, indicating that 71.62% of the variability in the number of RBIs can be explained by the variability of the number of home runs. It therefore would make sense to find the regression line.



We can also create a scatter plot on the calculator.

- Make sure there are no functions in the grapher (press **Y=** to check)
- Input the data in the usual way (we already have it there for this example)

- Press 2^{nd} and $Y=$ to get into the **STAT PLOT** menu
- Make sure only the plot we want is turned on
- Select the first graph in the first row and then make sure the **XList** and **YList** are correct
- Press $ZOOM$ 9

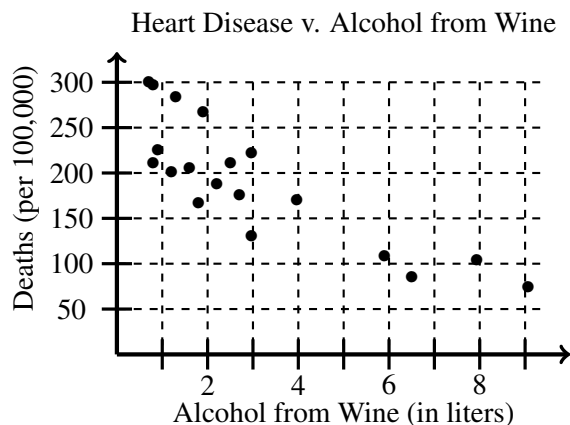
Much easier than plotting points ... we will soon discuss how to find the regression line, but first one more example.

Example 8.1.8 *There is some evidence that drinking moderate amounts of wine helps prevent heart attacks. The accompanying table gives data on yearly wine consumption (in liters of alcohol from drinking wine per person) and yearly deaths from heart disease (per 100,000 people) in 19 developing nations. Construct a scatter plot and describe what you see.*¹

Country	Alcohol	Deaths	County	Alcohol	Deaths
Australia	2.5	211	Austria	3.9	167
Belgium	2.9	131	Canada	2.4	191
Denmark	2.9	220	Finland	0.8	297
France	9.1	71	Iceland	0.8	211
Ireland	0.7	300	Italy	7.9	107
Netherlands	1.8	167	New Zealand	1.9	266
Norway	0.8	227	Spain	6.5	86
Sweden	1.6	207	Switzerland	5.8	115
United Kingdom	1.3	285	United States	1.2	199
West Germany	2.7	172			

The question we would ask here is whether or not wine consumption affects the number of yearly deaths from heart disease. So, we would have that the liters of wine would be our explanatory variable and the number of deaths from heart disease (per 100,000 people) would be the response variable.

Let's look at the scatter plot and see if it passes the eye test.



It looks pretty good, seemingly a strong negative correlation. Using technology, we see that $r = -.8428$. Further, we see that $r^2 = .7103$, telling us that 71.03% of the variability in deaths by heart disease can be explained by wine consumption. It would be worthwhile to find linear regression line to make predictions.

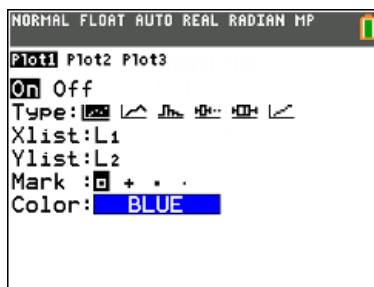
We do need to be careful here, however. This study only looks at wine consumption as a possible cause for lowering the risk of heart disease. But there are plenty of lurking variables present. First, only one type of alcohol consumption is considered; we are not looking at beer or hard alcohol consumption. There is also no mention of other health factors, such as smoking, exercise or diet. Whereas it looks like there is a strong relationship here, there are other factors to consider and therefore, we would at best be able to say that ‘there appears to be a strong negative correlation’.

Example 8.1.9 *The following are the high school and college grade point averages (on the 4.0 scale) for 15 random students. Make a scatter plot for the relationship between the high school GPAs and the college GPAs.*

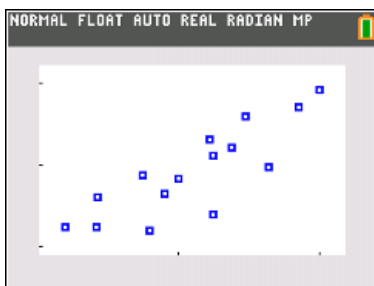
<i>Student</i>	<i>HS GPA</i>	<i>College GPA</i>
1	2.20	2.25
2	3.85	3.78
3	4.00	3.92
4	2.43	2.61
5	2.42	2.24
6	3.64	2.98
7	3.22	3.31
8	2.90	2.65
9	3.25	3.12
10	3.00	2.83
11	3.48	3.60
12	2.75	2.88
13	3.25	2.40
14	2.80	2.20
15	3.38	3.21

Solution First, we have to decide which data is going to be for the x -axis and which for the y -axis. In examples such as this, where both are response variables, it does not matter which we use for x and which we use for y . So, without loss of generality, we will use the high school GPA as the x and the college GPA as the y in each of the pairs. We know how to graph the points by hand, so we will plot them here using the calculator. To do so enter the high school GPAs under L_1 and the college GPAs under L_2 in the **STAT** menu under *EDIT*. Be sure that the ordered pairs from the data set are on the same horizontal line in the lists on the calculator. Then press **2nd** and **Y=** to get into the *STATPLOT* menu.

The scatter plot is the first of the six plots that we have to choose from. We then have to make sure that the $XList$ and $YList$ are L_1 and L_2 , respectively. At this point the screen should look like

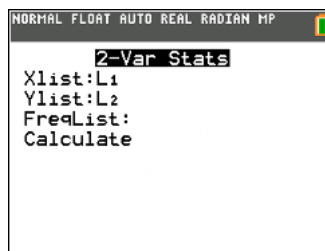
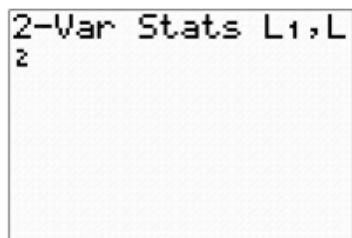


Notice that we also have a choice for what marks to use for each of the data elements. Remember that each of these marks represents an ordered pair from the list of data. We then press **ZOOM** and **9** as before to get the plot. What we should now have is

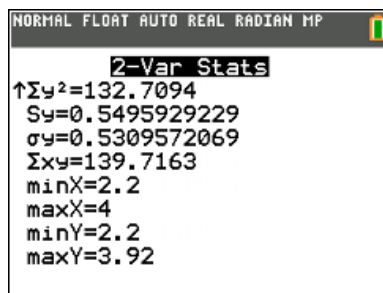
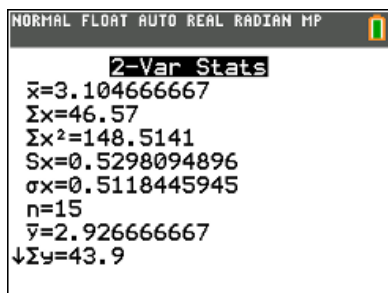


This is the scatter plot. If we want to look at individual values, we can use the **TRACE** command.

This probably seems a lot more complicated than it really is. We can get some of the information we need by using *2-Var Stats*. We will use the first example we were working with (the GPA example) to see how to calculate these statistics. To get the statistics we need, press **STAT** and then use the arrows to scroll to **CALC**. Press **2** and the screen will have nothing on it except *2-Var Stats*. We now need to tell the calculator where the two sets of data are that we want to use. The view on the right is what you will see on a newer calculator model, like the TI-84 Plus Color.



When we press **ENTER**, we get the mean and standard deviation for both sets of data, as well as other statistics.



Now we can calculate the correlation coefficient. Begin by pressing the **STAT** key and the use the arrows to scroll to **TESTS**. Here, we want the option named **LinRegTTest**. Press **ENTER**. The screen will now look like

```

NORMAL FLOAT AUTO REAL RADIAN MP
LinRegTTest
Xlist: L1
Ylist: L2
Freq: 1
β & ρ:  $\neq 0$  < 0 > 0
RegEQ:
Calculate

```

Since we are not using this for anything else, the only parameters we need to be concerned with is the Xlist and Ylist. Be sure that these indicate the correct lists for the data we are working with. Then use the arrows to scroll to *Calculate*. When we do so, we get a lot of information that we will not use, but as we scroll down the page we get to ‘ $r =$ ’ at the bottom. This is the correlation coefficient for the two variables.

```

NORMAL FLOAT AUTO REAL RADIAN MP
LinRegTTest
y=a+bx
β≠0 and ρ≠0
↑df=13
a=0.2236074515
b=0.8706439388
s=0.3100723414
r2=0.7044308446
r=0.8393037856

```

So, as we can see, the correlation coefficient for the GPA data is approximately $r = .84$, which indicates that there is a strong positive relationship between the high school GPA and college GPA.

Here, we see $r^2 \approx .70$, so any predictions made using the regression line associated with this data will be about 70% better than just using \bar{y} to guess values. And remember, this is a very good value for us to consider finding the regression line. If r^2 is close to 1, then the differences in the y values can almost completely be explained by the differences in the x values. If r^2 is close to 0, then almost none of the differences in the y values can be explained by the x values. Since we got $r^2 = .70$, almost two-thirds of the variability in college GPA is explained by the variability in high school GPA.

All this talk of regression lines and predictions and we aren’t even there yet. We will be shortly. But first ...

8.1.1 Points of Clarification for Correlation Coefficients

There are a few things that we need to be clear about with correlation coefficients. First, the correlation may be coincidental. A strong relationship indicated by the data does not mean there is a definite relationship between the two sets of data. Some possibilities are:

1. There is a direct cause-and-effect relationship (i.e. x causes y).
2. There is a reverse cause-and-effect relationship (i.e. y causes x).
3. The relationship may be caused by a third variable that was not part of the scatter plot. A variable of this type is called a **lurking variable**.

4. There may be a complicated interrelationship between more than just two variables. Depending on the situation, there may be many variables that impact the relationship.

The bottom line is that we have to be careful when interpreting the data. In the last example, the correlation coefficient would seem to indicate that there is a relatively strong relationship between the grades one gets in high school and the grades the same individual gets in college. We would think that someone who learns a lot in high school and possibly how to study will have that knowledge and skill carry over to college, so this would seem to be a direct cause-and-effect relationship. But scatter plots and correlation alone are not enough for us to definitely say that correlation implies causation.

8.2 Linear Regression

When there is a strong relationship between the two variables, we want to find some way to make predictions based on this relationship. We do this by using a **linear regression line**. That is, when there is a strong linear relationship between the two variables, we want to find the linear function that best fits the data (i.e. the linear function that fits the data with the least error). This is found by using the **method of least squares**. We need to identify which is the explanatory variable and which is the response variable and then we can find the linear equation using basic statistics from the data sets.

The equation of the regression line is of the form $\hat{y} = mx + b$ where

$$m = r \frac{s_y}{s_x}, \quad b = \bar{y} - m\bar{x}$$

So, we can find the linear regression line simply by using the mean and standard deviation for the individual sets of data and by knowing the correlation coefficient.

Example 8.2.1 Find the regression line for the relationship between the GPAs in the first example in this section.

Solution From our earlier work, we know

$$\begin{aligned}\bar{x} &= 3.105 & s_x &= .530 \\ \bar{y} &= 2.922 & s_y &= .558 \\ r &= .8411\end{aligned}$$

We calculate the slope to be

$$m = .8411 \frac{.558}{.53} \approx .8855$$

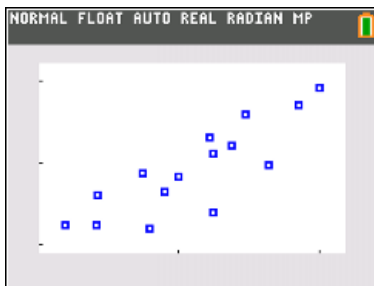
and the y-intercept to be

$$b = 2.932 - .8855(3.105) \approx .1825$$

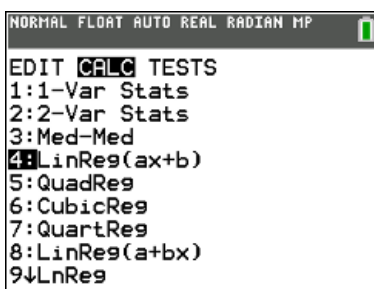
This means that the linear function that best fits the data from the first example is

$$\hat{y} = .8855x + .18255$$

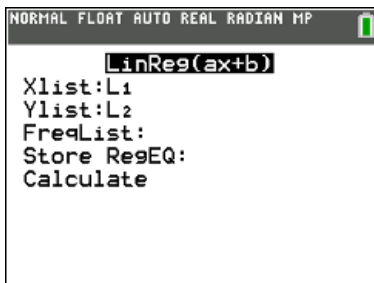
As you can probably guess, we can use the calculator to find this regression line directly. First, produce the scatter plot exactly as we did before. This way, once we have the regression line, we can plot them together and visually see the relationship between the data and the line. As a reminder, here is the scatter plot we found earlier for the GPA example.



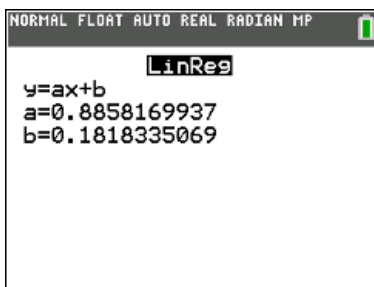
Now press **STAT** and scroll over to the *CALC* menu. The command we want is the 4th option. Notice that there are other regression options we could also find higher order regression lines. These would be parabolas that best fit the data, cubics that best fit the data, etc, but we will only concern ourselves with the linear function that best fits the data.



When the screen comes up, just make sure your lists are correct and do not worry about the other parameters. They do not do anything for us here.



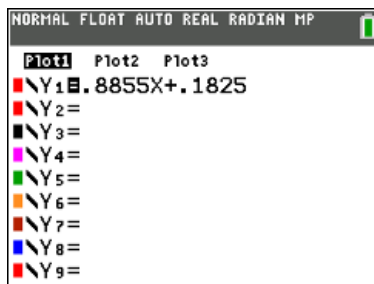
Pressing **ENTER** gives



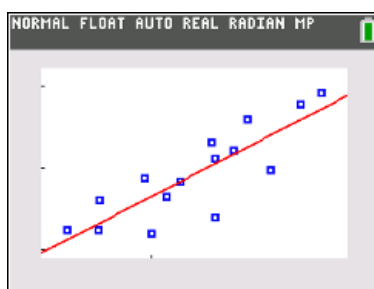
These are basically the same values we found when we calculated the slope and y-intercept by hand. The difference is because we rounded our values and the calculator did not. The implication here is that the

calculator is slightly more accurate.

Now, press **Y=** and enter the linear regression line on Y_1 , but be sure to leave on Plot1 (We know Plot1 is on because it is highlighted on the screen below.)



When we press **GRAPH**, we see the regression line plotted through the data.



8.2.1 Important things to note about regression lines

- The **response variable** is what we are studying.
- The **explanatory variable** may or may not affect the response variable - this is why we are looking to see if there is a relationship. If we already knew if there was a relationship, there would be no point in studying the two variables.
- The response variable (y) is dependent on the explanatory variable (x).
- We need to be careful in determining which is the explanatory variable and which is the response variable (if there is one of each) because the linear regression line found using the least squares method only considers distances in the y -direction. If the variables are reversed then the resulting regression line will be different and the predictions we make will be greatly affected.
- The regression line will always pass through the point (\bar{x}, \bar{y}) but may not pass through any of the other points.
- The closer to ± 1 the correlation coefficient is, the better the regression line will fit the data. But, all of the data points will be on the regression line only when there is a perfect relationship.

8.2.2 What good is having this regression line?

We said in the earlier section that we could use the linear regression lines to make predictions about the situation. For example, what GPA could we expect in college from a student who graduated high school with a 3.5? The linear regression line gives us the means to make a prediction about this value based on the research that has already been done. We can simply substitute $x = 3.5$ into our regression equation and the corresponding y -value will be the predicted college GPA.

$$\hat{y} = .5876(3.5) + 1.0675 \approx 3.12$$

So, we can expect a student with a 3.5 GPA in high school to have approximately a 3.12 GPA in college.

This is the point of a regression line. We can use it to predict the future or to find an the value of an explanatory variable for a value of a response variable that is not explicitly part of the data. The important thing to note is that the regression line is the linear function that fits the data *with the least error*. It is not going to be exact for most of the data. If we substitute an actual data point into the regression line, the approximating value will more often than not be different then the actual value. For example, using our regression line in the same example, if we substitute the value of $x = 4.0$, we get $\hat{y} = .5876(4.0) + 1.0675 \approx 3.42$ but the point in the data was $(4.0, 3.92)$. We have to remember that the regression line gives a prediction, but empirical data is more accurate.

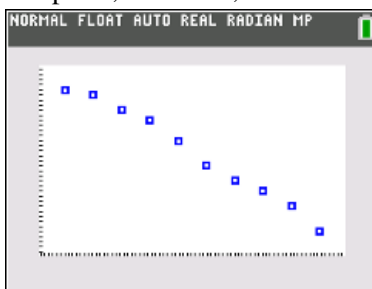
Example 8.2.2 *The percentage of smokers in a given city has declined over time. The following table gives the percentage of smokers in the city in 10 year increments.*

Year	1900	1910	1920	1930	1940	1950	1960	1970	1980	1990
%	48	47	44	42	38	33	30	28	25	20

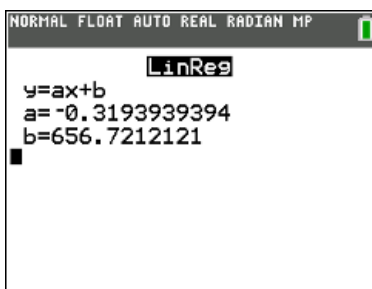
- Create a scatter plot for the given data.
- Find the correlation coefficient for the two variables.
- Find and interpret r^2
- If there is a strong relationship, find the regression line.
- Based on this data, what percentage of smokers can we predict in this city in the year 2007?

Solution

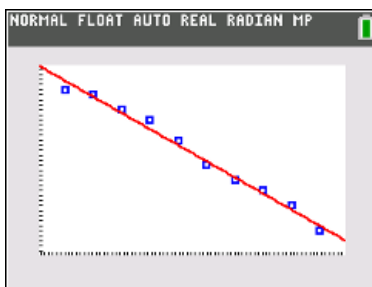
- To create the scatter plot, we enter the year under L_1 and the percentage under L_2 as we did before. Just a reminder, we get to the list by using the **Edit** option in the *EDIT* menu after we press **STAT**. Then we enter the *STATPLOT* menu by pressing **2nd** and **Y=** as we did before. In whichever plot we choose, we turn the plot on, choose the 1st option for scatter plot and make sure that the Xlist and Ylist are the lists on which we entered the data. Then we press, as before, **ZOOM** and **9**. The scatter plot for this data is



- b. Without even finding the correlation coefficient, we can see that there is a strong negative relationship. When we calculate r , we get $r = -.9935$.
- c. $r^2 = .9870$, which tells us that 98.7% of the variability in the number of smokers can be explained by which year we are considering.
- d. Since there is a strong relationship, we will find the regression line. Using the **CALC** menu under **STAT** we will select the 4th option, **LinReg** ($ax + b$). (Remember if the data is not under L_1 for the explanatory variable and L_2 for the response variable then we need to specify the lists so that the calculator knows where to get the information.) When we press **ENTER**, we get the values for the y -intercept and slope of the line.



Entering these under Y_1 allow us to see the line with the points as before.



- e. By using our regression line $\hat{y} = -.3194x + 656.72121$ we can predict that the percentage of smokers in this city in the year 2007 should be

$$\hat{y} = -.3194(2007) + 656.72121 \approx 15.87\%$$

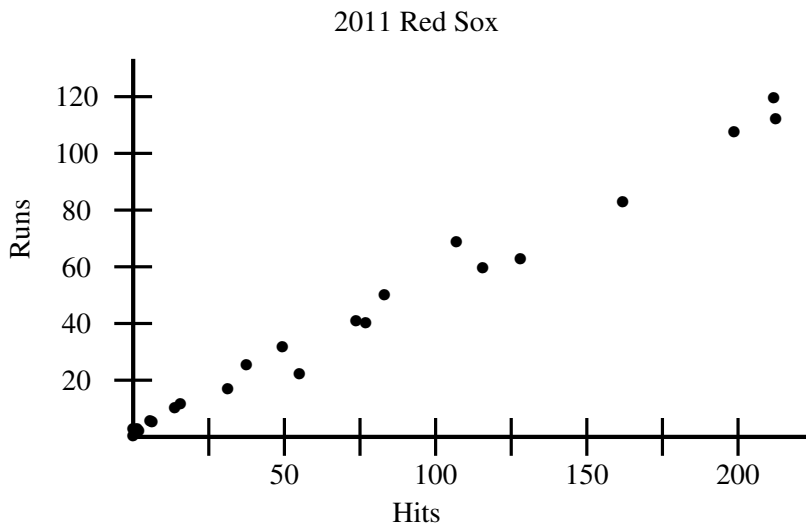
8.2.3 Residuals

In this next example, we will determine if a linear relationship is really appropriate through the use of residuals.

Example 8.2.3 Suppose we wanted to know if a player was expected to score more runs if he got more hits. To answer this question, we will use the roster of the 2011 Boston Red Sox.¹

<i>Name</i>	<i>Runs</i>	<i>Hits</i>
<i>Jarred Saltalamacchia</i>	52	84
<i>Adrian Gonzalez</i>	108	213
<i>Dustin Pedroia</i>	102	195
<i>Marco Scutaro</i>	59	118
<i>Kevin Youkilis</i>	68	111
<i>Carl Crawford</i>	65	129
<i>Jacoby Ellsbury</i>	119	212
<i>J.D. Drew</i>	23	55
<i>David Ortiz</i>	84	162
<i>Jed Lowrie</i>	40	78
<i>Josh Reddick</i>	41	71
<i>Jason Varitek</i>	32	49
<i>Darnell McDonald</i>	26	37
<i>Mike Aviles</i>	17	32
<i>Mike Cameron</i>	9	14
<i>Drew Sutton</i>	11	17
<i>Ryan Lavarney</i>	5	9
<i>Yamaico Navarro</i>	6	8
<i>Conor Jackson</i>	2	3
<i>Jose Iglesias</i>	3	2
<i>Lars Anderson</i>	2	0
<i>Joey Gathright</i>	1	0

Now, let's look at the scatter plot for the data.



Since the plot looks like there is a strong correlation, we would want to proceed. But before we do, we want to make sure all of the assumptions about regression lines are met.

- Both of the variables are quantitative.
- The data from the plot looks like the line we would produce would be straight enough.
- There does not seem to be any outliers that would dramatically affect the fit of our line.
- The spread of the data seems to be consistent for all x .

¹data obtained from www.baseballreference.com

Next, we find the correlation coefficient. When we plug all of the data into our technology, we get $r = .9942$. There is a very strong, positive correlation between hits and runs scored. Also, $r^2 = .9884$, which tells us that we can explain 98.84% of the variability in the runs total by the variability in the number of hits.

Another way to interpret this is in terms of standard deviations. For each standard deviation above the mean for the explanatory variable x , y will be r standard deviations above the mean of the response variable. Here, we would say that for every standard deviation we are above the mean number of hits, we will be .9942 standard deviations above the mean number of runs.

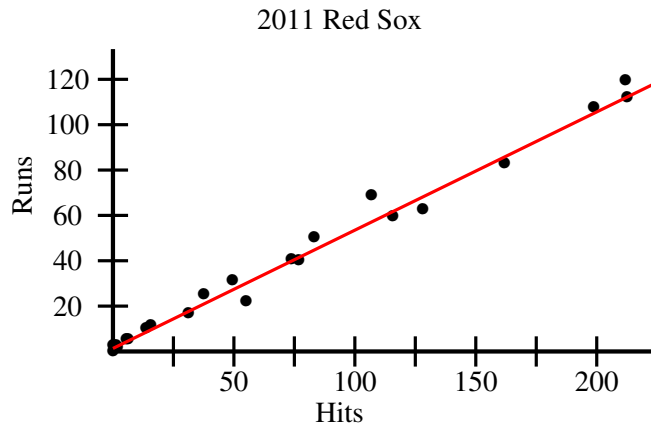
Another point we can make about the data based on r^2 is that $1 - r^2$ is the fraction of the variation in the original data left in the residuals - we will get to these shortly.

Since all of these are satisfied, we will continue on to find the formula of the regression line. Using our technology, we have

$$\hat{y} = b_0 + b_1x = 1.92 + .52x$$

The practical interpretation of the slope is that for each hit, we expect a player to score .52 additional runs. And, the practical interpretation of the y -intercept is that if a player has no hits, we expect 1.92 runs to be scored.

Now that we have the equation of the regression line, we will plot it with the scatter plot.



So, when we plot the regression line over the scatter plot, we see that the line is a good fit.

We want to make predictions with this line.

- If a player got 200 hits, how many runs would we expect them to have?
Here, we are given the x value and using our regression line, we find the predicted value.

$$\hat{y} = 1.92 + .52(200) \approx 105.92$$

So, we'd expect about 106 runs for a player with 200 hits.

- What if we wanted to know how many hits a player had if they scored 120 runs?
We are given the value of \hat{y} and want to find the value of x . So, we use our algebra skills ...

$$\begin{aligned}\hat{y} &= 1.92 + .52x \\ 120 &= 1.92 + .52x \\ 118.08 &= .52x \\ 227.08 &= x\end{aligned}$$

We expect about 227 hits.

Everything looks good, but we cannot be sure yet. The correlation coefficient and the significance level are all indicators, but correlation does not necessarily imply causation. A way we can get a better idea about how good the predictions will be is through the use of residuals.

Definition 8.2.4 The *residual* measures the difference between the observed and predicted values. The residual, e is found using

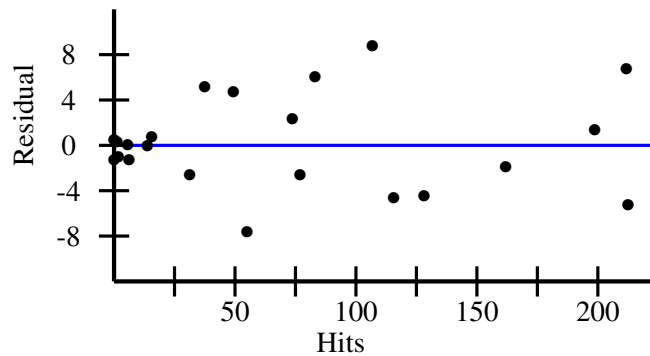
$$e = y - \hat{y}$$

Residuals give us the error in our predictions based on the observed values. For our example, we have

Name	Runs	Hits	\hat{y}	e
Jarred Saltalamacchia	52	84	45.6	6.4
Adrian Gonzalez	108	213	112.68	-4.68
Dustin Pedroia	102	195	103.32	-1.32
Marco Scutaro	59	118	63.28	-4.28
Kevin Youkilis	68	111	59.64	8.36
Carl Crawford	65	129	69	-4
Jacoby Ellsbury	119	212	112.16	6.84
J.D. Drew	23	55	30.52	-7.52
David Ortiz	84	162	86.16	-2.16
Jed Lowrie	40	78	42.48	-2.48
Josh Reddick	41	71	38.84	2.16
Jason Varitek	32	49	27.4	4.6
Darnell McDonald	26	37	21.16	4.84
Mike Aviles	17	32	18.56	-1.56
Mike Cameron	9	14	9.2	-.2
Drew Sutton	11	17	10.76	.24
Ryan Lavarney	5	9	6.6	-1.6
Yamaico Navarro	6	8	6.08	-.08
Conor Jackson	2	3	3.48	-1.48
Jose Iglesias	3	2	2.96	.04
Lars Anderson	2	0	1.92	.08
Joey Gathright	1	0	1.92	-.92

This doesn't seem to tell us much, but we will have something to analyze when we produce a residual scatter plot. This is a scatter plot of the residuals versus the x values. So, we are plotting the points (x_i, e_i) .

2011 Red Sox



What we want is the most boring scatter plot we could possibly have. We want to see

- Roughly horizontal
- Same scatter throughout
- No interesting slope or direction
- No outliers

This plot is far from a horizontal line, but the spread looks about the same throughout, so we feel like this is a good indication that a linear model is appropriate here.

A note on residuals: the mean of the least squares residuals always equals 0. \hat{y} and y differ by a percentage. It is $x\%$ different than what we would have predicted by \hat{y} . Because residuals tell us how far from the regression line our data falls, it can tell us how good the regression line actually is. This is because **the residual is measuring the error**.

We can also look at the residual standard deviation to see how consistent this data is. Our formula is

$$s_e = \sqrt{\frac{\sum e^2}{n-2}}$$

Here, we get $s_e \approx 4.125$.

When we look at our residuals, we see that most of the data points (14 of 22) lie within one standard deviation of 0.

$$\frac{14}{22} \approx 63.63\%$$

Of the remaining 7 points, all but one lie within 2 standard deviations of 0.

$$\frac{21}{22} \approx 95.45\%$$

So, this roughly satisfies the 68-95-99.7% rule. If we so chose, we could produce a histogram for the residuals, and we would be looking for a roughly symmetric, unimodal graph.

A few important points to keep in mind

1. An observation is influential for a statistical calculation if removing it would markedly change the results of the calculation. Point that are outliers in either the x or y direction are often influential points.
2. Correlation and least squares regression lines are not resistant
3. They only describe linear relationships.
4. There could be lurking variables. Those are ones that are not among the explanatory or response variables but may influence the interpretation of the relationship.
5. An association between an explanatory variable x and a response variable y , even if r is very strong, is not itself good evidence that changes in x actually cause changes in y . The phrase to remember is that **correlation does not necessarily imply causation**.

There are two different ways we can find the residuals and produce a residual scatter plot, and which we use is dependent on which values we need in our work. Either way, we first press **STAT** and then press **ENTER** on **EDIT** and then put the explanatory variable under L_1 and the response variable under L_2 .

- While in the spreadsheet, use the arrows to highlight L_3
- Type the regression equation $b_0 + b_1x$ in, where the variable is L_1 , which is found by pressing **2nd** and then **1**.
- Press **ENTER** These will be the predicted values for each of the x values.
- Scroll so that the cursor is highlighting L_4
- Type $L_2 - L_3$ and press **ENTER**. These will be the residuals.

To produce the residual scatter plot, go to the **STATPLOT** menu as usual and select the first plot. In theory, by the time you are looking for the residual scatter plot, you have already produced the regular scatter plot. The difference here is that for Y -List, select L_4 instead of L_2 . Then press **ZOOM** and **9** as usual. If all we are concerned about is the residuals and we don't need the predicted values, we would do the following set of steps. Assuming the data is already set in L_1 and L_2 as the explanatory and response variables for the situation, we next ...

- Move the arrows to highlight L_3
- Press **2nd** and **STAT** to get into the **LIST** menu
- Scroll down to **RESID** and press **ENTER**
- Now back in the spreadsheet, press **ENTER**. These will be the residuals.

To produce the residual scatter plot, go into the **STATPLOT** menu as usual and select the same plot you used before for the regular scatter plot. To make the residual scatter plot, change the Y List to L_3 by pressing **2nd** and **3** and then press **ZOOM** and **9**.

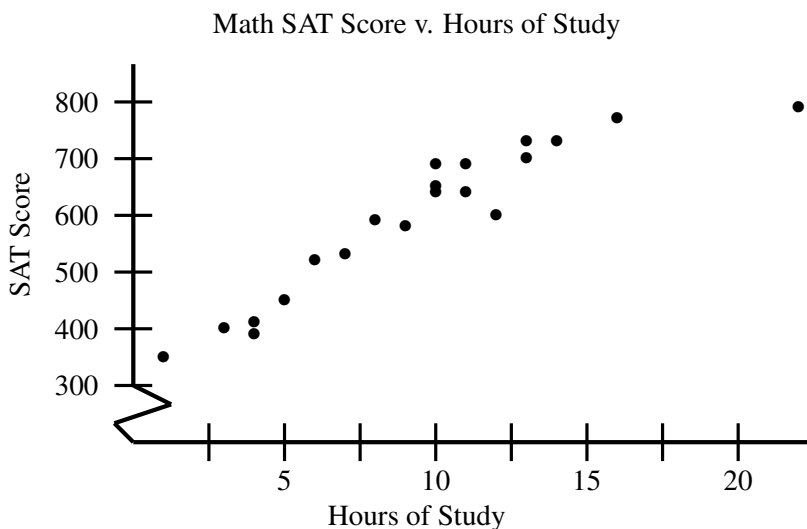
We can find the value of s_e quickly using the calculator as well. In fact, we already found it but we didn't need it at the time. When you use the **LinRegTTest** to find the coefficients for the regression line and to find r and r^2 , the value in the output where we find these labeled as 's' is the residual standard deviation. So, we actually require no additional work.

Example 8.2.5 We want to know if there is a relationship between the score on the math portion of the SAT exam and the number of hours studying for the test. The question is, "Does studying more increase the score on the exam?" The following data was taken from a study conducted of 20 students as they prepared and took the SAT exam.

Hours	4	9	10	14	4	7	12	22	1	3
Score	390	580	650	730	410	530	600	790	350	400
Hours	8	11	5	6	10	11	16	13	13	10
Score	590	640	450	520	690	690	770	700	730	640

Based on the scenario, the response variable is the Math SAT score and the explanatory variable is the hours of study.

Next, let's produce our scatter plot so we can see what we are dealing with.



We are feeling pretty good about this - it seems to have a strong, positive correlation. When we consider the conditions, are we still happy with this?

- Both variables are quantitative.
- Data looks reasonably straight.
- There do not seem to be any outliers.
- The spread seems to be consistent. Other than the one person who studied for 22 hours, the relationship seems very strong.

Now, we find the correlation coefficient to be $r = .9336$, which would lead us to believe there is a strong, positive correlation. And, $r^2 = .8716$, telling us that 87.16% of the variation in the Math SAT score can be explained by the variation in the hours of study.

Looks good so far, but is the data significant?

$$r\sqrt{n} = .9336\sqrt{20} \approx 4.17 > 3$$

So, the data is significant based on this criteria.

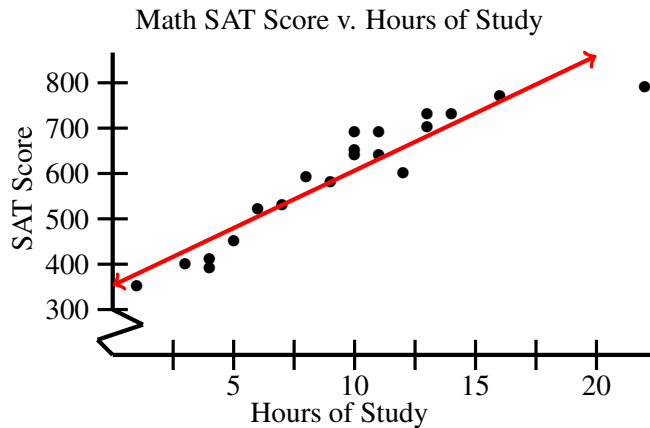
Next, using our technology, we find the equation of the regression line is

$$\hat{y} = 353.16 + 25.33x$$

A few questions we could ask:

- What is the practical interpretation of the slope $b_1 = 25.33$?
For each hour of study, we expect the person to get an additional 25.33 points on their score.
- What is the label for the slope?
Points per hour of study
- What is the practical interpretation of the y-intercept $b_0 = 353.16$?
If a person does not study, we expect their score on the Math portion of the SAT exam to be 353.16.

Now, let's look at the scatter plot with the regression line to make sure things still look good.



That data point where the person studied for 22 hours does look a little sketchy, but it does not seem so far out of whack that it seems to be an outlier.

Now, we will use this line to make a couple of predictions.

So what score would we expect for a person who studied for 10 hours?

$$\hat{y} = 353.16 + 25.33(10) = 606.46$$

So, since SAT scores are rounded to the nearest 10, we would expect about a 610.

If someone scored a 720, how many hours would we guess they studied?

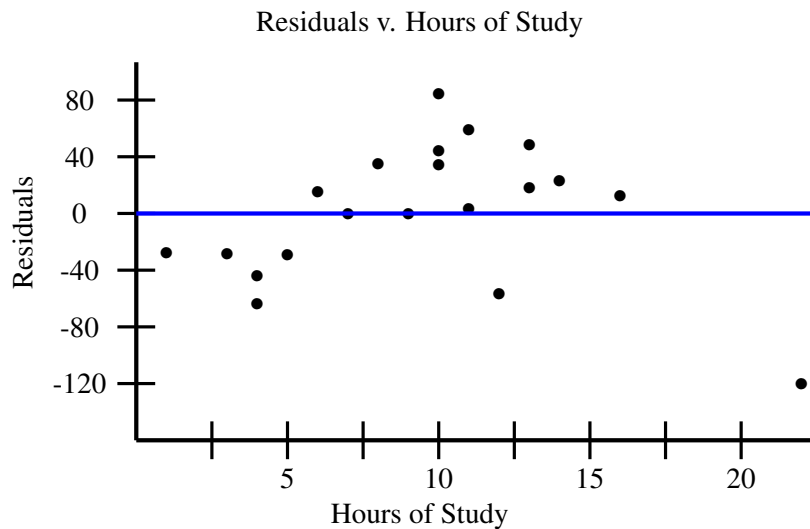
$$720 = 353.16 + 25.33x \Rightarrow x = 14.48 \text{ hours}$$

Now we turn our attention to residuals to see just how linear this relationship really is. Remember,

$$e = y - \hat{y}$$

Hours	4	9	10	14	4	7	12	22	1	3
Score	390	580	650	730	410	530	600	790	350	400
Prediction	454.48	581.13	606.46	707.78	454.48	530.47	657.12	910.42	379.49	429.15
Residual e	-64.48	-1.13	43.54	22.22	-44.48	-.47	-57.12	-120.42	-28.49	-29.15
Hours	8	11	5	6	10	11	16	13	13	10
Score	590	640	450	520	690	690	770	700	730	640
Prediction	555.8	707.78	479.81	505.14	606.46	631.79	758.44	682.45	682.45	606.46
Residual e	34.2	8.21	-29.81	14.86	83.54	58.21	11.56	17.55	47.55	33.54

These values are all over the place. Let's look at the residual scatter plot and see if there is any indication from that.



This gives us a pattern we didn't want to see, so this set of data is probably not very well suited for a linear model.

By looking at the standard deviation of the residuals, we see

$$s_e = \sqrt{\frac{\sum e^2}{n-2}} \approx 49.72$$

With a little calculation, we see that

$$\frac{15}{20} = 75\%$$

of the data within one standard deviation, above 68% we are expecting.

One final idea: suppose we have a person who studies for 7 hours and had a residual of -25. How can we find their score on the exam? What we need to solve is

$$e = y - \hat{y} \Rightarrow y = \hat{y} + e$$

So, we can find the predicted value, add the residual and this will tell us the person's score on the exam.

$$\hat{y} = b_0 + b_1x = 353.16 + 25.33(7) \approx 530.47$$

Then,

$$y = b_0 + b_1x + e = 530.47 + (-25) = 505.47$$

Since we need to round to the nearest whole number, we believe the person who studies for 7 hours got a 505 on the Math portion of the SAT exam.

8.3 Lurking Variables

Earlier in this chapter, we mentioned the idea of lurking variables and we will discuss more formally here.

Definition 8.3.1 *A lurking variable is a variable that may affect what we are studying but is not included in the response or explanatory variable.*

We need to be careful with lurking variables because they can provide us with a false sense about the relationship between the variables when we are looking at bivariate data. For example, in our first scatter plot example, we looked at the relationship between TV watching and a student's grade in a class. But we all know that there are many other factors when we consider the time we have to study. Some of us have children and they are certainly a priority and take a great deal of time and provide a great deal of distraction. Some of us actually study better with the background noise and distraction of TV and suffer when there is silence while studying. The impact of roommates on the ability to concentrate at times can affect study as well. But none of these factors are considered when looking at a potential relationship between a response variable and an explanatory variable.

Example 8.3.2 *Suppose you are studying whether a person's salary impacts how far they are willing to commute for work. The expectation is that people would not be willing to commute far for a low paying job but would for a high enough salary. What lurking variables could be present here?*

Solution There are many possible. A few could be:

- These two variables do not consider how much a person likes their job
- The two variables do not consider any benefits or perks of the job outside of salary

- The two variables do not consider how much time is spent in the office v. time spent working from home

This is one of the limitations with linear regression - we cannot include more than the two variables into consideration when looking at a situation. This is why, no matter how compelling the relationship may seem to be, the best we can say is that there appears to be a strong relationship.

8.4 Exercises

For each of the three questions, do the following:

- Create a scatter plot for the given sets of data.
- Find and interpret the correlation coefficient.
- Find and interpret r^2 .
- Find the equation of the regression line.
- Answer the question with the set of data.
- Produce a residual plot for the set of data.
- Do you feel that the regression line is a good predictor? Explain.

- The following list gives the power numbers for 10 Red Sox players for the 2006 season.²

Name	Home Runs	RBI
David Ortiz	54	137
Manny Ramirez	35	102
Jason Varitek	12	55
Trot Nixon	8	52
Kevin Youkilis	13	72
Coco Crisp	8	36
Willy Mo Pena	11	42
Mike Lowell	20	80
Mark Loretta	5	59
Alex Gonzalez	9	50

We want to know if there is a relationship between the number of home runs hit and the number of runs batted in. If a player hit 60 home runs, how many RBI could we expect the player to have had?

- Seeing as how we are all huge Red Sox fans, we are also interested in the relationship between the number of hits and the number of runs scored. Here are the associated numbers for the same 10 players from the 2006 Red Sox team.²

Name	Hits	Runs
David Ortiz	160	115
Manny Ramirez	144	79
Jason Varitek	87	46
Trot Nixon	102	59
Kevin Youkilis	159	100
Coco Crisp	109	58
Willy Mo Pena	83	36
Mike Lowell	163	79
Mark Loretta	181	75
Alex Gonzalez	99	48

If I was on the Sox in 2006 and got 200 hits, how many runs would you predict that I would have scored?

3. The following chart gives the passing yardage and touchdown passes for Tom Brady during his Patriots career so far.³

Year	Passing Yards	Touchdowns
2000	6	0
2001	2843	18
2002	3764	28
2003	3620	23
2004	3692	28
2005	4110	26
2006	3529	24
2007	4806	50
2008	76	0
2009	4398	28
2010	3900	36
2011	5235	39
2012	4827	34
2013	4343	25
2014	4109	33
2015	4770	36
2016	3554	28

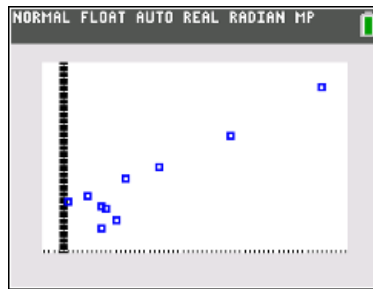
We are looking at the relationship between passing yards and touchdown passes. Based on this data, how many touchdowns can we expect Tom Brady to throw next season if he passes for 4000 yards?

²data obtained from www.baseballreference.com

³data obtained from www.espn.com

8.5 Solutions

1. a.



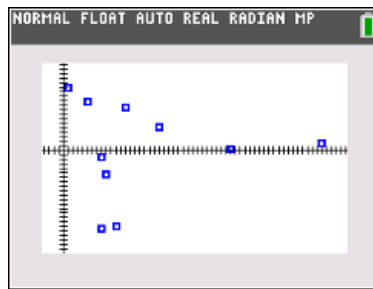
b. $r = .9368$, which represents a very strong, positive linear correlation.

c. $r^2 = .8775$, which means that 87.75% of the variation in the number of RBIs can be explained by the variation in the number of home runs.

d. $\hat{y} = 1.783x + 39.267$

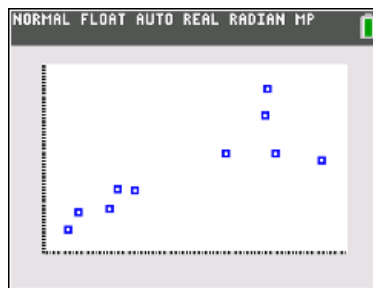
e. $\hat{y} = 1.783(60) + 39.267 = 146.247$, so we would expect about 146 RBI for a player who hit 60 home runs.

f.



g. Since the residual plot does not produce a horizontal line, we do not feel this regression line is a good indicator, even though the correlation coefficient was strong.

2. a.



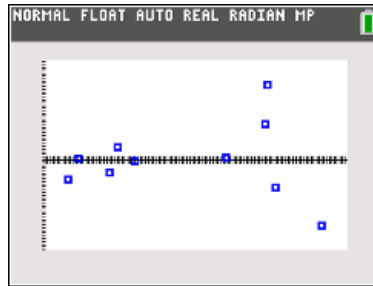
b. $r = .8308$, which indicates a strong, positive correlation.

c. $r^2 = .6903$, which indicates that 69.03% of the variation in the number of runs scored can be explained by the number of hits.

d. $\hat{y} = .5705x - 3.931$

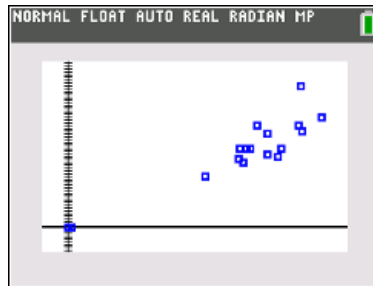
e. $\hat{y} = .5705(200) - 3.931 = 110.169$, so we would expect a player who got 200 hits to score about 110 runs.

f.



g. This regression line does not appear to be a good predictor. It looked good for the first group of points, but the last 4 deviate from the horizontal line we were hoping to see.

3. a.



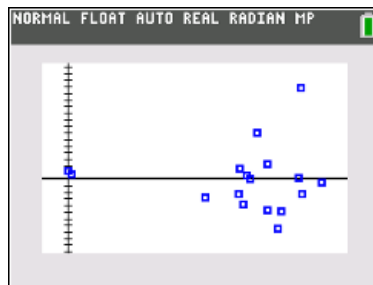
b. $r = .9190$, which represents a strong, positive linear relationship.

c. $r^2 = .8446$, which means that 84.46% of the variation in the number of touchdowns can be explained by the number of passing yards.

d. $\hat{y} = .0078x - 1.374$

e. $\hat{y} = .0078(4000) - 1.374 = 29.826$, so we would expect roughly 30 touchdowns if TB throws for 4000 yards next season.

f.



g. Based on that the residual plot is not producing data that can be represented by a horizontal line, this regression line does not appear to be a good predictor of touchdowns thrown.

Chapter 9

Probability

9.1 The Basic Laws of Probability

9.1.1 Definitions and Examples

Definition 9.1.1 *Probability is the likelihood that a random phenomenon occurs. WE often express probability as the proportion of the times the outcome would occur in a very long sequence of repetitions.*

There are two important parts in this definition that are cannot be understated. The first is the word here is **random**. For a phenomenon to be random, individual outcomes are uncertain but there are nonetheless regular distributions of relative frequencies over a large number of trials.

The second part that needs to be stressed is that probability is based on a very long sequence of outcomes. Probability is the long term expectation even though the short term situation is unpredictable - this is the Law of Large Numbers, which was introduced by Jacob Bernoulli in 1689. He is one of a family of sons and their father who did a lot of work in math and physics.

Theorem 9.1.2 *The Law of Large Numbers*

If an experiment is repeated a large number of times, the empirical (experimental) probability of a particular outcome approaches a fixed number as the number of repetitions increases.

Empirical (experimental) probability is the percent we get from doing an actual experiment, such as flipping a coin. If we have a long series of trials, provided we have a random event, we will approach the theoretic probability. So the probability you flip a coin and gets heads is not just because there are two sides to the coin but because if we flipped a coin for a very, very long time and recorded the sequence of heads and tails, we would see that as the number of flips approached infinity, the closer to 50% we would get.

Example 9.1.3 *Suppose you flip a coin 3 times and record the sequence of heads and tails. What is the probability of getting 2 tails?*

Solution To solve, we need to define a few terms along the way.

Definition 9.1.4 *An experiment is any procedure that can be infinitely repeated and has a well-defined set of possible outcomes, known as the sample space.*

Definition 9.1.5 *An outcome is a possible result of an experiment. Each possible outcome of a particular experiment is unique, and different outcomes are mutually exclusive (only one outcome will occur on each trial of the experiment).*

Definition 9.1.6 *The sample space of an experiment is the set of all possible outcomes.*

In our example, the sample space is

$$S = \{HHH, HHT, HTH, THH, THT, TTH, HTT, TTT\}$$

Definition 9.1.7 *An event is a subset of the sample space.*

Our event is getting two tails.

When calculating probabilities, we need to know the number of outcomes in the event and the number of outcomes in the sample space. Then,

$$P(A) = \frac{N(A)}{N(S)}$$

In our example, the event space is

$$S = \{THT, TTH, HTT\}$$

which gives the probability

$$P(2T) = \frac{3}{8}$$

Example 9.1.8 *Suppose we have a standard die. What is the probability of rolling a 5 or a 6?*

Solution Our sample space here is

$$S = \{1, 2, 3, 4, 5, 6\}$$

and the event is

$$E = \{5, 6\}$$

Therefore, the probability is

$$P(E) = \frac{2}{6} = \frac{1}{3}$$

Note: What we are calculating here are the theoretic probabilities since we are using logic to determine the values rather than experimentation.

Example 9.1.9 *Suppose we have a standard die. What is the probability of rolling less than 5?*

Solution If we look at the last example, we found that the probability of rolling at least 5 was $\frac{1}{3}$. What we are looking for here is every other possible roll besides what was considered in the last scenario. So,

$$P(\bar{E}) = \frac{4}{6} = \frac{2}{3}$$

Notice that these two events (at least 5, less than 5) cannot occur at the same time and are ‘opposites’ of each other.

$$P(5 \text{ or } 6) = \frac{2}{6} = \frac{1}{3} = 1 - P(1, 2, 3, 4)$$

We would call these events complements. A complement of an event is all of the possible outcomes from the sample space besides those in the event.

Example 9.1.10 *Suppose you are getting dressed and need to get a pair of socks. Because you are too lazy to go turn a light on, you are searching in the dark. There are 10 pairs of white socks, 8 pairs of blue socks and 4 pairs of black socks in the drawer. What is the probability you randomly grab a pair of socks and they are white?*

Solution Here, it is easiest to approach the problem directly because we only need to know how many socks there are in the drawer and how many of them are white.

$$P(W) = \frac{10}{22} = \frac{5}{11}$$

Example 9.1.11 *Suppose you are getting dressed and need to get a pair of socks. Because you are too lazy to go turn a light on, you are searching in the dark. There are 10 pairs of white socks, 8 pairs of blue socks and 4 pairs of black socks in the drawer. What is the probability you randomly grab a pair of socks and they are not blue?*

Solution Here, because we want *not blue*, it is easier for us to use the complement.

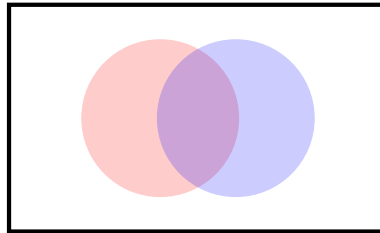
$$\begin{aligned} P(\text{not } B) &= P(W \text{ or } Bl) = P(W) + P(Bl) \\ &= \frac{10}{22} + \frac{4}{22} \\ &= \frac{14}{22} \end{aligned}$$

Notice that you could have done either of these directly or either using complements. The choice of which to use is up to you, for the most part, as long as you follow the correct rules for finding probabilities.

9.1.2 Subsets

Definition 9.1.12 *The intersection of sets A and B , denoted $A \cap B$, is the set of all elements that are simultaneously in A and B .*

Definition 9.1.13 *The union of sets A and B , denoted $A \cup B$, is the set of all elements that are in A or in B . This set includes all elements in A but not B , in B but not A and those in A and B simultaneously.*



The intersection of the two sets is the purple region where the two circles overlap. The union of the two sets is all of the colored region.

Example 9.1.14 *Suppose we have the following sets:*

$$A = \{1, 2, 3\}$$

$$B = \{1, 3, 5\}$$

$$C = \{2, 4\}$$

- $A \cup B = \{1, 2, 3, 5\}$
- $A \cap C = \{2\}$
- $\bar{A} = \{4, 5\}$
- $\bar{B} = \{2, 4\} = C$

Note that $P(A) + P(\bar{A}) = 1$.

So how does this relate to probability?

Example 9.1.15 Using the previous example, what is the probability a randomly selected number is in set A ?

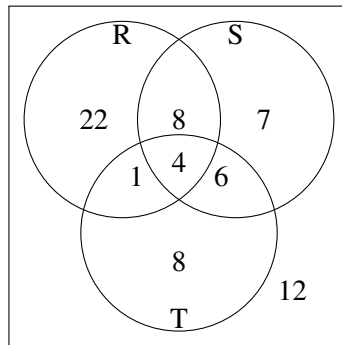
Solution Since we have 5 elements in our universal set and 3 in set A , we have $P(A) = \frac{3}{5}$.

So why use Venn diagrams if we can just count the number of elements in the set? They are helpful with visualizations ...

Example 9.1.16 Let R , S , and T be subsets of the universal set U . Use the data given below to fill in the number of elements in each basic region.

$n(U) = 68$, $n(R \cup S \cup T) = 56$, $n(R \cap \bar{T}) = 30$, $n(R \cap S) = 12$, $n(R \cap T) = 5$, $n(S \cap \bar{T}) = 15$, $n(\bar{S} \cap T) = 9$, $n(R \cap S \cap T) = 4$.

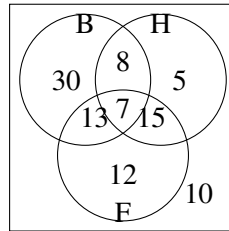
Solution To solve, we work from the inside out. First, we look at the common intersection of all three sets, which is given to contain 4 elements. Then, we look at the intersection of two sets. If we have $n(\bar{S} \cap T) = 9$, for example, but the intersection of all three contains 4 elements, there are 5 elements in this section of the diagram. Working through all regions in this manner,



So if we wanted to know what is the probability of selecting an element that is only in one set, we could now look at the diagram and see that there are 37 elements that satisfy this and $P(\text{one set}) = \frac{37}{68}$.

Example 9.1.17 A survey was made to determine the popularity of hockey, baseball and football. Of those surveyed, 35% watched hockey, 58% watched baseball and 47% watched football, 15% watched hockey and baseball, 20% watched baseball and football and 22% watched hockey and football. 7% watched all three.

- a. What is the probability a random person only watches baseball?
 b. What is the probability a random person watches hockey and football but not baseball?



So, we see $P(\text{baseball only}) = 30\%$ and $P(\text{hockey and football but no baseball}) = 15\%$.

9.1.3 Back to Discrete Probability

Example 9.1.18 Suppose we have a standard deck of cards. What is the probability of drawing one card and it is a queen or a 7?

- The sample space is $S = \{A, K, Q, J, 10, 9, 8, 7, 6, 5, 4, 3, 2\}$
- The event space is $E = \{Q, 7\}$
- The probability is therefore $P(E) = \frac{2}{13}$

Notice that there are different scenarios in the event, but they cannot happen at the same time. We say that events are mutually exclusive in a case such as this.

$$P(Q \cup 7) = P(Q) + P(7) = \frac{8}{52} = \frac{2}{13}$$

Definition 9.1.19 Two events A and B are said to be mutually exclusive if they have no elements in common. In symbols we say that $A \cap B = \emptyset$. Then,

$$P(A \cup B) = P(A) + P(B)$$

So, if we wanted to find the probability of rolling one die and getting a 3 or a 4, we would have mutually exclusive events because we cannot roll a 3 *and* a 4 at the same time. This will have implications on the probability, once we get to the rules.

9.1.4 The Basic Rules

For an event A ,

- $0 \leq P(A) \leq 1$
- $\sum P(A_i) = 1$ for mutually exclusive events A_i
- If $P(A) = 0$ we say A is the impossible event.

- If $P(A) = 1$ we say A is the certain event.
- For any event A , $P(\bar{A}) = 1 - P(A)$. This is the complement rule.

Example 9.1.20 Choose a student in grades 9-12 at random and ask if he or she is studying a language besides English. Here is the distribution:

Language	Spanish	French	German	Other	None
Probability	.26	.09	.03	.03	.59

Solution We first want to make sure that this is a legitimate probability by making sure the basic probability rules are satisfied. They are all here, so we could ask some questions.

What is the probability that a randomly selected student is studying a language other than English? we find this by adding up the percent of students studying other languages, which is .41 or 41%.

We could also ask questions that have us find probabilities like $Pr(\text{French} \cup \text{German} \cup \text{Spanish})$. To solve this, since it is asking us to find those who study French *or* German *or* Spanish, we would just need to add the appropriate probabilities to get .38. Alternatively, we could have looked at our answer in the last part and subtracted the .03 that represented other languages and we would have reached the same probability.

Example 9.1.21 Suppose the chances of you getting into an accident on your way to school are $\frac{1}{100}$ on any given day. What is the probability you get into an accident on the way to school at least once in a 3 day span?

Solution We wouldn't want to solve this directly because there are too many scenarios to consider. We would have to consider an accident on any one day in a three day span, then twice in any three day span and then finally all three days in any three day span, and there are many three day spans we would have to look at. But if we use the complement rule instead, this becomes much easier. Since there is a $\frac{1}{100}$ chance of an accident, there is a $\frac{99}{100}$ chance of no accident.

$$P(\text{no accident}) = \frac{99}{100} \cdot \frac{99}{100} \cdot \frac{99}{100} = \frac{970229}{1000000}$$

which leads to

$$P(\overline{\text{no accident}}) = 1 - \frac{970229}{1000000} = \frac{29771}{1000000}$$

That is, the chances of getting into at least one accident are $\approx .0298$.

Example 9.1.22 Suppose you have a bowl with 8 oranges, 6 apples and 3 peaches on the counter. You randomly grab a piece of fruit on your way out of the house. What is the probability it is not an apple?

Solution This is another case where complements come into play. We know

$$P(O \cup P) = \frac{8}{17} + \frac{3}{17} = \frac{11}{17}$$

and that getting an apple is the complement of getting an orange or a peach, so

$$P(A) = 1 - P(O \cup P) = 1 - \frac{11}{17} = \frac{6}{17}$$

So

$$P(\bar{A}) = 1 - P(A) = 1 - \frac{6}{17} = \frac{11}{17}$$

Now, you may be thinking ‘why did we do two extra steps to get back to the same answer?’ and the reason is the thought process - not all problems will get us back to the exact same answer like this but the way we rationalized through this will always work.

Example 9.1.23 Suppose there is a $\frac{1}{10}$ chance it snows. If it does snow, there is a $\frac{2}{9}$ chance you study. If it does not snow, there is a $\frac{3}{5}$ chance you study. What is the probability you study?

Solution Well, we can quickly see that

$$P(\text{snow} \cap \text{study}) = \frac{1}{10} \cdot \frac{2}{9} = \frac{1}{45}$$

because we have an ‘and’ situation, so we multiply the events’ probabilities. Similarly, we have

$$P(\text{no snow} \cap \text{study}) = \frac{9}{10} \cdot \frac{3}{5} = \frac{27}{50}$$

But how do we combine these? Since we cannot have both scenarios happening at the same time, we add these disjoint probabilities.

$$P(\text{study}) = \frac{1}{45} + \frac{27}{50} = \frac{10}{450} + \frac{243}{450} = \frac{253}{450}$$

Rule 9.1.24 For mutually exclusive events A and B , $P(A \cup B) = P(A) + P(B)$.

This becomes more complicated if the events are not mutually exclusive, however, as we have to figure out how to deal with the shared elements.

Example 9.1.25 What is the probability in a standard deck of cards of randomly selecting a card that is a queen or a spade?

Solution If this was previous way to think about the problem, we would simply have

$$P(Q \cup S) = P(Q) + P(S) = \frac{4}{52} + \frac{13}{52} = \frac{17}{52}$$

But there is a problem with this. Since we could have a queen of spades, we need to consider that we counted that card twice. To take this into account, we subtract off one of the times we counted the queen of spades.

$$P(Q \cup S) = P(Q) + P(S) - P(Q \cap S) = \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}$$

This idea is called the Principle of Inclusion-Exclusion.

Definition 9.1.26 If two events A and B are not mutually exclusive, then

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Notice that if A and B are mutually exclusive, $A \cap B = \emptyset$. And $P(\emptyset) = 0$ since we are looking for the probability that nothingness occurs. This might seem strange, but if we think of this in terms of sets and that probability is the ratio of how many elements in the event to the number of elements in the set, there are no elements in the event and this gives 0 probability.

This rule does work with what we discussed earlier, though, because if A and B are mutually exclusive, we have

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B)$$

Example 9.1.27 Suppose we know that the percent of commuters on campus is .75 and that the percent of nursing major is .27. Are these events mutually exclusive?

Solution No, since $P(C) + P(N) = 1.02$ and we know it is impossible to have a probability that exceeds 100%.

Example 9.1.28 If I told you that the percent of students who are nursing majors and commute is .17, what percent of all SSU students either commute or are majoring in nursing?

Solution Since we know from the previous example that the events are not mutually exclusive, we use the Principle of Inclusion-Exclusion here.

$$P(N \cup C) = P(N) + P(C) - P(N \cap C) = .27 + .75 - .17 = .85$$

Example 9.1.29 If we have an urn that contains 5 white ping pong balls, 4 black ping pong balls and 8 red ping pong balls, how many more white ones do we need to add to that the probability of taking one out and getting a white one is .40?

Solution Right now, we have

$$P(W) = \frac{5}{17}$$

What we want is to have a probability of .4. The way to approach this is to look at ratios that equal .4.

$$P(W) = .4 = \frac{4}{10} = \frac{8}{20} = \frac{12}{30} = \dots$$

In order for us to get where we want, we have to do equal additions to the numerator and denominator until we get a ratio fitting the form above. If we add 3 white ones, we will have the desired probability. This means we will have a total of 8 white ones but 20 total - for each white one we add, we will be adding one to the total number as well.

One final point. What we always want to keep in mind is that probability only works because we have randomness - we don't know which item will be selected ahead of time - and because we have no memory. The reason that if you flip a coin and get tails that you don't automatically get heads next time is because the coin flip could care less what happened last time and only has that the probability of any one flip of .5.

9.1.5 Conditional Probability

The idea behind conditional probability is that we know some information that affects the probability of the situation. When we have numbers to work with, then we can use the following formula:

Rule 9.1.30 *Conditional Probability*

The probability of B given that we know A has already occurred is

$$P(B|A) = \frac{P(B \cap A)}{P(A)}$$

Example 9.1.31 Find the probability of A given B if $P(A) = .2$, $P(B) = .3$ and $P(A \cup B) = .25$.

Solution Seeing the word ‘given’ should tell us that this problem involves conditional probability. But we are not given $P(A \cap B)$. We need to use the Principle of Inclusion and Exclusion to find it.

$$\begin{aligned} P(A \cup B) &= P(A) + P(B) - P(A \cap B) \Rightarrow P(A \cap B) = P(A) + P(B) - P(A \cup B) \\ &= .2 + .3 - .25 = .25 \end{aligned}$$

This leads us to

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{.25}{.3} \approx .83$$

Example 9.1.32 Suppose a math teacher gave two tests in a class. 25% of the class passed both exams and 42% of the class passed the first exam. What percent of the students who passed the first test then passed the second test?

Solution We are looking for a conditional probability because we are looking for the percent of students who passed the second exam but we are only concerned with those students who already passed the first exam.

$$P(\text{Second}|\text{First}) = \frac{P(\text{First} \cap \text{Second})}{P(\text{First})} = \frac{.25}{.42} = .6$$

So, 60% of the class who passed the first test went on to pass the second one.

Example 9.1.33 A jar contains black and white marbles. Two marbles are chosen without replacement. The probability of selecting a black marble and then a white marble is 0.34, and the probability of selecting a black marble on the first draw is 0.47. What is the probability of selecting a white marble on the second draw, given that the first marble drawn was black?

Solution Again, we see the work ‘given’ and we think conditional probability.

$$P(\text{White}|\text{Black}) = \frac{P(\text{Black} \cap \text{White})}{P(\text{Black})} = \frac{.34}{.47} = .72$$

The probability is 72%.

Example 9.1.34 Suppose $P(A) = .4$, $P(B) = .25$ and $P(A|B) = .5$. What is $P(B|A)$?

Solution We want $P(B|A)$ but are given $P(A \cap B)$. We need to use algebra to draw out the part we want.

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \Rightarrow P(A \cap B) = .5 \times .25 = .125$$

So, now we can find what we want.

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{.125}{.4} = .3125$$

Having prior knowledge does not mean we actually have information that changes the probability of a situation, though.

Definition 9.1.35 *Independence* means that having prior knowledge doesn't change the probability of the situation. If A and B are independent events, then $P(A) = P(A|B)$.

Example 9.1.36 *What is the probability of flipping a coin and getting tails if you know you already flipped and got tails the first time?*

Solution Using the conditional probability formula, we have

$$P(T|T) = \frac{P(T \cap T)}{P(T)} = \frac{\frac{1}{4}}{\frac{1}{2}} = \frac{1}{2} = P(T)$$

9.1.6 Exercises

1. Suppose you have a 4-sided die and you are going to roll it three times and record the sequence of numbers that come up.
 - (a) Write out the sample space using set notation.
 - (b) Write out the event space for the event of having all even numbers (remember, order will matter).
 - (c) What is the probability of rolling all even numbers?
2. Suppose you flip a fair coin 10 times. Describe the sample space (do not write out all of the possibilities).
3. You quickly grab a piece of fruit out of a bowl in your kitchen on your way out the door. There were 5 apples, 3 bananas and 4 oranges in the bowl. What is the probability you did not grab an orange, seeing how that would be difficult to eat while driving?

4. Given the sets

$$W = \{b, c, e\}, X = \{a, b, c, d, e\}, Y = \{a, b, c\}, Z = \{a, c, e\}$$

Find the following:

- (a) The universal set U
 - (b) $W \cup X$
 - (c) $Y \cap Z$
 - (d) $\overline{W} \cap Z$
 - (e) $\overline{X \cup Y}$
5. Find the probability, given a standard deck of 52 cards, that
 - (a) you draw one card and it is red.
 - (b) you draw one card and it is a club or a jack.
 - (c) you draw one card and it is not a face card.
 6. Is the percent of people who work on Wednesdays mutually exclusive with the percent of people who take classes on Wednesdays? Explain.
 7. Suppose that 35% of SSU students are taking an English class this semester and that 45% of SSU students are taking a math class. If 40% of SSU students are taking neither, what percent of students are taking a math and an English class this semester?
 8. You randomly poll people outside your dorm as to whether or not they will vote in the next election. So far, 13 have said they will and 4 have said they will not. If all of the people you will ask from now on will say yes, how many more will you need to ask so that the percent of people voting will be at least 90%?
 9. Suppose 50% of your friends drink coffee and 40% of your friends drink soda. Suppose further that 25% of your friends like both. What is the probability you randomly select a friend who drinks soda given that they drink coffee? Is this different than the probability a random friend drinks coffee given that they drinks soda?

9.1.7 Solutions

- Suppose you have a 4-sided die and you are going to roll it three times and record the sequence of numbers that come up.

(a)

$$S = \{(1, 1, 1), (1, 1, 2), (1, 1, 3), (1, 1, 4), (1, 2, 1), (1, 2, 2), (1, 2, 3), (1, 2, 4), \\ (1, 3, 1), (1, 3, 2), (1, 3, 3), (1, 3, 4), (1, 4, 1), (1, 4, 2), (1, 4, 3), (1, 4, 4), \\ (2, 1, 1), (2, 1, 2), (2, 1, 3), (2, 1, 4), (2, 2, 1), (2, 2, 2), (2, 2, 3), (2, 2, 4), \\ (2, 3, 1), (2, 3, 2), (2, 3, 3), (2, 3, 4), (2, 4, 1), (2, 4, 2), (2, 4, 3), (2, 4, 4), \\ (3, 1, 1), (3, 1, 2), (3, 1, 3), (3, 1, 4), (3, 2, 1), (3, 2, 2), (3, 2, 3), (3, 2, 4), \\ (3, 3, 1), (3, 3, 2), (3, 3, 3), (3, 3, 4), (3, 4, 1), (3, 4, 2), (3, 4, 3), (3, 4, 4), \\ (4, 1, 1), (4, 1, 2), (4, 1, 3), (4, 1, 4), (4, 2, 1), (4, 2, 2), (4, 2, 3), (4, 2, 4), \\ (4, 3, 1), (4, 3, 2), (4, 3, 3), (4, 3, 4), (4, 4, 1), (4, 4, 2), (4, 4, 3), (4, 4, 4)\}$$

(b) $E = \{(2, 2, 2), (4, 4, 4)\}$

(c) $P(E) = \frac{N(E)}{N(S)} = \frac{2}{64} = \frac{1}{32}$

- The sample space consists of sequences of heads and tails, where there are a total of 10 H 's and T 's. The number of H 's ranges from 0 to 10, with the number of T 's being the complement of the number of H 's.

- There are a total of 12 pieces of fruit in the bowl and $P(\text{orange}) = \frac{4}{12}$. So, $P(\overline{\text{orange}}) = 1 - \frac{4}{12} = \frac{8}{12} = \frac{2}{3}$.

- (a) $U = \{a, b, c, d, e\}$

(b) $W \cup X = \{a, b, c, d, e\} = U$

(c) $Y \cap Z = \{a, c\}$

(d) $\overline{W} \cap Z = \{a\}$

(e) $\overline{X \cup Y} = \{\}$

- Find the probability, given a standard deck of 52 cards, that

(a) $P(\text{red}) = \frac{26}{52}$

(b) $P(\text{club} \cup \text{jack}) = \frac{13}{52} + \frac{4}{52} - \frac{1}{52} = \frac{16}{52}$

(c) $P(\overline{\text{face card}}) = 1 - P(\text{face card}) = 1 - \frac{12}{52} = \frac{40}{52}$

- The events are not mutually exclusive because a person could certainly work on Wednesdays and take classes on Wednesdays.

- $P(M \cap E) = P(M) + P(E) - P(M \cup E) = .45 + .35 - .6 = .2$

- We need to exceed $\frac{9}{10} = \frac{18}{20} = \frac{27}{30} = \frac{36}{40}$. Notice that the number of people who said no is 4, and if everyone from this point on says yes, there will always be 4 people who said no. So, if we asked 23 more people, this would give us the ration $\frac{36}{40}$ that we seek.

9. $P(S | C) = P(S \cap C) \cdot P(C) = .25 \cdot .4 = .1$

The probabilities will be different because the probability of liking coffee and liking soda are different. Since the probability of liking coffee and soda are the same in both situations, the difference is that we are multiplying by different probabilities (.5 or .4) depending on which conditional probability we need. We also know that the probabilities will be different because the events are not independent; that is, $P(S \cap C) = .25 \neq .2 = P(S) \cdot P(C)$.

9.2 Tree Diagrams and the Multiplication Rule

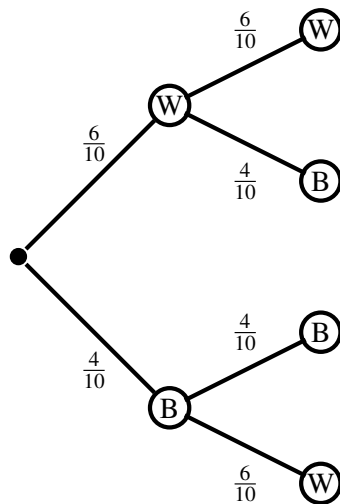
When we think of multiplying probabilities, we should be thinking ‘and’. We have seen this earlier but now we visualize probabilities using tree diagrams.

Example 9.2.1 *Suppose we have an urn that has 6 white ping pong balls and 4 black ping pong balls. Suppose we select two balls at random. What is the probability they are both white if*

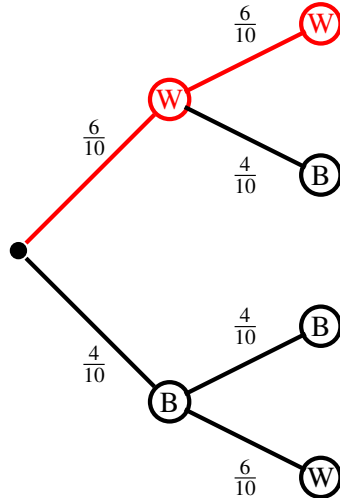
1. *there is replacement?*
2. *there is no replacement?*

Solution When we talk about replacement, it is exactly what it sounds like. Picture that you take a ping pong ball out of the urn, record what color it is you got, and then put it back. This would mean we are replacing the ball and the probability of getting white the next pull would be exactly the same. In contrast, if we take the ball out, record the color and then throw it away, we are not replacing and it makes the probability of getting a white one one the next pull different.

So first, we will set up the tree diagram for the replacement situation. Notice that at each stage, the probabilities (viewed vertically) add up to one. Also notice that the denominator in each case is 10, which is how many ping pong balls we started with.



So, to find the probability both are white, we follow the path that has us select white ping pong balls on both pulls.



To calculate a probability, we can apply the Multiplication Property.

Definition 9.2.2 *The Multiplication Property* The probability of an event at the end of a path is equal to the probabilities along the path multiplied together.

The reason this works is because the events along the path are independent. As we said before, independent events are not affected by prior events. A more useful formula for the probability of independent events is

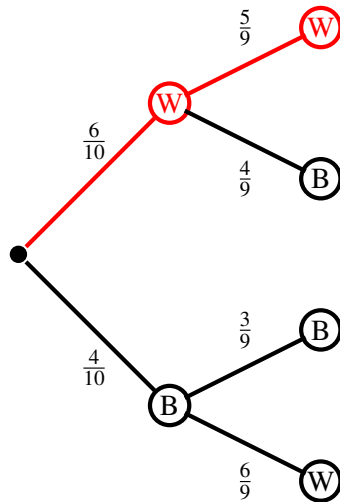
Rule 9.2.3 *If events A and B are independent, then*

$$P(A \cap B) = P(A) \cdot P(B)$$

So in our example, we want to know

$$P(W \cap W) = \frac{6}{10} \cdot \frac{6}{10} = \frac{36}{100} = \frac{9}{25}$$

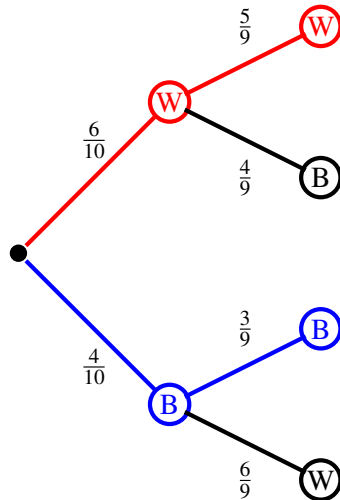
Now, if we do not have replacement, we will still use the multiplication rule and we will still follow the path that leads us to two white balls, but some of the probabilities along the branches changes.



Notice the probabilities in the second level all have 9 in the denominator. This is because we have one less ping pong ball in the urn when we get to the second pull. So, we have

$$P(W \cap W) = \frac{6}{10} \cdot \frac{5}{9} = \frac{30}{90} = \frac{1}{3}$$

What if we wanted to get the same color with no replacement? We are not being specific about which color, which means there is more than one scenario that works. We cannot get both white and both black at the same time, so we need to find the probability of each possibility.



Since this is an 'or' situation, we now apply the addition principle.

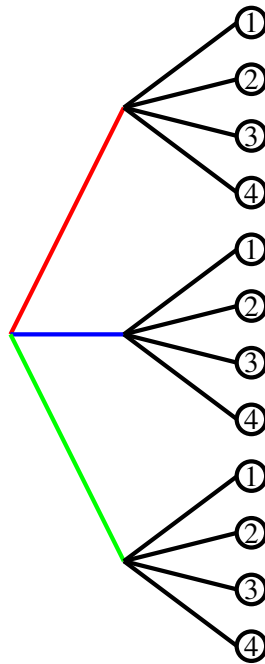
$$P((W \cap W) \cup (B \cap B)) = \frac{30}{90} + \frac{12}{90} = \frac{42}{90} = \frac{7}{15}$$

9.2.1 Exercises

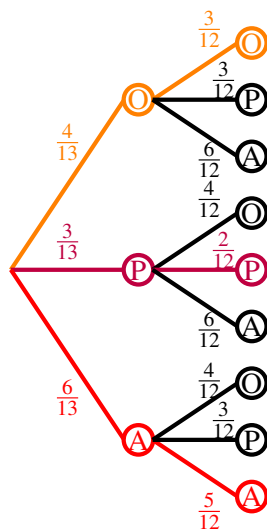
1. At the Topsfield Fair, a game asks you to choose a color from the set $\{red, blue, green\}$ and then a number from the set $\{1, 2, 3, 4\}$ to see what prize you win. Draw a tree diagram for the possible outcomes of this game.
2. As you run out of the house, you grab two pieces of fruit from a bowl containing 4 oranges, 3 plums and 6 apples. Construct a probability tree diagram and use it to find the probability both pieces of fruit are the same kind.
3. Suppose we have three urns: one contains three white balls and two black ones. The second one has 2 of each and the third one has 3 black ones. Suppose we randomly select an urn and then select one ball. What is the probability the ball is black?
4. A family has three children, sons Owen and Jack, and daughter Stephanie. Suppose the family has game night on Tuesdays and movie night on Saturdays and that they randomly select a child to choose the entertainment for the night each time. If the same child can be selected for both in a given week, create a tree diagram for this situation and find the probability that one boy and one girl are selected this week.
5. When Charlie is coaching your baseball team, there is a 50% chance you will pitch but when Jack is coaching, there is only a 30% chance you will pitch. You have a game today and don't know which coach but there is a 60% chance that Charlie will be the coach. What is the probability you pitch today? (Hint: this requires multiple probability ideas from this section to solve.)

9.2.2 Solutions

1.



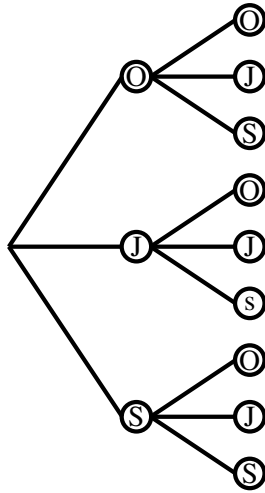
2.



$$P(\text{same fruit}) = \frac{4}{13} \cdot \frac{3}{12} + \frac{3}{13} \cdot \frac{2}{12} + \frac{6}{13} \cdot \frac{3}{12} = \frac{48}{132}$$

$$3. P(\text{black}) = \frac{1}{3} \left(\frac{2}{5} + \frac{2}{4} + \frac{3}{3} \right) = \frac{114}{180}$$

4.



If a boy is selected first, then in either of these two situations, there is a $\frac{1}{3}$ chance of selecting the female second. If the female is selected first, then there is a $\frac{2}{3}$ chance of selecting a male second. And, there is a $\frac{1}{3}$ chance of any of the children being selected first. This gives

$$P(\text{one male and one female}) = 2 \cdot \frac{1}{3} \cdot \frac{1}{3} + \frac{1}{3} \cdot \frac{2}{3} = \frac{4}{9}$$

5. Let P be the event you pitch, let C be the event Charlie coaches and let J be the event Jack is coaching. We have multiple disjoint conditional probabilities to consider here. We could model this with a tree diagram if we chose but here we will just worry about the probabilities through formulas.

$$P(P \cap C) = P(P | C) \cdot P(C) = .5 \cdot .6 = .3$$

$$P(P \cap J) = P(P | J) \cdot P(J) = .3 \cdot .4 = .12$$

$$P(P) = P(P \cap C) \cup P(P \cap J) = .42$$

9.3 Binomial Probabilities

Earlier in this chapter, we talked of discrete probabilities and said that the probability was found by taking the ratio of the number of elements in the event space and the number of elements in the sample space. But these situations relied on us finding the likelihood of one event that happened one way. When there were scenarios, we have to consider all of the different ways that the event could occur.

For example, suppose we wanted to flip a coin twice. If we asked for the probability that we get tails on the first flip and then heads on the second, we could just multiply the probability of the events ($\frac{1}{2}$) in each case to find the probability of the event. But what if we instead ask for the probability that we wanted to find was for flipping a coin twice and getting one head and one tail. Since we are not specific about the sequence, we have to consider all possible sequences of heads and tails that have exactly one of each. So even though the probability of each event is $\frac{1}{2}$ and we multiply them to find the probability of them both happening, $\frac{1}{4}$ is not the probability because HT and TH both satisfy what we want in our event. The probability we want is $\frac{1}{4}$ for each and since they are independent events, we can add to get $\frac{1}{2}$.

Now, this seems simple enough - we can find the probability of a specific sequence and then just count all of the possible sequences. when we think of flipping a coin twice, we only have $S = \{HH, HT, TH, TT\}$, so only four possibilities, two of which work. But ... what if we wanted to know about flipping a coin 20 times? 50 times? We certainly don't want to have to write out all of those possible sequences. This is where binomial probabilities come into play.

9.3.1 Rules for Using Binomial Probabilities

So when should we use binomial probabilities? We do when we have exactly two mutually exclusive outcomes over a known and fixed number of trials. We often see this type of probability associated with looking at how many tails we get over a series of coin flips or probability of hitting a certain number of free throws when given a fixed number of attempts. We also need a short memory in that the trials must be independent.

The formal criteria is as follows:

1. We must have Bernoulli trials

Definition 9.3.1 *Bernoulli trials are independent, repeated trials with exactly two possible outcomes. We call these outcomes successes and failures.*

2. The number of observations n must be fixed. We have to know exactly how many trials there will be before we begin and it cannot be a variable number. It cannot be an open-ended question we are asking.
3. The probability of 'success' is p and the probability of 'failure' is $q = 1 - p$. These must be fixed as well.

If all of these criteria are met, we say the random variable X has the binomial distribution and the parameters we need to give are n and p . The standard notation is $B(n, p)$.

Before we go too far, it should be pointed out that we don't want to get hung up on 'success' v. 'failure'. These are just the terms we use to distinguish between the event we are concerned with and the complementary event.

Formula 9.3.2 If $X \sim B(n, p)$ then

$$P(X = k) = \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

Remember, $k!$ is the notation for a factorial, which indicates that we need to multiple the first k positive integers together.

Example 9.3.3 Suppose you flip a fair coin 5 times. What is the probability that we get tails exactly 3 times.

Solution We can see that this is a situation suitable for our binomial probability formula because we have a fixed number of trials (5), we have the probability of 'success' and 'failure' are fixed (fair coin, so $p = .5$) and the trials are independent.

The reason we need the binomial probability formula here is because there are many ways we can get 3 tails in 5 flips and we would have to find all of those ways and add the respective probabilities to get our desired answer. By using this formula, the coefficient gives us the number of ways the event can occur. Therefore, we have

$$\begin{aligned} P(X = 3) &= \frac{5!}{3! \cdot (5-3)!} \left(\frac{1}{2}\right)^3 \left(1 - \frac{1}{2}\right)^2 \\ &= \frac{5!}{3! \cdot 2!} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^2 \\ &= \frac{120}{6 \cdot 2} \cdot \frac{1}{8} \cdot \frac{1}{4} \\ &= 10 \cdot \frac{1}{32} \\ &= \frac{5}{16} \end{aligned}$$

So, the probability that we get 3 tails in 5 flips is $\frac{5}{16} = .3125$.

Example 9.3.4 The game of roulette consists of wedges on a wheel, where each wedge has a different color and number. The even integers 2-36 are black and the odd integers 1-35 are red. The remaining two wedges are green and are assigned 0 and 00. Suppose you are at the casino and decide to bet black 3 times in a row. What is the probability that you won all three attempts?

Solution This is a binomial probability because we either win or lose and there is no third option. The number of trials here is fixed at $n = 3$ and the probability of success is always $P(\text{black}) = \frac{18}{38}$. So, we have

$$\begin{aligned} P(X = 3) &= \frac{3!}{3!(3-3)!} \left(\frac{18}{38}\right)^3 \left(\frac{20}{38}\right)^0 \\ &= \frac{6}{6 \cdot 1} \frac{729}{6859} \cdot 1 \approx .106 \end{aligned}$$

So, there is a 10.6% chance you will all three times.

Note: $0! = 1$.

Example 9.3.5 *What is the probability that, while sitting at that same roulette wheel, you win at least once in three attempts if you bet green?*

Solution We want to know the probability of winning at least once, which is the complement of not winning at all. This is easier, since it is just one probability to find instead of finding the probability of winning exactly one time, then two times and finally three times and adding them together since those would be mutually exclusive events.

Since we want to bet green, we would have $p = \frac{2}{38}$ and we are still working with $n = 3$.

$$\begin{aligned} P(X = 0) &= \frac{3!}{0!(3-1)!} \left(\frac{2}{38}\right)^0 \left(\frac{36}{38}\right)^3 \\ &= \frac{6}{1 \cdot 6} \frac{5832}{6859} \\ &\approx .8502 \end{aligned}$$

So, we would not win at all 85.02% of the time. But, we wanted to know the probability we won at least one, which is the complementary event. So, we have

$$P(X \geq 1) = 1 - .8502 = .1498$$

That is, we have a 14.98% chance to win at least once betting on green.

Example 9.3.6 *Suppose you find that you only eat breakfast on average 20% at home of the time. What is the probability that you eat breakfast at home 5 times in a week?*

Solution We will call it a ‘success’ if you eat breakfast at home, so $p = .2$ here. We have a fixed n at 7, since there are that many days in a week. Whether you eat breakfast at home one day has no impact on another, so the trials are independent and therefore we can use our binomial probability formula.

$$\begin{aligned} P(X = 5) &= \frac{7!}{5! \cdot 2!} (.2)^5 (.8)^2 \\ &= \frac{5040}{120 \cdot 2} \cdot \frac{1}{3125} \cdot \frac{64}{100} \\ &= \frac{336}{71875} \\ &\approx .00467 \end{aligned}$$

There is a .467% chance you eat breakfast at home five times in a week.

9.3.2 What Happens As n Gets Larger

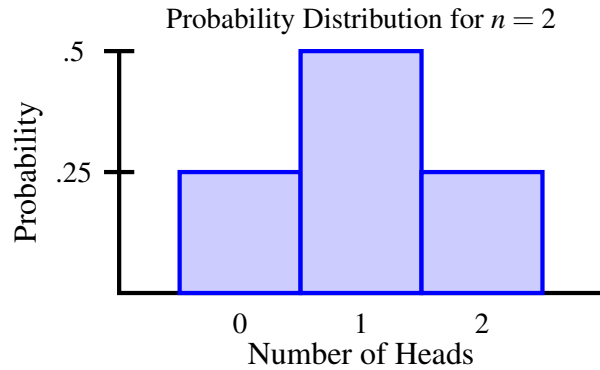
Let’s look at the probability distribution for the coin flip example for increasing values for n .

For $n = 2$, we have

$$P(X = 0) = \frac{2!}{0!(2-0)!} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^2 = .25$$

$$P(X = 1) = \frac{2!}{1!(2-1)!} \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^1 = .5$$

$$P(X = 2) = \frac{2!}{2!(2-2)!} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^0 = .25$$



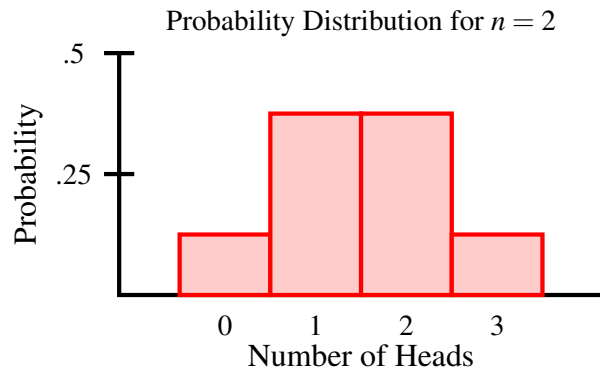
And for $n = 3$, we have

$$P(X = 0) = \frac{3!}{0!(3-0)!} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^3 = .125$$

$$P(X = 1) = \frac{3!}{1!(3-1)!} \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^2 = .375$$

$$P(X = 2) = \frac{3!}{2!(3-2)!} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^1 = .375$$

$$P(X = 3) = \frac{3!}{3!(3-3)!} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^0 = .125$$



And now, for $n = 4$, we have

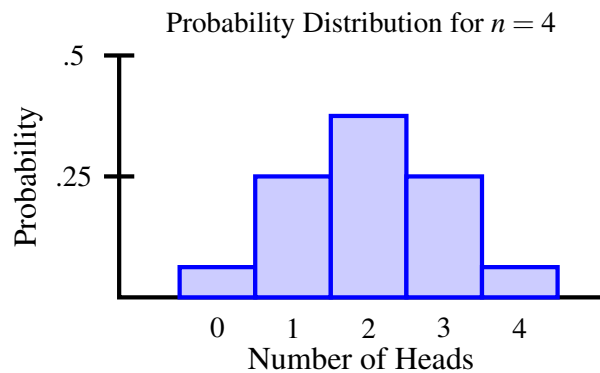
$$P(X = 0) = \frac{4!}{0!(4-0)!} \left(\frac{1}{2}\right)^0 \left(\frac{1}{2}\right)^4 = .0625$$

$$P(X = 1) = \frac{4!}{1!(4-1)!} \left(\frac{1}{2}\right)^1 \left(\frac{1}{2}\right)^3 = .25$$

$$P(X = 2) = \frac{4!}{2!(4-2)!} \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^2 = .375$$

$$P(X = 3) = \frac{4!}{3!(4-3)!} \left(\frac{1}{2}\right)^3 \left(\frac{1}{2}\right)^1 = .25$$

$$P(X = 4) = \frac{4!}{4!(4-4)!} \left(\frac{1}{2}\right)^4 \left(\frac{1}{2}\right)^0 = .0625$$



There are a couple of things to notice:

- The sum of the probabilities in each case is always 1
- The distributions are all symmetric
- The distributions are becoming more 'bell-shaped'

We will see in the next chapter why this is so important.

9.3.3 Central Tendencies of Binomial Distributions

We can calculate descriptive statistics for binomial distributions but not using the formulas we used earlier in this course.

Definition 9.3.7 The mean μ for the binomial distribution $B(n, p)$ is given by

$$\mu = np$$

The rationale behind this is fairly straight forward. Since there are only two outcomes, and the probability of success is p , we would expect np successes in n trials.

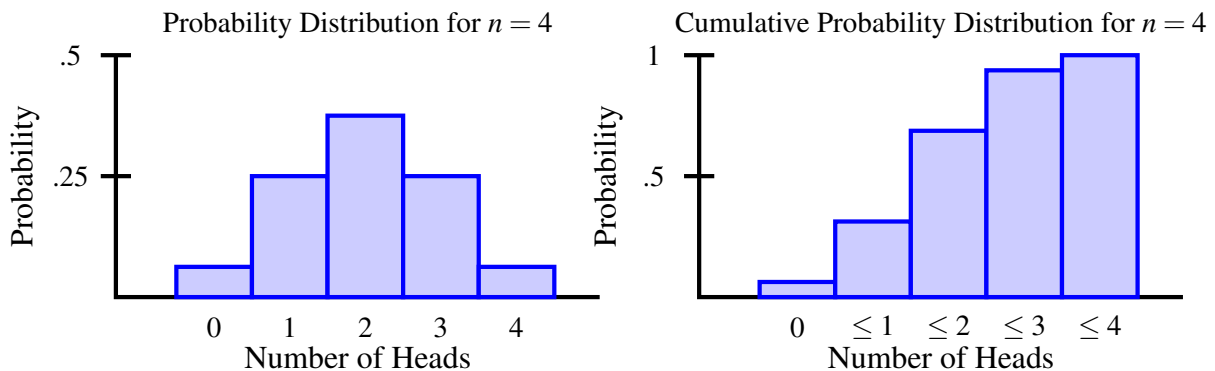
Definition 9.3.8 The standard deviation σ for the binomial distribution $B(n, p)$ is given by

$$\sigma = \sqrt{np(1-p)}$$

The formula here is logical too. The only variation in the outcome is between the probability of success, p , and the probability of failure, $1-p$. Over n trials, the variation would therefore be $np(1-p)$. And, the standard deviation is the square root of variance, so ...

9.3.4 Cumulative Binomial Probabilities

In the examples earlier, we were concerned with one event. Here, we are concerned with all of the events up to a certain point or all events after a certain point. Visually, if we look back at the coin flipping situation, we have



To solve these, we can solve directly, which will involve adding probabilities. This is not a problem, however, because the events are mutually exclusive. For example, we cannot get *exactly* three heads and *exactly* four heads at the same time. Alternatively, we could consider using compliments as that makes it easier in some situations.

Example 9.3.9 You head back to the roulette table and decide this time to bet red. What is the probability you win at least twice in three games?

Solution The probability of success on any one roll of the wheel is $\frac{18}{38}$ as before. Since we have two scenarios

in our situation and two that are not, using compliments does not save us any work.

$$\begin{aligned}
 P(X = 2) &= \frac{3!}{2!(3-2)!} \left(\frac{18}{38}\right)^2 \left(\frac{20}{38}\right)^1 \\
 &= \frac{6}{2 \cdot 1} \cdot \frac{81}{361} \cdot \frac{10}{19} \\
 &\approx .3543 \\
 P(X = 3) &= \frac{3!}{3!(3-3)!} \left(\frac{18}{38}\right)^3 \left(\frac{20}{38}\right)^0 \\
 &= \frac{6}{6 \cdot 1} \cdot \frac{729}{6859} \\
 &\approx .1063
 \end{aligned}$$

Since the events are mutually exclusive, we can add the probabilities, giving

$$P(X \geq 2) = .3543 + .1063 = .4606$$

Example 9.3.10 Suppose the probability of a newborn being female is 53%. What is the probability that a family of 5 children has at least 3 girls?

Solution We have three scenarios to find probabilities for if we find the solution directly or if we use compliments, so we might as well go directly at this one.

$$\begin{aligned}
 P(X = 3) &= \frac{5!}{3!2!} (.53)^3 (.47)^2 \approx .3289 \\
 P(X = 4) &= \frac{5!}{4!1!} (.53)^4 (.47)^1 \approx .1854 \\
 P(X = 5) &= \frac{5!}{5!0!} (.53)^5 (.47)^0 \approx .0418
 \end{aligned}$$

Since the events are mutually exclusive, we can add the associated probabilities, giving

$$P(X \geq 3) = .3289 + .1854 + .0418 = .5552$$

These problems can be serious, depending on how many different scenarios need to be considered. If there was only a way to use technology ...

9.3.5 Binomial Probabilities on the TI-Series Calculator

There are two choices when finding binomial probabilities in the calculator. The difference between the two is whether or not we are looking for the probability of a value or of a range. We will illustrate each through examples. Whereas we can use the formulas to find these probabilities, technology makes it much easier for us when we have large numbers of trials.

Example 9.3.11 Find the probability that we get exactly 12 heads in 20 flips of a fair coin.

Solution First, we go onto the Distribution menu (2^{nd} and **vars**) and scroll to **binompdf**, which stands for 'binomial probability distribution function'.

```
NORMAL FLOAT AUTO REAL RADIAN MP
DISTR DRAW
3↑invNorm(
4:invT(
5:tPdf(
6:tcdf(
7:χ²pdf(
8:χ²cdf(
9:Fpdf(
0:Fcdf(
A binompdf(
```

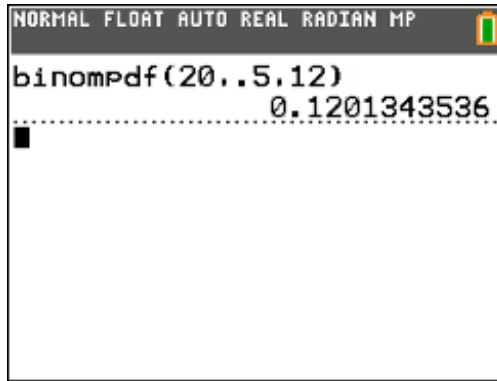
When we press **enter**, we are given an option screen where we can input the parameters we need, which include n , p and x , which is the number of successes. The following screen shows the values we want here.

```
NORMAL FLOAT AUTO REAL RADIAN MP
binompdf
trials:2
p:.5
x value:12
Paste
```

After scrolling to **Paste** and pressing enter, we have

```
NORMAL FLOAT AUTO REAL RADIAN MP
binompdf(20,.5,12)
```

And pressing **enter** again gives the desired probability.

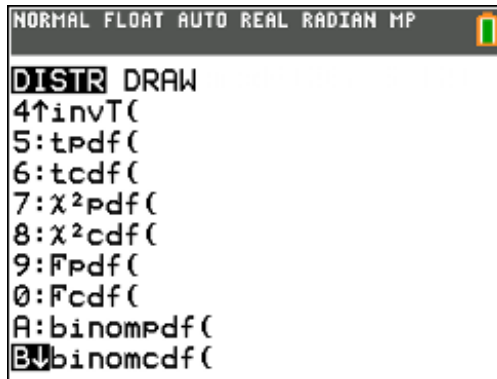


Note: If you have a different version of the calculator, the syntax is

$$\text{binompdf}(n, p, x)$$

Example 9.3.12 Find the probability that, in 20 flips of a fair coin, we get at least 12 heads.

Solution Here, we need the cumulative probability, so we need to use the next function in that same window.



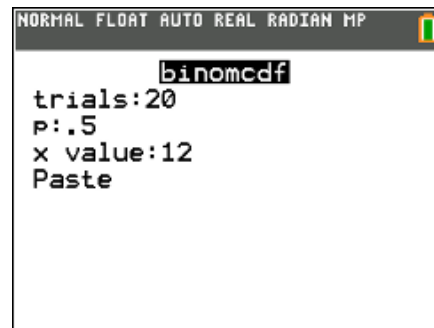
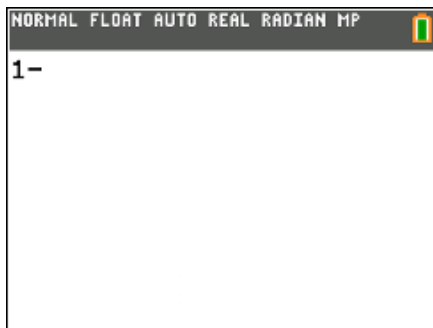
This function represents the binomial cumulative distribution function. But before we get there, we have to be careful with the input here. The syntax

$$\text{binomcdf}(n, p, k)$$

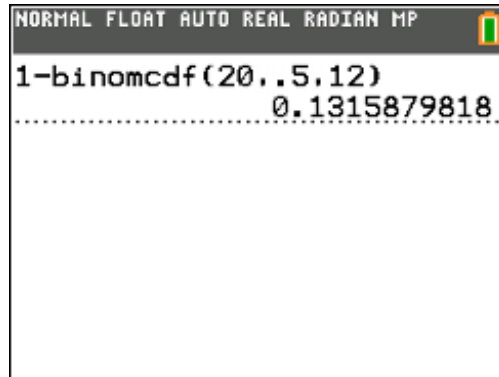
gives the probability $P(X \leq k)$. So, if we want $P(X \geq k)$, we need the compliment. In our case, we want

$$1 - \text{binomcdf}(20, .5, 12)$$

What we will do is start by putting 1- on the screen and then go into the menu.



Now, we input the values we need, scroll to **Paste** as we did before and then press *enter* again to get the probability we need.



Note: Since we are talking about discrete values, $P(X \geq k)$ is the same as $P(X > k - 1)$. This is how we adjust when finding these probabilities in the calculator.

9.3.6 Exercises

1. Determine which of the following are situations in which we should use binomial probabilities:
 - (a) You flip a coin 10 times and want to find the probability that there are exactly 3 tails.
 - (b) You flip a coin 10 times and want to find the probability that there are at least 3 tails.
 - (c) You take 6 free throws in a game and want to find the probability that you make 5 of them.
 - (d) You roll a standard die 20 times and want to find the probability that you roll a 5 nine times.
 - (e) You want to find the probability that it will rain on 3 of the next 4 Fridays.
2. 35% of Salem State students identify themselves as athletic. What is the probability that you randomly select 6 students and 4 of them identify themselves as athletic?
3. If we can assume that the probability of snow on the first week of January is 20% for each day, what is the probability it snows 5 times during that first week?
4. Suppose a certain new cold medicine is found to be effective in 75% of adults. What is the probability that among 8 randomly selected adults with colds, the drug is effective for exactly 6 of them?
5. For the last example, find the mean and the standard deviation of the distribution.
6. Suppose a certain new cold medicine is found to be effective in 75% of adults. What is the probability that among 8 randomly selected adults with colds, the drug is effective for at least 6 of them?
7. Suppose that among two teams in your intramural kickball league, there is a 25% chance team 1 beats team 2. If they play 5 times over the course of the season, what is the probability that team 1 wins exactly two games?

9.3.7 Solutions

1. Determine which of the following are situations in which we should use binomial probabilities:

- (a) Binomial probability
- (b) Not a binomial probability because there are multiple values for the desired number of successes. This could be solved using multiple binomial probabilities and the addition rule, however.
- (c) This depends on how you establish the probability. In real life, the probability of making a free throw changes as your percentage of free throw success changes. But, if we assume a fixed probability of success for the game, we can solve this using the ideas of binomial probabilities.
- (d) Binomial probability
- (e) Not a binomial probability - the probability of 'success' (rain) is not fixed

2. $P(X = 4) = {}_6C_4(.35)^4(.65)^2 = .095$

3. $P(X = 5) = {}_7C_5(.2)^5(.8)^2 = .0043$

4. $P(X = 6) = {}_8C_6(.75)^6(.25)^2 = .311$

5.

$$\begin{aligned}\bar{x} &= 6(.75) = 4.5 \\ \sigma &= \sqrt{6(.75)(.25)} = 1.061\end{aligned}$$

6.

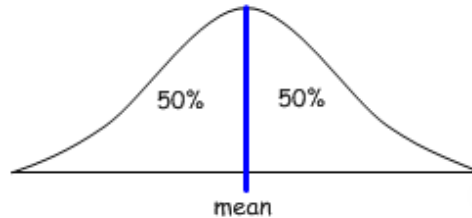
$$\begin{aligned}P(X \geq 6) &= P(X = 6) + P(X = 7) + P(X = 8) \\ &= {}_8C_6(.75)^6(.25)^2 + {}_8C_7(.75)^7(.25)^1 + {}_8C_8(.75)^8(.25)^0 \\ &= .311 + .267 + .100 \\ &= .678\end{aligned}$$

7. $P(X = 2) = {}_5C_2(.25)^2(.75)^3 = .264$

Chapter 10

Normal Distributions

Earlier, we talked about some ways we can describe distributions in numerical and pictorial ways. But sometimes we want to use the data to explore the likelihood of an event occurring. If the data is irregular, this can be very difficult. But when the data follows some kind of regular pattern, we can use the mean and standard deviation to find the probability of an event. Real-life situations are governed by **Normal distributions**. These are the symmetric bell curves that we generally think of when we think of distributions. They look like

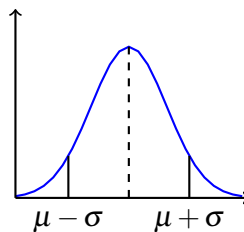


These are the curves which have the majority of the observations in the middle, indicated by the single peak, and then the observations get less and less frequent as we get further from the mean. The area under the curve represents 100% of the observations. Since the curve is symmetric, the mean and median are the same and therefore 50% of the observations lie above the mean and 50% lie below the mean. What this means for us is that we can predict the likelihood of an event occurring based on how far away from the mean that observation lies.

We will first look at how this works by approximating probabilities using the **Empirical Rule** and then we will look at how we can find more exact probabilities associated with this Normal distribution curve.

10.1 The Normal Distribution Curve

The Normal distribution curve probably looks familiar to you.

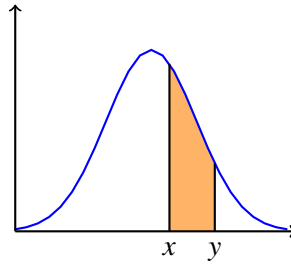


The Normal curve has the following properties:

1. Area under curve is 1
2. Mean equals the median
3. Unimodal
4. Bell-shaped
5. inflection points at ± 1 standard deviation

- 6. The mean of the distribution is given by μ
- 7. The standard deviation of the distribution is given by σ

So how exactly do we use this curve to find probabilities?



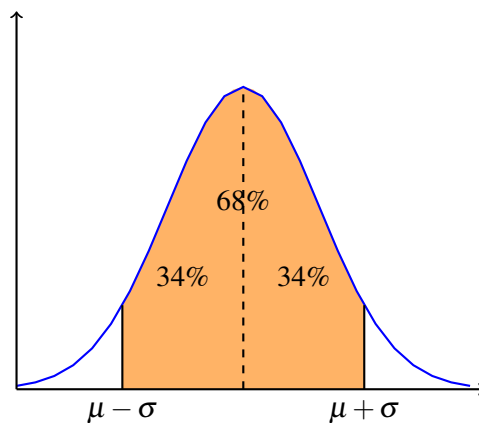
The probability of an observation occurring between x and y , we need a way to find this area.

10.2 The Empirical Rule

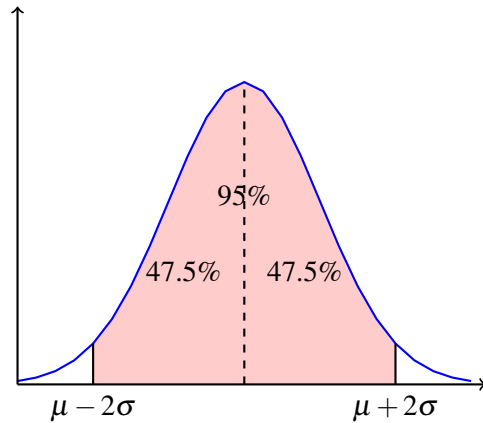
When we are dealing with exactly 1, 2 or 3 standard deviations from the mean, we know the percentages of observations that lie in a given range. This approximation tool is known as the Empirical Rule, or the 68 – 95 – 99.7 Rule.

We do have to be careful in making sure we have a symmetric, unimodal distribution if we are not told that the distribution is Normal. The probabilities associated here are only valid for Normal distributions.

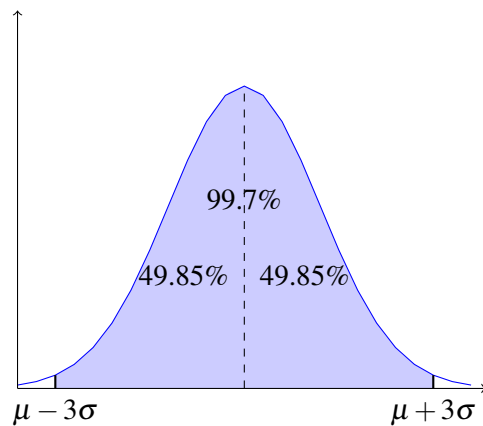
We know that in a Normal distribution with mean μ and standard deviation σ , 68% of all observations lie within one standard deviation of the mean.



We know that in a Normal distribution with mean μ and standard deviation σ that 95% of all observations lie within 2 standard deviations.



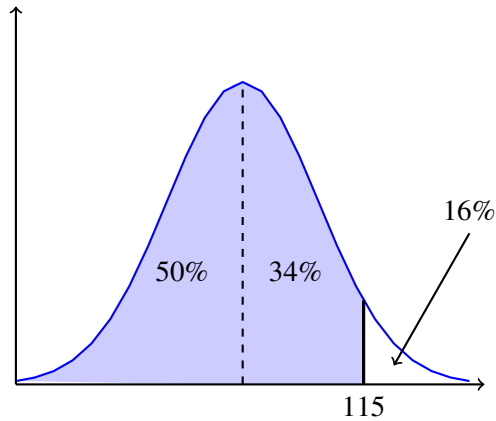
Finally, we know that in a Normal distribution with mean μ and standard deviation σ that 99.7% of all observations lie within 3 standard deviations of the mean.



Let's see how we use this by looking at all possible types of questions we could ask in the following examples.

Example 10.2.1 *IQ scores are Normally distributed with $X \sim N(100, 15)$. Find $P(X \geq 115)$.*

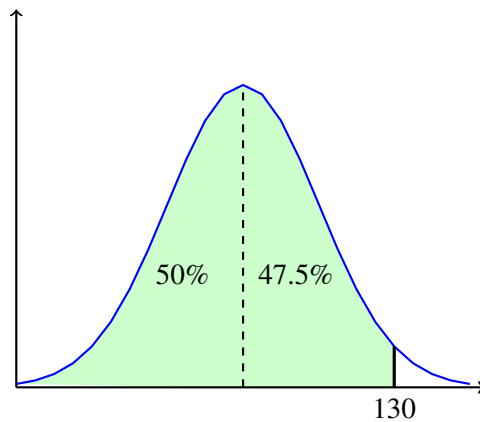
Solution $P(X \geq 115)$ means that we want to know how much of the curve lies more than one standard deviation above the mean. But, the way the percents are expressed with this rule, it is often easier to think about how much area is to the left of the line in question (here, 115 which is one standard deviation larger than 100) and then decide if we want to keep the area we shade or take the complement (what is left over out of 100%).



Here, we need the right tail, so $P(X \geq 115) = 16\%$.

Example 10.2.2 IQ scores are Normally distributed with $X \sim N(100, 15)$. Find $P(X \leq 130)$.

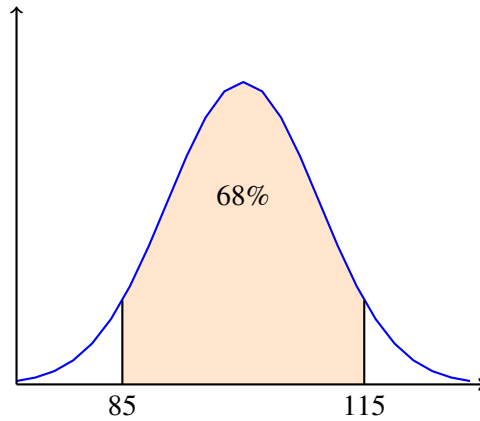
Solution Again, by being told that the distribution is Normal, we can use the bell curve and find the area of the shaded region. We want here to know how much area there is to the left of 130, which is 2 standard deviations above the mean of 100.



So, here, since we want to the left because we want the probability less than the given value, $P(X \leq 130) = 97.5\%$.

Example 10.2.3 IQ scores are Normally distributed with $X \sim N(100, 15)$. Find $P(85 \leq X \leq 115)$.

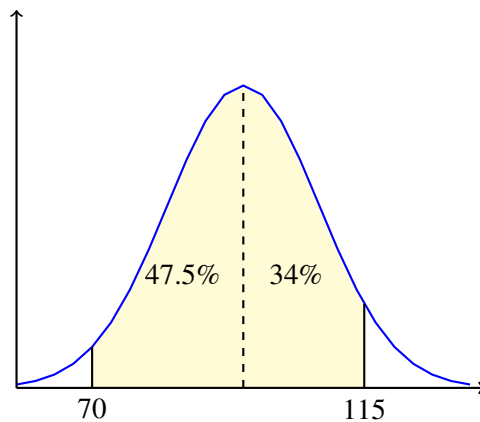
Solution Here, we want to find the area between the two given values. Notice that one of them is one standard deviation below the mean of 100 and the other is one standard deviation above the mean. This means that we have a symmetric interval that corresponds to one of the values named in our rule.



So, we have here that $P(85 \leq X \leq 115) = 68\%$. But, what if the interval is not symmetric?

Example 10.2.4 *IQ scores are Normally distributed with $X \sim N(100, 15)$. Find $P(70 \leq X \leq 115)$.*

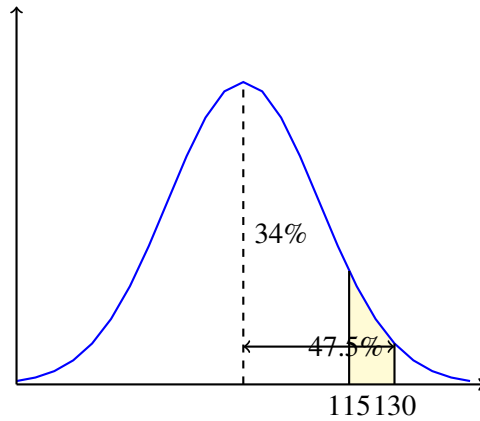
Here, we need to think about what is happening on either side of the mean line. It is not going to be the same area because on the left, we are two standard deviations away but on the right, we are only one standard deviation away.



So, when we add the areas in these regions, we get that $P(70 \leq X \leq 115) = 81.5\%$.

Example 10.2.5 *IQ scores are Normally distributed with $X \sim N(100, 15)$. Find $P(115 \leq X \leq 130)$.*

This one differs from the others because the whole region we are concerned with is on the same side of the mean line but does not include the tail. The way we want to approach a situation like this is to figure out how much area there is to the left of each of the lines, here one and two standard deviations above the mean, and then find their difference.



So, the difference here gives $P(115 \leq X \leq 130) = 13.5\%$.

10.3 Using z-Scores

How can we more accurately predict the likelihood of events? We can use something called a **z-score** to find the probability associated with a particular score. If we know how many standard deviations away from the mean a particular element is, we can make a prediction based on the pattern that exists with the rest of the observations. The z-score is the result we obtain when we convert the observations from our particular data set to a standardized scale. So, the z-score is how many standard deviations an observation lies from the center of the distribution. Based on this distance, we can give a probability, or likelihood, that the event could occur.

If X is a Normally distributed random variable with mean μ and standard deviation σ , denoted as

$$X \sim N(\mu, \sigma)$$

the z-score is calculated by the formula

$$z = \frac{x - \mu}{\sigma}$$

where x is the value of the observation.

Example 10.3.1 Find the z-score for $x = 15$ if $\mu = 12$ and $\sigma = 2.3$.

Solution Here, we have

$$z = \frac{15 - 12}{2.3} = \frac{3}{2.3} = 1.30$$

which means that in this situation, 15 is 1.3 standard deviations above the mean. We know it is above the mean because the z-score is positive.

Note: We always express z-scores with 2 decimal places because of the table we use to look up probabilities.

Example 10.3.2 Suppose the average number of home runs for a player on the Red Sox roster last season was 22 with a standard deviation of 4.2. Dustin Pedroia hit 15 home runs last season. How many standard deviations is his total from the mean?

Solution Using the same formula, we have

$$z = \frac{15 - 22}{4.2} = \frac{-7}{4.2} = -1.67$$

So, Pedroia's home run total was 1.67 standard deviations below the mean.

10.3.1 Probabilities with z -Scores

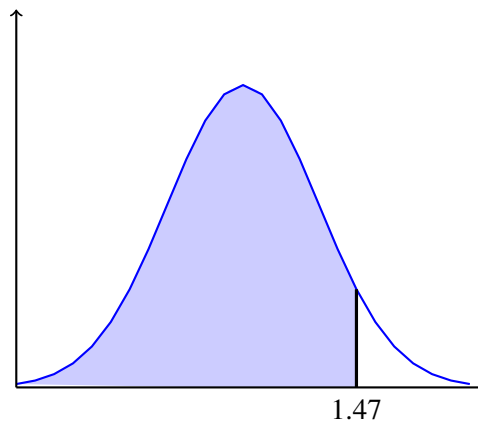
When we work with a number of standard deviations from the mean that is not exactly 1, 2, or 3 away, we use z -scores and the Normal table. The next examples will show the three possibilities that we could be presented with when trying to solve problems of this type.

Example 10.3.3 Suppose we are looking at IQ scores again $N(100, 15)$. Find the percent of people with IQ scores less than 122.

Solution To solve, we need to find z -scores and use the Normal probability table.

$$z = \frac{122 - 100}{15} = \frac{22}{15} = 1.47$$

We are looking for 'less than' so the visual is



Now we go to our table.

[Click here for Normal Distributions Table for \$z\$ -scores](#)

The table gives the probability that a randomly selected value lies to the *left* of the z -score. So here, since that is what we are looking for, we only need to look up the value in the table.

$$P(X \leq 122) = P(z \leq 1.47) = .9292$$

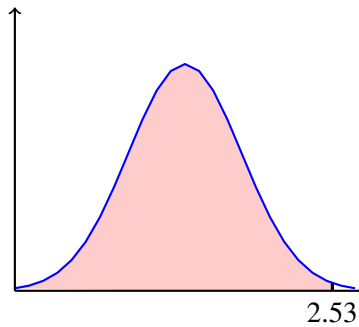
This probability is larger than 50%, as we would expect.

Example 10.3.4 Suppose we are looking at IQ scores again $N(100, 15)$. Find the percent of people with IQ scores greater than 138.

Solution For the IQ greater than 138, we start in a similar way. We want to find $P(X \geq 138)$ and since this is not exactly 1, 2 or 3 standard deviations from the mean, we need to use z -scores.

$$z = \frac{138 - 100}{15} = \frac{38}{15} = 2.53$$

Visually, we have



So when we look the value up in the table, we get $P = .9943$. This is not the answer we want, though, because the table gives the probability a random value is to the left of the boundary and we want to the right. So, we have to subtract the value we looked up from 1 (100% written in decimal form) to get the answer we need.

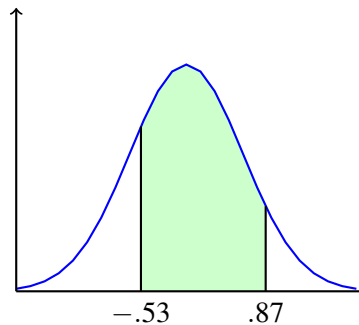
$$P(X \geq 138) = P(z \geq 2.53) = 1 - .9943 = .0057$$

Example 10.3.5 Suppose we are looking at IQ scores again $N(100, 15)$. Find the percent of people with IQ scores between 92 and 113.

Solution Here, we need to find the z -score associated with each of the values.

$$z_1 = \frac{92 - 100}{15} = \frac{-8}{15} = -.53$$
$$z_2 = \frac{113 - 100}{15} = \frac{13}{15} = .87$$

And now the visual.



So what we are looking for is the area between the two curves. How do we get there? Remember when we did this using the Empirical Rule - we subtracted the areas to the left of the respective boundaries and we will do the same here. The area to the left of $z = .87$ is .8078 and the area to the left of $z = -.53$ is .2981. So, we have

$$P(92 \leq X \leq 113) = P(-.53 \leq z \leq .87) = .8078 - .2981 = .5097$$

A note: when finding probabilities, **never subtract z-scores**. We want to find the difference in areas and not in the number of standard deviations. For example, in this last problem, if we would have subtracted z scores, we would be looking up $z = .87 - (-.53) = 1.40$ and that corresponds to a probability of .9192, nowhere near the correct value.

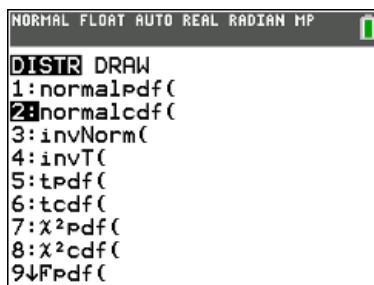
10.3.2 Probabilities Using Technology

As with much of what we have learned, we can find values for these probabilities on the TI-series calculator as well. We will illustrate how using the following example.

Example 10.3.6 *The Mathematics portion of the SAT exam is Normally distributed with a mean of 500 and a standard deviation of 100. What is the probability of getting a score of at least 650?*

Solution Since these are continuous curves, the probability of any one score is virtually 0, regardless of the score. We have to look at the probability that a score lies in some range. In this case, we are looking for scores between 650 and 800. But, the Normal distribution curve will not stop at 800; it continues on and considers all possible values up to infinity as the upper bound for the range. So we will have to ‘trick’ the calculator into using an infinite value as the upper bound so that we can get as accurate a probability as possible.

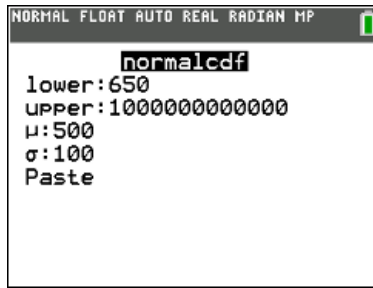
Pictorially, we are looking for the probability that the score in question lies in the following range :



We can find the z-score using the formula $z = \frac{x-\mu}{\sigma}$, where μ is the mean of the population, σ is the standard deviation of the population and x is the value of the observation. (Note: if we are finding the mean and standard deviation from the given data then we denote these values as \bar{x} and s . There is a slightly different formula when using empirical data, which we will get to shortly.) In this case, we have that

$$z = \frac{650 - 500}{100} = 1.5$$

So, we are looking for what percent of the observations lie more than 1.5 standards deviations above the mean. We could look this value up using a table, but we can obtain basically the same value (but a bit more accurate) from the calculator. The way in which we do this is by using the **normalcdf** command. The cdf stands for the cumulative distribution function. This function is a running total of the area in the given region. To use this command, we begin by getting into the menu by pressing **2nd** and **VARs** and then selecting the normalcdf command.



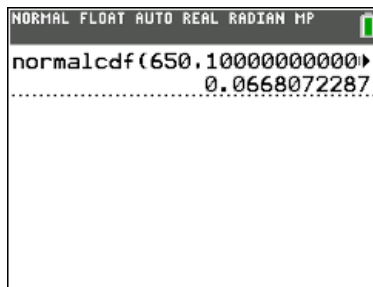
The format of the input for the command is as follows:

Normalcdf(lower bound, upper bound, mean, standard deviation)

In our case, we know the lower bound (650), the mean (500) and the standard deviation (100), but do not have an upper bound. This is where tricking the calculator comes into play. We need to use a ridiculously huge number for the upper bound to simulate infinity (if we had a similar situation but needed an infinite lower bound, we would make up some ridiculously huge negative number to represent negative infinity). So, for example, we could use



When we paste this in and then press **ENTER**, we get that the probability of getting a score greater than 650 on the Math portion of the SAT exam is approximately 6.68%.



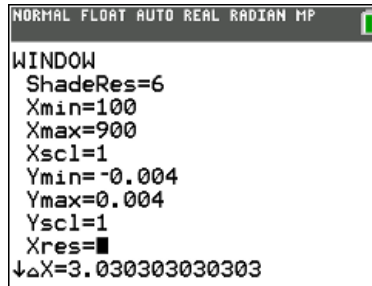
We can also use a different command to get a visual representation of the area in question, but we should get the same probability. This command is called **ShadeNorm**. In order to use this, however, we have to set the window so that the whole picture appears on the screen and also we have to make sure that the **STATPLOTS** are all Off. If they are not then we will get multiple pictures on the screen and it will possibly be confusing.

Once the **STATPLOTS** are all turned off, open the **WINDOW** menu. Use the following values to set the window (you can type in the formulas and the TI-83/84 will calculate the values for you).

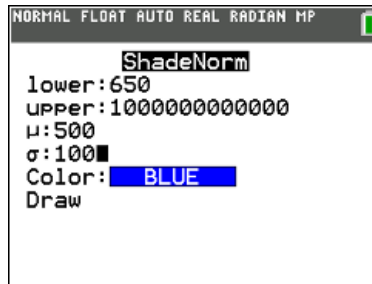
1. Set $X_{min} = \text{mean} - 4 * \text{standard deviation}$ and $X_{max} = \text{mean} + 4 * \text{standard deviation}$.

- Set the $Y_{min} = -.4/\text{standard deviation}$ and $Y_{max} = .4/\text{standard deviation}$.

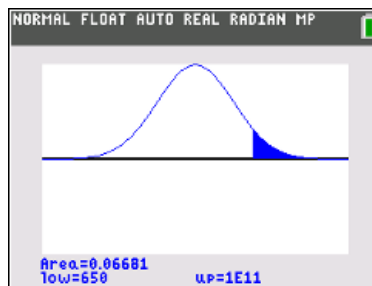
If this is done for this example, the following will be the values on the **Window** screen.



Now that the window is appropriately sized, press **2nd**, then **VARS** and then scroll to the *DRAW* menu. The first option is **ShadeNorm**. Open this command and then use the same format for the input as with the normalcdf command. For this example, we get



When we press the **ENTER** key, we get the picture we were looking for and also the associated probability of the shaded region.



Note: You cannot clear the picture out of memory simply by pressing the **CLEAR** key. If you do, this picture will be on the screen when you try to produce another plot. They will be on top of each other and make it quite confusing. In order to clear out the picture, press **2nd** and **PRGM** to get to the *DRAW* menu and then select **ClrDraw** to remove the picture.

Example 10.3.7 A car manufacturer found that a certain model was uncomfortable for women shorter than 159 cm. Women's heights are $N(161.5, 6.3)$. Find the percent of women for which the model is uncomfortable.

Solution Here, we want to find $P(X \leq 159)$. Since there is only one boundary point, there is only one z -score to calculate and look up.

$$z = \frac{159 - 161.5}{6.3} = -.40$$

So, we look up $z = -.40$ in the table and we see

$$P(X \leq 159) = P(z \leq -.40) = .3446$$

If we would have used the calculator instead, we would have had to use an arbitrarily large value for the lower endpoint, giving syntax like `normalcdf(-100000000,159,161.5,6.3)`

This returns .3457, which is reasonably close to our answer by hand.

Example 10.3.8 *A banker studying customer service needs finds that the number of times a person uses an ATM machine in a year are Normally distributed $N(30, 11.4)$. Find the percent of customers who use ATMs between 40 and 50 times per year.*

Solution Here, we are looking for the percent of customers falling between two endpoints, so we want $P(40 \leq X \leq 50)$. Since we have two endpoints, we have two z -scores to deal with.

$$z_1 = \frac{40 - 30}{11.4} = .88$$

$$z_2 = \frac{50 - 30}{11.4} = 1.75$$

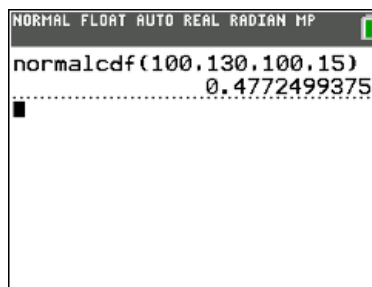
So, we need

$$\begin{aligned} P(40 \leq X \leq 50) &= \\ P(.88 \leq z \leq 1.75) &= \\ &= .9599 - .8106 \\ &= .1493 \end{aligned}$$

If we would have used the calculator, we would have gotten $P = .1505$, again reasonably close enough to our answer.

Example 10.3.9 *What is the probability that a randomly selected person has an IQ between 100 and 130 given that IQ scores are Normally distributed with a mean of $\mu = 100$ and a standard deviation of $\sigma = 15$?*

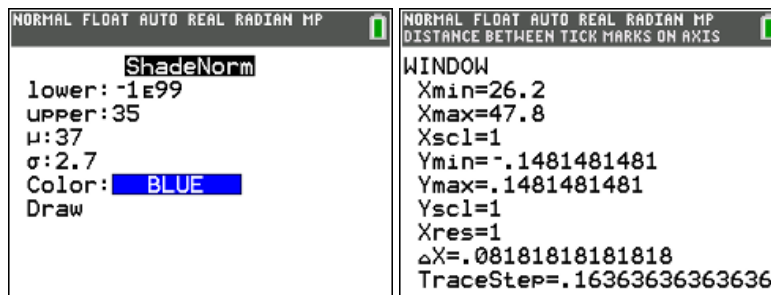
Solution We are dealing with the mean and standard deviation for the population in this example. We have a lower bound of 100 and an upper bound of 130. When we use the normalcdf command, we get the following:



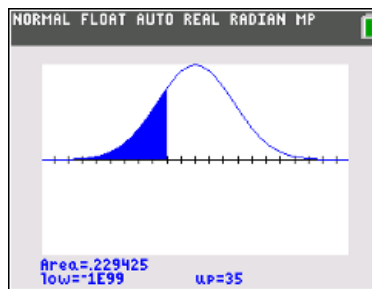
This shows that if we randomly select a person, their IQ will be between 100 and 130 47.72% of the time.

Example 10.3.10 *The average amount of time it takes for a smoker to quit is 37 days (this is a made-up value) with a standard deviation of 2.7 days. If the amount of time is Normally distributed, what percentage of smokers take less than 35 days to quit smoking?*

Solution Since we are talking about the area under a Normal distribution curve, the probability and the percent of people are represented the same way. So, just like before, we want to find the percent of the curve that is less than 35 days (somewhere between 0 and 1 standard deviation below the mean. We will find this using the ShadeNorm command. The upper limit is 35 days and the lower limit will be some ridiculously large negative value to represent negative infinity. When we use this command, the screen and the Window should look



The plot will look like



So, if we have a very large sample of smokers, approximately 22.94% of them will quit in less than 35 days.

10.3.3 What do the probabilities mean?

Before we continue, we want to make sure we understand what the probabilities we are calculating actually mean. It can be misleading to say that the probability of getting a 650 on the SAT Math portion is 6.68% or that the chance of the Red Sox winning the World Series is 99%. What do these things really mean? First off, they are entirely different they are both probabilities but they represent entirely different concepts. The latter is called a personal probability and is not based on any kind of experimentation or trial. There is no scientific proof that the Red Sox will win it is just an opinion. The former is based on empirical data. But it does not mean that out of every 100 people that 6.68 of them will score at least 650. Probabilities are based on a long series of trials many, many trials. The probability we calculated in the last example means that in a long series of trials, a randomly selected person will have a score of at least 650 on the Math portion of the SAT exam 6.68% of the time.

10.4 The Central Limit Theorem

Depending on what information we have, we need to use slightly different formulas to find the probabilities of events occurring. If we are given the mean and standard deviation for a Normal population (since we cannot calculate them) then we say that the population is Normally distributed with a mean of μ and a standard deviation of σ . When dealing with a population, we are dealing with the probability that one observation lies in a certain range. But when we are using a sample, we have a smaller standard deviation because averages vary less than individual values. The reason we can use this to help us find probabilities is because of the Central Limit Theorem. This is an integral theorem in the study of probability. We will not prove it here but we will state the theorem.

Theorem 10.4.1 *The Central Limit Theorem states that if we draw a simple random sample of size n from a population with mean μ and standard deviation σ then the sampling distribution of the mean \bar{x} is approximately Normal when n is large (at least 30). We express this symbolically as*

$$\bar{x} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Here, a simple random sample of size n is a collection of n elements chosen from the population in such a way that every single element in the population has an equal chance of being selected.

So what does this do for us in practice? Notice that the standard deviation is smaller since we are dividing by the square root of the size of the sample. The standard deviation is smaller since averages have a smaller variation than an individual value. In these situations, we are looking at the probability that the average value for the simple random sample falls within a prescribed range whereas we earlier were looking at the probability that an individual value fell within a given range.

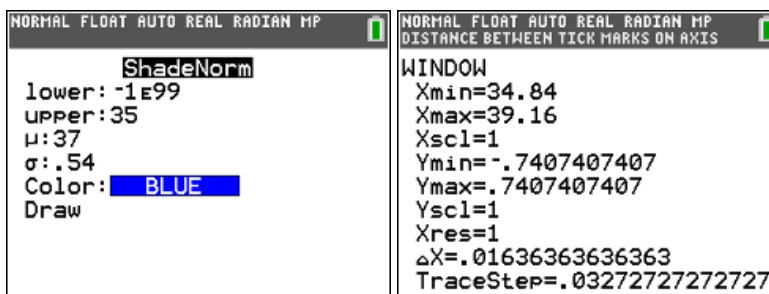
Another important note: we can use Normal probability calculations to answer questions about the sample means from many observations even when the population distribution is not Normal. Nowhere in the statement of the theorem does it say we can only use Normal distributions, and the reality is that as the number of samples gets large, the distribution becomes more and more Normal.

Example 10.4.2 *If we poll 25 people who recently quit smoking, what is the probability that their average quit time was less than 35 days?*

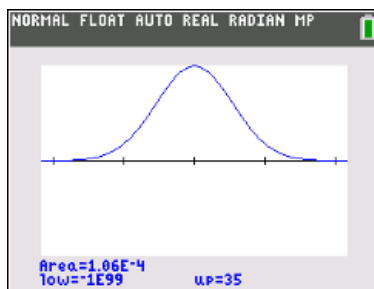
Solution Using the information from the previous example combined with the Central Limit Theorem we have that

$$\bar{x} \sim N\left(37, \frac{2.7}{\sqrt{25}}\right) \sim N(37, .54)$$

Using the same approach as the last example we get the following screens



The shaded area will be different because of the significantly smaller standard deviation.



Here, the probability that the average quit time for the 25 smokers is not very likely at all to be less than 35 days. In fact, the probability is .0106.

Clearly, having a simple random sample makes a large difference when finding the probability. The larger the simple random sample, the smaller the standard deviation and consequently, the smaller the chances of being far from the mean.

Before we go on, there are a few conditions as to when we can use the Central Limit Theorem.

1. We must have the sample taken randomly from the population, that there are no biases and that the sample represents the population.
2. The sample size, n , must be no bigger than 10% of the size of the population.
3. If there are no strong outliers or strong skewness, 30 is a minimum sample size and if there are strong outliers or skewness, 55 is the magic number.

Example 10.4.3 If SAT scores are $N(1026, 209)$, what is the probability that a randomly selected student scores greater than 1110?

Solution Here, we are looking for $P(x \geq 1110)$. We first find the z -score and look up the associated probability, but we have to remember that because we are looking for the probability that the student score is greater than 1110, we need to subtract the probability we look up from 1.

$$z = \frac{1110 - 1026}{209} = .40$$

$$P(x \geq 1110) = P(z \geq .40) = 1 - .6554 = .3446$$

Using technology, we have

$$P(x \geq 1110) = .3439$$

Example 10.4.4 Now, what is the probability that if we averaged the scores for 80 students, their average would be greater than 1110?

Here, we need to apply the Central Limit Theorem. The mean remains unchanged but the standard error is much smaller.

$$SE = \frac{209}{\sqrt{80}} = 23.37$$

Now we find the z -score

$$z = \frac{1110 - 1026}{23.37} = 3.59$$

And then, as before, we look up the probability and subtract from 1 because of the direction of the inequality.

$$\begin{aligned} P(\bar{x} \geq 1110) &= P(z \geq 3.59) \\ &= 1 - .9999 \\ &= .0001 \end{aligned}$$

Technology gives $P(\bar{x} \geq 1110) = P(z \geq 3.59) = .00016$, a more exact answer.

Example 10.4.5 Suppose that the length of radio edits of songs is 300 seconds with a standard deviation of 30 seconds. Suppose you have 10 song playlist for workouts. What is the probability that the mean length of those 10 songs is between 310 and 330 seconds?

Solution Here, we have 10 songs who's mean we are concerned with, which tells us that we need to apply the Central Limit Theorem.

$$\bar{x} \sim N\left(300, \frac{30}{\sqrt{10}}\right)$$

We need to calculate two z -scores here since we are looking for the probability between two given scores.

$$\begin{aligned} z_1 &= \frac{310 - 300}{\frac{30}{\sqrt{10}}} = 1.05 \\ z_2 &= \frac{330 - 300}{\frac{30}{\sqrt{10}}} = 3.15 \end{aligned}$$

We look these up in our table and subtract to get

$$P(310 \leq \bar{x} \leq 330) = P(1.05 \leq z \leq 3.15) = .9992 - .8531 = .1461$$

If we would have used technology, we would have gotten a similar probability.

$$P(310 \leq \bar{x} \leq 330) = .1451$$

Example 10.4.6 There is only a 10% percent chance that the average song length for the playlist is longer than what length?

Solution When solving a problem like this, we have to think backwards. We know the distribution from the last example but instead of knowing the song length, we know the percent. We need the z -score, which we can find by searching the body of the table. When we do, we see that the top 10%, which is the same as the bottom 90%, corresponds to $z = 1.28$. We use this in our z -score formula to solve for \bar{x} .

$$\begin{aligned} 1.28 &= \frac{\bar{x} - 300}{\frac{30}{\sqrt{10}}} \\ \bar{x} &= 312.14 \end{aligned}$$

Example 10.4.7 A soda bottler says their cans hold 12 oz. of soda. The amount the bottling machine puts in is 12.1 oz with a standard deviation of .12 oz. What fraction of all cans sold are underweight?

Solution Here, we need to find $P(x < 12)$. First, we need to find the z -score.

$$z = \frac{12 - 12.1}{.12} = -.83$$

Since we want to know the probability we are less than 12 ounces, we only need to look up this score in the table.

$$P(x < 12) = P(z < -.83) = .2033$$

If we used technology here, we would have gotten a similar probability.

$$P(x < 12) = .2023$$

Example 10.4.8 A soda bottler says their cans hold 12 oz. of soda. The amount the bottling machine puts in is 12.1 oz with a standard deviation of .12 oz. What is the probability that the mean weight of 6 cans is below the stated amount?

Solution Here, we are looking for $P(\bar{x} < 12)$ and need to apply the Central Limit Theorem with $n = 6$. Following the same process as before, we have

$$z = \frac{12 - 12.1}{\frac{.12}{\sqrt{6}}} = -2.04$$
$$P(\bar{x} < 12) = P(z < -2.04) = .0207$$

And here, technology gives $P(\bar{x} < 12) = .0206$, which is very close to our answer.

Example 10.4.9 A soda bottler says their cans hold 12 oz. of soda. The amount the bottling machine puts in is 12.1 oz with a standard deviation of .12 oz. What is the probability that the mean weight for a case (24 cans) is below where it should be?

Solution This problem is similar to the last, with the difference being that $n = 24$. Because n is much larger, we get a much smaller standard error

$$SE = \frac{.12}{\sqrt{24}} = .02449$$

and consequently, a much larger z -score.

$$z = \frac{12 - 12.1}{\frac{.12}{\sqrt{24}}} = \frac{12 - 12.1}{.02449} = -4.08$$

A z -score this large is off the table, giving us a very small probability.

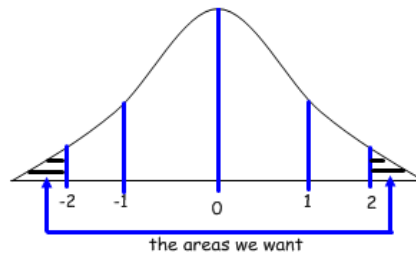
$$P(\bar{x} < 12) = P(z < -4.08) = .0001$$

10.4.1 The Standard Normal Distribution

A **standard Normal distribution** has a mean of 0 and a standard deviation of 1. Other than this, we calculate probabilities in the same manner as with other Normal distributions.

Example 10.4.10 Find the percent of elements that would be more than 2 standard deviations from the mean in a standard Normal population.

Solution Since we are not told whether we are looking for elements greater than the mean or smaller than the mean, we need to find both. We will use the normalcdf function to calculate these probabilities. What we want is the area of the tails taken together.



For the lower tail we are looking for the area between negative infinity and -2. For the upper tail we are looking for the area between 2 and positive infinity. The calculator gives the following for areas :

```
NORMAL FLOAT AUTO REAL RADIAN MP
normalcdf(-1E99,-2,0,1)
.....
.022750062
normalcdf(2,1E99,0,1)
.....
.022750062
```

Notice that the area of both tails is the same. This is because the Normal distribution curve is symmetric. So, we could have found the area of just one tail and then doubled it and we would have gotten the same result: the percent of elements in the tails of the standard Normal curve that are at least 2 standard deviations from the mean is approximately 4.55%. This symmetry property is not limited to the standard Normal population, however. As long as the distribution curve is Normal, the symmetry property holds.

10.5 Exercises

1. Suppose that the mean commute time to Salem State University is 24 minutes with a standard deviation of 6 minutes.
 - (a) What percentage of students have a mean commute time of more than 33 minutes?
 - (b) If we polled 20 students, what is the probability that their mean commute time is more than 33 minutes?
2. A largely Normally distributed population has a mean of 4.50 and a standard deviation of 1.05.
 - (a) Find the probability that a randomly selected score is less than 5.00.
 - (b) If a sample of size $n = 40$ is randomly selected, find the probability that the sample mean \bar{x} is less than 5.00.
3. A study of the time high school students spend working each week at a job found that the mean is 10.7 hours and the standard deviation is 11.2 hours. If 42 high school students are randomly selected, find the probability that their mean weekly work time is less than 12 hours.
4. The average time it takes to drive around Boston to find a parking space, seeing how I refuse to use garages, is 14 minutes with a standard deviation of 8.5 minutes.
 - (a) What is the probability that I find a parking spot in under 7 minutes?
 - (b) What is the probability that in 8 trips to the city, the mean time for finding a parking spot for those trips is less than 7 minutes?
5. Suppose that the amount of time it takes for the average statistics student to finish one of the exams is 115 minutes with a standard deviation of 28 minutes.
 - (a) What is the probability that a class of 15 students finish with an average time of less than 100 minutes?
 - (b) What is the probability that a class of 30 students finish with an average time of less than 100 minutes?
6. Suppose I tracked the length of my tee shots for an entire golf season and found that my average drive was 245 yards with a standard deviation of 37 yards. Suppose further that the drives are Normally distributed.
 - (a) What is the probability that a randomly selected drive exceeded 280 yards?
 - (b) What is the probability that my average drive exceeded 280 yards for an entire round of 18 holes?
7. Stephanie has frequent conference calls and over the course of time, the calls take an average of an hour with a standard deviation of 12 minutes. If she has 3 conference calls this week, what is the probability that their average time will be less than 45 minutes?
8. The Providence Grays play doubleheaders on almost every single date. Suppose that the time of games is Normally distributed with a mean of 156 minutes and a standard deviation of 22 minutes. What percent of games will be completed in less than 2 hours?

9. Suppose that the length of time for the Red Sox-Yankees playoff games are as follows (in minutes):

$\{185, 221, 201, 193, 212, 231, 229\}$

If we can assume that the mean and standard deviation for these times is representative of the mean time for any Red Sox-Yankees game, what is the probability that a randomly selected game is completed in under 3 hours?

10. Use the Empirical rule to approximate the percent of students who score less than 65 on an exam that has a mean score of 85 with a standard deviation of 10, given that the exam scores are Normally distributed.
11. Suppose the mean commute distance of SSU students is 12 miles with a standard deviation of 3.3 miles. Approximate the percent of students whose commute is between 5.4 miles and 15.3 miles using the Empirical rule if we can assume we have a Normal distribution.
12. The mean amount of time spent on homework is Normally distributed with a mean of 17.2 hours with a standard deviation of 4.8 hours. Using the Empirical rule, how many hours of study give the range for the middle 95% of students?

10.6 Solutions

1. (a)

$$z = \frac{33 - 24}{6} = 1.50$$

$$P(X > 33) = P(z > 1.50) = 1 - .9332 = .0668$$

(b)

$$z = \frac{33 - 24}{\frac{6}{\sqrt{20}}} = 6.71$$

$$P(\bar{x} > 33) = P(z > 6.71) = 1 - .9999 = .0001$$

2. (a)

$$z = \frac{5 - 4.5}{1.05} = .48$$

$$P(X < 5) = P(z < .48) = .6844$$

(b)

$$z = \frac{5 - 4.5}{\frac{1.05}{\sqrt{40}}} = 3.01$$

$$P(\bar{x} < 5) = P(z < 3.01) = .9987$$

3.

$$z = \frac{12 - 10.7}{\frac{11.2}{\sqrt{42}}} = .75$$

$$P(\bar{x} < 12) = P(z < .75) = .7734$$

4. (a)

$$z = \frac{7 - 14}{8.5} = -.82$$

$$P(X < 7) = P(z < -.82) = .2061$$

(b)

$$z = \frac{7 - 14}{\frac{8.5}{\sqrt{8}}} = -2.33$$

$$P(\bar{x} < 7) = P(z < -2.33) = .0099$$

5. (a)

$$z = \frac{100 - 115}{\frac{28}{\sqrt{15}}} = -2.07$$

$$P(\bar{x} < 100) = P(z < -2.07) = .0192$$

- (b)
(c)

$$z = \frac{100 - 115}{\frac{28}{\sqrt{30}}} = -2.93$$
$$P(\bar{x} < 100) = P(z < -2.93) = .0017$$

6. (a)

$$z = \frac{280 - 245}{37} = .95$$
$$P(X > 280) = P(z > .95) = 1 - .8289 = .1711$$

- (b)

$$z = \frac{280 - 245}{\frac{37}{\sqrt{18}}} = 4.01$$
$$P(\bar{x} > 280) = P(z > 4.01) = 1 - .9999 = .0001$$

- 7.

$$z = \frac{45 - 60}{\frac{12}{\sqrt{3}}} = -2.17$$
$$P(\bar{x} < 45) = P(z < -2.17) = .0150$$

- 8.

$$z = \frac{120 - 156}{22} = -1.64$$
$$P(X < 120) = P(z < -1.64) = .0505$$

9. First, we need to find the mean and standard deviation for this data set.

$$\bar{x} = 210.29$$
$$s = 17.90$$

Now we can find the z -score and associated probability.

$$z = \frac{180 - 210.29}{17.9} = -1.69$$
$$P(X < 180) = P(z < -1.69) = .0455$$

10. We are two standard deviations below the mean, so we need the area that is outside of the 95% range. This lower tail represents 2.5%.
11. 15.3 is one standard deviation above the mean, which gives 34% between the mean and this point. 5.4 is two standard deviations below the mean, which gives 47.5% between the mean and this point. Combining, we have that the percent of students is $34\% + 47.5\% = 81.5\%$.

12. The middle 95% of students are those within 2 standard deviations of the mean.

$$17.2 + 2(4.8) = 26.8$$

$$17.2 - 2(4.8) = 7.6$$

The middle 95% of students study between 7.6 hours and 26.8 hours.

Chapter 11

Rates of Change and Derivatives

11.1 Average Rates of Change

Earlier in the course, we talked about how to find the slope of a linear function. We used the formula

$$m = \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$

for the slope between two points on the line. And we knew that the slope was always going to be the same, regardless of which two points we chose. But what about for other curves? We will see later in this chapter that there are applications that rely on our ability to find rates of change for functions where the rate changes. Let's investigate first using an example.

Example 11.1.1 *Suppose a grapefruit is thrown straight up in the air at $t = 0$ seconds. The grapefruit leaves the thrower's hand at high speed, it slows down until it reaches its maximum height and then turns and picks up speed until it hits the ground.*

t (sec)	0	1	2	3	4	5	6
y (feet)	6	90	142	162	150	106	30

Why does the grapefruit travel faster from $0 \leq t \leq 1$ than from $1 \leq t \leq 2$? This is because of gravity. The longer the grapefruit is in the air on the way to the maximum height, the slower the rate of speed (the rate of change of distance compared to time) will be.

Suppose we wanted to find the average rate of change on the interval $0 \leq t \leq 1$? We would need to find the change in position at those two times and divide that by the change in time.

$$\frac{y(1) - y(0)}{1 - 0} = \frac{90 - 6}{1} = 84 \frac{\text{ft}}{\text{sec}}$$

If we look at the average rate of change on the interval $1 \leq t \leq 2$, we expect the rate to be smaller, and in fact it is.

$$\frac{y(2) - y(1)}{2 - 1} = \frac{142 - 90}{1} = 52 \frac{\text{ft}}{\text{sec}}$$

What about the average velocity from $0 \leq t \leq 2$? It would stand to reason that the rate here should be between the other two, since it is considering the whole interval instead of the first second or the second second alone.

$$\frac{84 + 52}{2} = 68 \frac{\text{ft}}{\text{sec}}$$

Now, where does the grapefruit reach its maximum height? It would be nice if we could answer this question, but here we cannot because we are not given a function - we are only given some data points. It is possible that the maximum height is 162 feet at $t = 3$ seconds. But we cannot know for sure. Maybe the grapefruit is still climbing and reaches the maximum between $t = 3$ and $t = 4$ seconds. Maybe it reaches the maximum between $t = 2$ and $t = 3$ seconds and it is hitting the height of 162 feet on the way back down. The best we could guess is that the maximum height occurs between $t = 2$ and $t = 4$ seconds.

Example 11.1.2 *A baseball is thrown from center field to the catcher. The distance traveled of the ball at time t is given below.*

time (in seconds)	0	.5	1	1.5	2	2.5
distance (in ft)	0	65	120	170	215	255

Find the average velocity over each one second interval. Explain why the sequences of averages makes sense in the context of the problem.

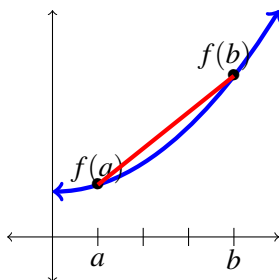
Solution Remember, velocity equals the change in distance divided by change in time. Here, we have

$$\begin{aligned}\frac{120 - 0}{1 - 0} &= 120 \frac{\text{ft}}{\text{sec}} \\ \frac{170 - 65}{1.5 - .5} &= 105 \frac{\text{ft}}{\text{sec}} \\ \frac{215 - 120}{2 - 1} &= 95 \frac{\text{ft}}{\text{sec}} \\ \frac{255 - 170}{2.5 - 1.5} &= 85 \frac{\text{ft}}{\text{sec}}\end{aligned}$$

We can see that the ball is slowing down as it gets closer to the catcher. This is because it is being thrown towards the catcher at an angle but not straight in the air. As the ball travels, because the direction of the motion is not directly with gravity or against it, the ball is being slowed by both gravity and air resistance. The ball is not coming to a stop while in flight like when we throw the ball straight up in the air but it is slowing down.

In general, what we are doing is finding the *slope of the secant line* between the given endpoints. The general formula in terms of position is

$$\frac{\Delta y}{\Delta x} = \frac{f(b) - f(a)}{b - a}$$



Notice that this formula looks a lot like that for the slope of a line - that is because *it is* the formula for the slope of a line. It is just expressed using function values in terms of $f(x)$ instead of in terms of y . It will work essentially the same, however, as we will see in the next example.

Example 11.1.3 Find the average rate of change of $f(x) = 3x - 1$ over $-1 \leq x \leq 3$.

Solution What we are being asked to find here is the slope of the secant line - that is, the average rate of change between the endpoints. We are approximating the slope of the curve by using two points that were chosen, here $x = -1$ and $x = 3$.

$$\begin{aligned}\text{a.r.o.c.} &= \frac{f(3) - f(-1)}{3 - (-1)} \\ &= \frac{8 - (-4)}{4} \\ &= \frac{12}{4} \\ &= 3\end{aligned}$$

Example 11.1.4 Find the average rate of change of $f(x) = x^2$ over $2 \leq x \leq 5$.

Solution We use a similar formula to the last here.

$$\begin{aligned}\text{a.r.o.c.} &= \frac{f(5) - f(2)}{5 - 2} \\ &= \frac{25 - 4}{3} \\ &= 7\end{aligned}$$

In general, what we want to be able to do is approximate the *instantaneous* rate of change, or rate of change at one point. You may be asking yourself how it is possible to find this, considering that we have been talking about finding the average rate of change over an interval. We can do so using the **difference quotient**, which is given by

Formula 11.1.5 The Difference Quotient

$$\frac{\Delta y}{\Delta x} = \frac{f(x+h) - f(x)}{h}$$

The idea here is that we want to consider the average rate of change from some point x to another point h units away. Then, we suppose that we can let h approach 0, since we cannot let $h = 0$ or we get an undefined expression. Let's see how this works in an example.

Example 11.1.6 Find the average rate of change of $f(x) = x^2 + x$ over $[1, 1+h]$.

Solution We will use the difference quotient here.

$$\begin{aligned}\text{a.r.o.c.} &= \frac{f(1+h) - f(1)}{1+h-1} \\ &= \frac{(1+h)^2 + (1+h) - (1^2 + 1)}{h} \\ &= \frac{1 + 2h + h^2 + 1 + h - 1 - 1}{h} \\ &= \frac{3h + h^2}{h} \\ &= \frac{h(3+h)}{h} \\ &= 3 + h\end{aligned}$$

What this tells us is that for this quadratic function when we are around $x = 1$, the rate of change is roughly 3 units more than the distance away from $x = 1$ that we are. This is only valid, however, when we are close to $x = 1$. We cannot say that we get a slope of approximately 3 when we are at $x = 7$, for example. At that

point, we would be looking for the average rate of change on $[7, 7 + h]$, which would be

$$\begin{aligned}\text{a.r.o.c.} &= \frac{f(7+h) - f(7)}{7+h-7} \\ &= \frac{(7+h)^2 + (7+h) - (7^2 + 7)}{h} \\ &= \frac{49 + 14h + h^2 + 7 + h - 49 - 7}{h} \\ &= \frac{15h + h^2}{h} \\ &= \frac{h(15+h)}{h} \\ &= 15 + h\end{aligned}$$

So, for x values near 7, the slope of this quadratic function would be approximately 15.

11.1.1 Exercises

1. You drop a rock off a cliff into the water, which is 200 feet below. The height of the rock, at time t , is given in the table below.

t in seconds	0	.5	1	1.5	2	2.5	3	3.5
height in ft	200	196	184	164	136	100	56	4

- (a) Find the average velocity over the 3.5 second interval.
- (b) Find the average velocity from $t = 1$ to $t = 3$.
- (c) Do you expect the average velocity over the interval from $t = 0$ to $t = 2$ to be larger or smaller than that from (b)? Explain your thoughts and then find this average velocity to confirm your idea.
- (d) Why is the velocity negative here?
2. Find the average rate of change of $f(x) = 2x - 5$ on the interval $-2 \leq x \leq 4$.
3. Find the average rate of change of $f(x) = 2x^2 - 3x$ over the interval
- (a) $1 \leq x \leq 4$
- (b) $1 \leq x \leq 3$
- (c) $1 \leq x \leq 2$
4. Find the average rate of change of $f(x) = x^2 - 2x + 1$ on the interval
- (a) $2 \leq x \leq 3$
- (b) $-1 \leq x \leq 3$
- (c) $0 \leq x \leq 2$
5. Find the average rate of change of $f(x) = x^2 - 2x + 1$ on the interval $3 \leq x \leq 3 + h$.

11.1.2 Solutions

1. (a) $\frac{4-200}{3.5-0} = \frac{-196}{3.5} = -56 \frac{\text{ft}}{\text{sec}}$
- (b) $\frac{56-184}{3-1} = \frac{-128}{2} = -64 \frac{\text{ft}}{\text{sec}}$
- (c) The velocity will be greater in magnitude from 1 to 3 seconds. The initial velocity is $0 \frac{\text{ft}}{\text{sec}}$ and due to freefall acceleration, the magnitude of the velocity increases as time increases until we reach the ground.

$$\frac{136-200}{2-0} = \frac{-64}{2} = -32 \frac{\text{ft}}{\text{sec}}$$

- (d) Velocity is negative because we are finding the velocity of a falling object. Velocity is considered positive as the object's height gets higher.

$$2. \frac{f(4)-f(2)}{4-(-2)} = \frac{3-(-9)}{4-(-2)} = \frac{12}{6} = 2$$

$$3. (a) \frac{f(4)-f(1)}{4-1} = \frac{32-12-2+3}{3} = 7$$

$$(b) \frac{f(3)-f(1)}{3-1} = \frac{18-9-2+3}{2} = 5$$

$$(c) \frac{f(2)-f(1)}{2-1} = \frac{8-6-2+3}{1} = 3$$

$$4. (a) \frac{f(3)-f(2)}{3-2} = \frac{9-6+1-(4-4+1)}{1} = 3$$

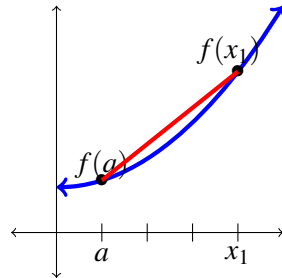
$$(b) \frac{f(3)-f(-1)}{3-(-1)} = \frac{9-6+1-(1+2+1)}{4} = 0$$

$$(c) \frac{f(2)-f(0)}{2-0} = \frac{4-4+1-1}{2} = 0$$

$$5. \frac{f(3+h)-f(3)}{3+h-3} = \frac{(3+h)^2-2(3+h)+1-(3^2-2(3)+1)}{h} = \frac{9+6h+h^2-6-2h+1-9+6-1}{h} = \frac{4h+h^2}{h} = 4+h$$

11.2 Instantaneous Rates of Change

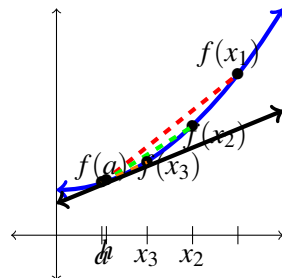
We have been building towards this idea - how can we find the rate of change at one point instead of over an interval? Here is the idea of this in pictures. Suppose we want to approximate the slope at the point $x = a$. We can find the slope of the secant line like we did before using some other point $x = x_1$ as the other end point of the line segment.



We want to see how the slope of successive secant lines change as the second point gets closer and closer to the $x = a$.



Eventually, we get arbitrarily close to $x = a$, say at $x = a + h$, for the second point.



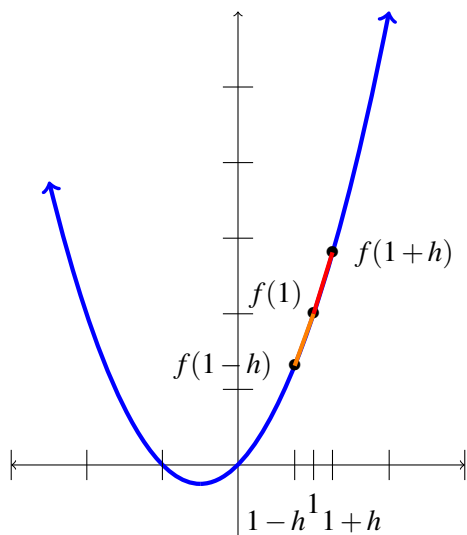
Because we cannot tell the difference between $x = a$ and $x = a + h$ when h is extremely small, we are looking at the slope of the tangent line at $x = a$ - the line that only touches the curve at the point $x = a$ - rather than the slope of the secant line.

Let's investigate how this works algebraically by looking back at an earlier example.

Example 11.2.1 Find the instantaneous rate of change of $f(x) = x^2 + x$ at $x = 1$.

Solution If we were looking at a linear function, we could consider any interval around the point in question and find the instantaneous rate of change since that is the defining characteristic of a linear equation - we get

the same slope between any two points. But since we have a quadratic function, we will not get the same slope when we look at intervals on either side of the point $x = 1$. So, we will consider values $x = 1 + h$ and $x = 1 - h$ for various small values of h . Visually, we are looking at ...



When we look at values for h , we need ones close to 0 so that we are looking only at the interval around $x = 1$. We will plug values into the difference quotient and use the values we see to approximate the instantaneous rate of change.

h	.1	.01	.001	-.1	-.01	-.001
$\frac{(1+h)^2+(1+h)-(1^2+1)}{h}$	3.1	3.01	3.001	2.9	2.99	2.999

So, our guess for the instantaneous rate of change of $f(x) = x^2 + x$ at the point $x = 1$ is that the rate of change is 3.

Example 11.2.2 Find the instantaneous rate of change of the function $f(x) = 3 - 2x^2$ at the point where $x = 4$.

Solution We will again use the difference quotient and small values for h on both sides of $x = 4$.

h	.1	.01	.001	-.1	-.01	-.001
$\frac{3-2(4+h)^2-(3-4^2)}{h}$	-17.62	-16.16	-16.0016	-14.42	-15.84	-15.9840

It appears that the instantaneous rate of change at $x = 4$ is -16 .

Rather than make tables like this every time, we want to see what happens to the difference quotient at a given point by supposing that the value of h approaches 0. We will look into this soon.

Example 11.2.3 Using the difference quotient, estimate the instantaneous rate of change of the function $f(x) = 3x^2 + 4x - 2$ at the point where $x = 3$.

Solution We approach as we did in the last section. But we will introduce new notation here too - when we are finding the instantaneous rate of change, one way we can denote this is using $f'(x)$. We will talk of this more later and formally define it then.

$$\begin{aligned}
 f'(3) &\approx \frac{3(3+h)^2 + 4(3+h) - 2 - (3(3)^2 + 4(3) - 2)}{h} \\
 &= \frac{27 + 18h + 3h^2 + 12 + 4h - 2 - 27 - 12 + 2}{h} \\
 &= \frac{22h + 3h^2}{h} \\
 &= \frac{h(22 + 3h)}{h} \\
 &= 22 + 3h
 \end{aligned}$$

So, if we have values of h that are closer and closer to 0, we would have an instantaneous rate of change of 22. That is, $f'(3) = 22$.

Example 11.2.4 Find the slope of $f(x) = x^3 - 2$ at the point where $x = 2$.

Solution This will work the same way, just a bit more algebra. But remember, the instantaneous rate of change at a point is the same as the slope at that point, which is same as the slope of the tangent line to the curve at the point in question.

$$\begin{aligned}
 f'(x) &\approx \frac{f(x+h) - f(x)}{h} \\
 f'(2) &\approx \frac{f(2+h) - f(2)}{h} \\
 &= \frac{(2+h)^3 - 2 - (2^3 - 2)}{h} \\
 &= \frac{8 + 12h + 6h^2 + h^3 - 2 - 8 + 2}{h} \\
 &= \frac{12h + 6h^2 + h^3}{h} \\
 &= 12 + 6h + h^2
 \end{aligned}$$

What happens to this as $h \rightarrow 0$? We get a slope of 12 at the point $(2, 6)$.

11.2.1 Exercises

1. Using values for h that get successively smaller, find the instantaneous rate of change of the function $f(x) = -x^2 + 3x$ at the point where $x = -3$.
2. Using values for h that get successively smaller, find the instantaneous rate of change of the function $f(x) = 5x - 3x^2$ at the point where $x = 2$.
3. Using the difference quotient and algebra, find the average rate of change of $f(x) = (x + 2)^2$ on the interval $[4, 4 + h]$ and then use the simplified result to find the instantaneous rate of change at $x = 4$.
4. Using the difference quotient and algebra, find the average rate of change of $f(x) = 2x - 4x^2$ on the interval $[2, 2 + h]$ and then use the simplified result to find the instantaneous rate of change at $x = 2$.
5. Using the difference quotient and algebra, find the average rate of change of $f(x) = 2x^2 - 5$ on the interval $[-3, -3 + h]$ and then use the simplified result to find the instantaneous rate of change at $x = -3$.

11.2.2 Solutions

1.

$$\begin{aligned}\frac{f(-3+h) - f(-3)}{h} &= \frac{-(-3+h)^2 + 3(-3+h) - (-(3)^3 + 3(-3))}{h} \\ &= \frac{-9 + 6h - h^2 - 9 + 3h + 9 + 9}{h} \\ &= \frac{9h - h^2}{h} \\ &= 9 - h\end{aligned}$$

h	$9 - h$	h	$9 - h$
.1	8.9	-.1	9.1
.01	8.99	-.01	9.01
.001	8.999	-.001	9.001

It seems that as $h \rightarrow 0$, the rate of change approaches 9 when $x = -3$.

2.

$$\begin{aligned}\frac{f(2+h) - f(2)}{h} &= \frac{5(2+h) - 3(2+h)^2 - (5(2) - 3(2)^2)}{h} \\ &= \frac{10 + 5h - 12 - 12h - 3h^2 - 10 + 12}{h} \\ &= \frac{-7h - 3h^2}{h} \\ &= -7 - 3h\end{aligned}$$

h	$-7 - 3h$	h	$-7 - 3h$
.1	-7.3	-.1	-6.7
.01	-7.03	-.01	-6.97
.001	-7.003	-.001	-6.997

3.

$$\begin{aligned}\frac{f(4+h) - f(4)}{h} &= \frac{(4+h+2)^2 - (4+2)^2}{h} \\ &= \frac{36 + 12h + h^2 - 36}{h} \\ &= \frac{12h + h^2}{h} \\ &= 12 + h\end{aligned}$$

As $h \rightarrow 0$, $12 + h \rightarrow 12$. Therefore, the instantaneous rate of change of $(x+2)^2$ when $x = 4$ is 12.

4.

$$\begin{aligned}\frac{f(2+h) - f(2)}{h} &= \frac{2(2+h) - 4(2+h)^2 - (2(2) + 4(2)^2)}{h} \\ &= \frac{4 + 2h - 16 - 16h - 4h^2 - 4 - 16}{h} \\ &= \frac{-14h - 4h^2}{h} \\ &= -14 - 4h\end{aligned}$$

As $h \rightarrow 0$, $-14 - 4h \rightarrow -14$. Therefore, the instantaneous rate of change of $2x - 4x^2$ when $x = 2$ is -14 .

5.

$$\begin{aligned}\frac{f(-3+h) - f(-3)}{h} &= \frac{2(-3+h)^2 - 5 - (2(-3)^2 - 5)}{h} \\ &= \frac{18 - 12h + 2h^2 - 18 + 5}{h} \\ &= \frac{-12h + 2h^2}{h} \\ &= -12 + 2h\end{aligned}$$

As $h \rightarrow 0$, $-12 + 2h \rightarrow -12$. Therefore, the instantaneous rate of change of $2x^2 - 5$ when $x = -3$ is -12 .

11.3 Derivatives

We have been finding derivatives for a couple of sections now but we did not define them as such until now.

Definition 11.3.1 The derivative of the function $f(x)$ with respect to the variable x is the function f' whose value at x is

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

provided the limit exists.

If $f'(x)$ exists for some x , f is differentiable at x . If $f'(x)$ exists for all x , then f is differentiable.

This notation includes something you may have never seen before - a limit. The idea behind limits is this - we want to think about the behavior of the function as we get closer and closer to the limiting value. This is often a point where we cannot just plug in the value in question. Fortunately for us, the need for limits will mostly be situations in which we can algebraically eliminate the need for complicated limits without needing a battery of techniques.

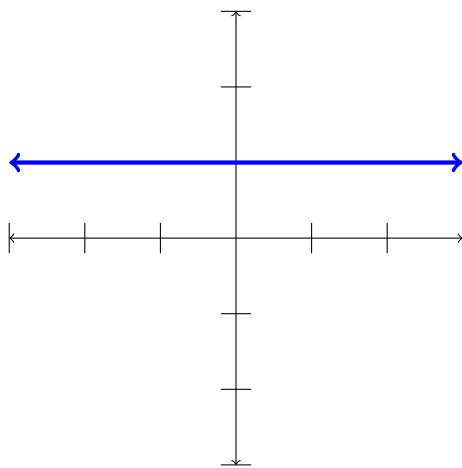
Notation 11.3.2 $f'(x)$ for a given function $y = f(x)$ can also be written as $\frac{dy}{dx}$ or $\frac{d}{dx} f(x)$.

So, thinking back to function notation, when we say $f'(a)$, we are talking about the the derivative of the function $f(x)$ at the point $x = a$.

Let's find the derivative of a couple common types of functions. We will start with power functions, which are those of the form $f(x) = x^n$ for some integer n .

Example 11.3.3 Find the derivative of $f(x) = x^0$.

Solution Remembering back to our properties of exponents, we remember that $x^0 = 1$. When we plot this function, we have



We see a horizontal line and we know that horizontal lines have a slope of 0. We can generalize this too.

Rule 11.3.4 For any constant c , we have

$$\frac{d}{dx} c = 0$$

Example 11.3.5 Find the derivative of $f(x) = x$.

Solution This should look familiar, since we are looking here for the derivative of a linear function. But we know that the derivative is the same as asking for the slope, and since the slope of a linear function is the same at all points, we know $f'(x) = 1$ here.

In general, we have

Rule 11.3.6 For any constant a ,

$$\frac{d}{dx} ax = a$$

Example 11.3.7 Find the derivative of $f(x) = x^2$.

Solution Here we will use the difference quotient, not unlike what we did earlier in this chapter. What will be different is that we will let $h \rightarrow 0$ to find the instantaneous rate of change. But, there is a problem. When we think of the difference quotient, we should think of having an h in the denominator and substituting in $h = 0$ would give us an undefined expression. So, we will need to algebraically simplify to make this issue disappear.

$$\begin{aligned} f'(x) &= \lim_{h \rightarrow 0} \frac{(x+h)^2 - x^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{x^2 + 2xh + h^2 - x^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{2xh + h^2}{h} \\ &= \lim_{h \rightarrow 0} \frac{h(2x + h)}{h} \\ &= \lim_{h \rightarrow 0} 2x + h \end{aligned}$$

At this point, we have removed the problem spot, so we can let $h \rightarrow 0$ and when we do, we see that $f'(x) = 2x$.

Example 11.3.8 Find the derivative of $f(x) = x^3$.

Solution We will proceed in a similar manner.

$$\begin{aligned}
 f'(x) &= \lim_{h \rightarrow 0} \frac{(x+h)^3 - x^3}{h} \\
 &= \lim_{h \rightarrow 0} \frac{x^3 + 3x^2h + 3xh^2 + h^3}{h} \\
 &= \lim_{h \rightarrow 0} \frac{3x^2h + 3xh^2 + h^3}{h} \\
 &= \lim_{h \rightarrow 0} \frac{h(3x^2 + 2xh + h^2)}{h} \\
 &= \lim_{h \rightarrow 0} 3x^2 + 2xh + h^2 \\
 &= 3x^2
 \end{aligned}$$

So, we have the following:

- $\frac{d}{dx} x^1 = 1$
- $\frac{d}{dx} x^2 = 2x$
- $\frac{d}{dx} x^3 = 3x^2$

Notice the pattern - in each case, the derivative is one power less than the original function, and the coefficient is the exponent from the original function.

Rule 11.3.9

$$\frac{d}{dx} x^n = nx^{n-1}$$

This rule works for negative exponents also. For example,

$$\frac{d}{dx} \frac{1}{x} = \frac{d}{dx} x^{-1} = -x^{-2} = -\frac{1}{x^2}$$

We can apply this to sums and differences too. Here are some rules that can help find derivatives.

Rule 11.3.10 *Constant Multiple Rule*

For any differentiable function $y = f(x)$,

$$\frac{d}{dx} cf(x) = c \frac{d}{dx} f(x) = cf'(x)$$

Since the constant is not the variable we are concerned with, we can factor it to the outside of the derivative. In practice, we temporarily forget about the constant, find the derivative of the function, then multiply our derivative by that constant.

Example 11.3.11

$$\begin{aligned}
 \frac{d}{dx} 5x^3 &= 5 \frac{d}{dx} x^3 \\
 &= 5(3x^2) \\
 &= 15x^2
 \end{aligned}$$

Rule 11.3.12 Sum and Difference Rule

$$\frac{d}{dx} [f(x) \pm g(x)] = f'(x) \pm g'(x)$$

This allows us to find the derivative of each term separately and then just combine the derivatives using the operation we are given. Why does this work?

$$\begin{aligned} & \lim_{h \rightarrow 0} \frac{f(x+h) \pm g(x+h) - (f(x) \pm g(x))}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \pm \frac{g(x+h) - g(x)}{h} \\ &= \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} \pm \lim_{h \rightarrow 0} \frac{g(x+h) - g(x)}{h} \\ &= f'(x) \pm g'(x) \end{aligned}$$

Example 11.3.13

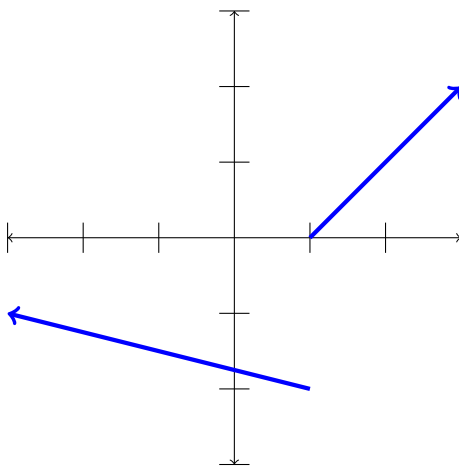
$$\begin{aligned} \frac{d}{dx} (4x^4 - x^2 + 3) &= \frac{d}{dx} 4x^4 - \frac{d}{dx} x^2 + \frac{d}{dx} 3 \\ &= 4 \frac{d}{dx} x^4 - 2x + 0 \\ &= 4(4x^3) - 2x \\ &= 16x^3 - 2x \end{aligned}$$

We do have formulas for other operations, like products and quotients, but we will not concern ourselves with those functions in this class.

11.3.1 When Derivatives Do Not Exist

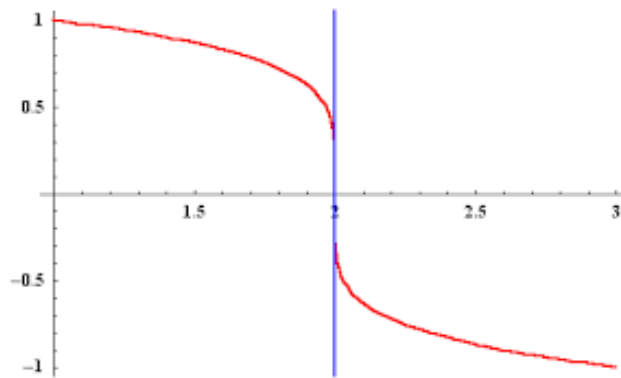
We want to make sure that when we are trying to find a derivative that one actually exists. There are three key times when one does not.

- The derivative does not exist at a point if the function is not continuous at a point. If we have a break in the curve, we cannot take the derivative because we won't get the same slope approaching the same point from both sides.

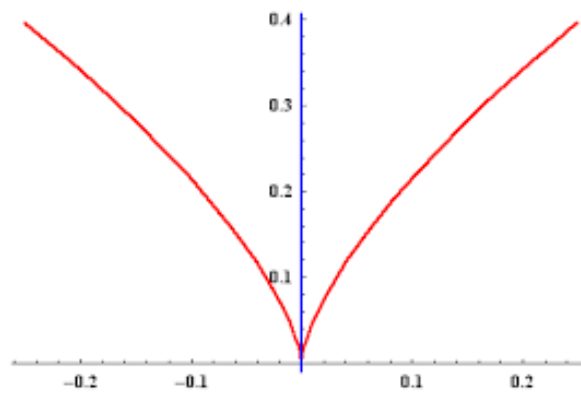


This function is not differentiable at $x = 1$ but is everywhere else.

- The derivative does not exist at a point $x = a$ if the function has a vertical tangent at that point.



- The derivative does not exist at a point $x = a$ if the function has a cusp at that point - a point where the slope on one side goes to ∞ and on the other goes to $-\infty$.



Note: this actually applies to all points where there is a sharp corner because the slope of the function on one side of the sharp corner is different than that on the other side. We cannot have a derivative where there is a drastic change in slope.

11.3.2 Exercises

1. Find the derivative of $f(x) = 4x^2 - x$.
2. Find the derivative of $g(x) = 2 - 5x + 3x^2$,
3. Find the derivative of $h(x) = 3x^{12} + 3x^{11}$.
4. Find the derivative of $j(x) = 3x - \frac{2}{x}$.
5. Find $f'(2)$ for $f(x) = 5x + 2x^2$.
6. Find $g'(-3)$ for $g(x) = -4x^3 + 2x - 1$.
7. Find $h'(4)$ for $h(x) = \frac{4}{x^2} + 3x^2$.
8. Find $j'(h + 2)$ for $j(x) = x^3 - 4x^2 + 2x$.

11.3.3 Solutions

1. $f'(x) = 8x - 1$

2. $g'(x) = -5 + 6x$

3. $h'(x) = 36x^{11} + 33x^{10}$

4. $j'(x) = 3 + \frac{2}{x^2}$

5.

$$f'(x) = 5 + 4x$$
$$f'(2) = 5 + 4(2) = 13$$

6.

$$g'(x) = -12x^2 + 2$$
$$g'(-3) = -12(-3)^2 + 2 = -106$$

7.

$$h'(x) = -\frac{8}{x^3} + 6x$$
$$h'(4) = -\frac{8}{4^3} + 6(4) = -\frac{1}{8} + 24 = \frac{191}{24}$$

8.

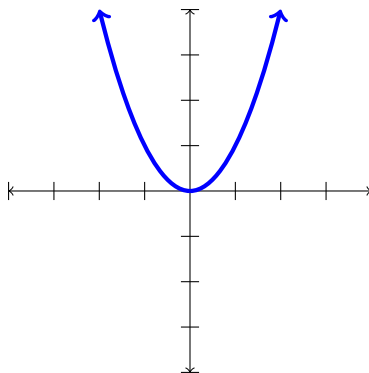
$$j'(x) = 3x^2 - 8x + 2$$
$$j'(h+2) = 3(h+2)^2 - 8(h+2) + 2$$
$$= 3h^2 + 12h + 12 - 8h - 16 + 2$$
$$= 3h^2 + 4h - 2$$

11.4 Extrema Points

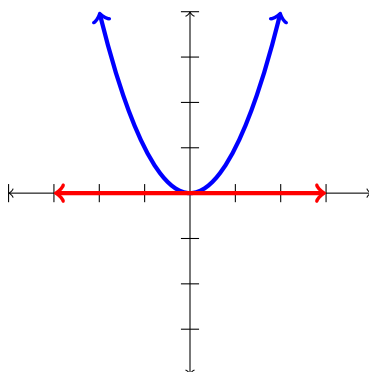
So why are we spending this time on slopes and rates of change? When looking at applications where we can model using functions, if we can find where the rate of change is zero, it tells us something important about the situation. Let's look at a couple examples so we can see how important this is.

Example 11.4.1 Find where $f(x) = x^2$ has a tangent line with a slope of 0.

Solution Let's plot the function to see what we are dealing with.



Where does this function have a horizontal tangent line? Remember, horizontal lines have a slope of 0.



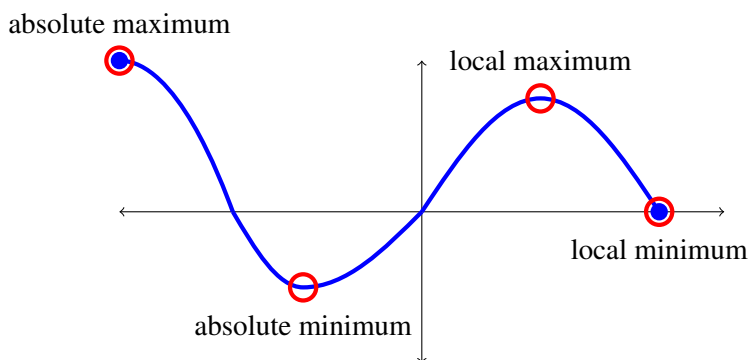
At this point $(0,0)$ we have a horizontal tangent line and notice that this point is also the smallest value on the curve. That makes this a local minimum. And, because it is the only minimum point, we can say that it is also an absolute minimum.

Definition 11.4.2 The absolute minimum, or global minimum, value of a function f occurs at a point $c \in \mathcal{D}(f)$ if $f(x) \geq f(c) \forall x \in \mathcal{D}(f)$.

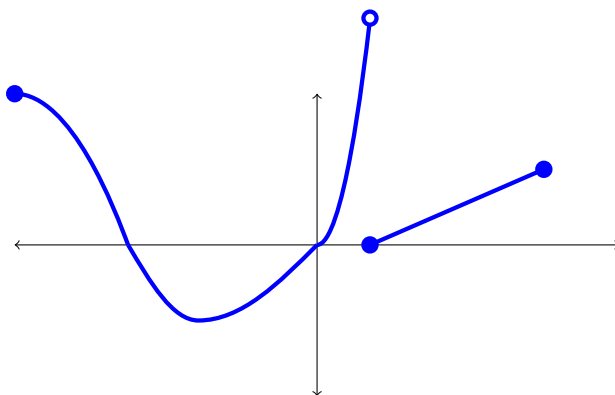
This is just a fancy way of saying that the absolute minimum is the smallest point of all points around it. We have an analogous definition for the absolute maximum as well.

Definition 11.4.3 The absolute maximum, or global maximum, value of a function f occurs at a point $c \in \mathcal{D}(f)$ if $f(x) \leq f(c) \forall x \in \mathcal{D}(f)$.

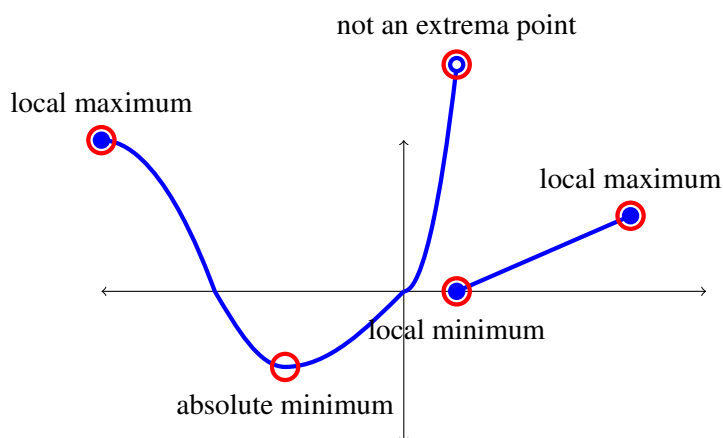
When we consider all of the places that have a horizontal tangent, they may all be local extrema, but only one value can correspond to the absolute maximum (if there is one) and one value can correspond to the absolute minimum (again, if there is one). The graph below illustrates this.



Example 11.4.4 Classify the extrema points in the following graph.



Solution We have to be careful here. If a point is closed, it can be an extrema point. But if it is an open point, we can never actually get to the value and therefore, it can never be an extrema point.



Calling these minimums and maximums is at the heart of why they are important. When we model with a function, we may be interested in optimizing the situation and being able to find the extrema point is a way to do so. By being able to find derivatives, we can algebraically solve for where these extrema points reside.

Example 11.4.5 Find any extrema points for the function $f(x) = x^2 - 3x$.

Solution A horizontal tangent line has a slope of 0, so we can find where a function has a horizontal tangent by setting the derivative, which represents the slope, equal to zero and solving for x .

$$\begin{aligned}f'(x) &= 2x - 3 = 0 \\2x &= 3 \\x &= \frac{3}{2}\end{aligned}$$

So, we get what is called critical point at $x = \frac{3}{2}$.

Definition 11.4.6 Interior points of a domain where $f'(x) = 0$ or $f'(x)$ is undefined are called critical points.

The question is, how do we know without graphing whether or not the critical point is a minimum or a maximum? We will do so using the derivative.

Definition 11.4.7 The First Derivative Test

Suppose $f'(x) = 0$.

- If f' changes from positive to negative at a critical point then the critical point is a local maximum.
- If f' changes from negative to positive at a critical point then the critical point is a local minimum.
- If f' does not change sign at a critical point then that critical point is not an extrema point.

the idea here is this: if the slope is positive, the function is increasing and if it is negative then the function is decreasing. So if we increase to a point and then decrease after the point, the point is a local maximum. But, if we decrease to a point and increase after it, then the point is a local minimum. If we increase before and after then the point is not an extrema point; similarly, if we decrease to a point and continue to decrease after then the point cannot be an extrema point either.

Back to our problem. Since $f'(1) = -1 < 0$ and $f'(2) = 1 > 0$, our point is a local minimum. So, the critical point $x = \frac{3}{2}$, being the only critical point, is an absolute minimum.

Example 11.4.8 Find any extrema points for the function $f(x) = x^3 - 3x^2 + 3x$.

Solution First, we differentiate.

$$\begin{aligned}f'(x) &= 3x^2 - 6x + 3x^2 - 6x + 3 = 0 \\3(x^2 - 2x + 1) &= 0 \\x^2 - 2x + 1 &= 0 \\(x - 1)^2 &= 0 \\x - 1 &= 0 \\x &= 1\end{aligned}$$

So, our critical point here is at $x = 1$.

Now, we check the sign of the derivative at points around $x = 1$. We see that $f'(2) = 1 > 0$ and $f'(0) = 1 > 0$, so the critical point is not an extrema point.

It is important to note that we don't need to worry here about how close to the critical point we are with the points we choose to test the first derivative. We know that a function cannot 'change direction' - it cannot have an extrema point - unless there is a critical point. So, we only need to check any point (as long as the function is continuous) to the left of the critical point and any point to the right of the critical point to make the determination.

Example 11.4.9 For $f(x) = x^3 - 9x^2 - 48x + 52$, find all critical points, evaluate the function at the critical points and endpoints, and classify those points.

1. $-5 \leq x \leq 12$
2. $-5 \leq x \leq 14$
3. $-5 \leq x < \infty$

Solution We start with $f'(x)$ and finding the critical points.

$$\begin{aligned} f'(x) &= 3x^2 - 18x - 48 = 0 \\ 3(x^2 - 6x - 16) &= 0 \\ (x - 8)(x + 2) &= 0 \\ x &= 8, -2 \end{aligned}$$

Both points are in our domain, so we will consider the values at these two points as well as at the endpoints.

$f(-5) = -58$	$f(-5)$ is a local minimum
$f(-2) = 104$	$f(-2)$ is the absolute maximum
$f(8) = -396$	$f(8)$ is the absolute minimum
$f(12) = -92$	$f(12)$ is a local maximum

Before we finish the example, an important point needs to be made. The reason we only need to check the endpoints and critical points is because we have a continuous function. This basically means that we didn't have to lift the pencil to draw the curve. If we did, we would not be guaranteed that the function even has absolute extrema points.

Theorem 11.4.10 The Extreme Value Theorem

Let f be continuous of $[a, b]$. Then f attains both its absolute minimum and maximum on $[a, b]$.

Notice that this does not guarantee absolute extrema when we don't have a closed interval or when the function is not continuous ... but back to the example.

For $-5 \leq x \leq 14$, we have

$f(-5) = -58$	$f(-5)$ is a local minimum
$f(-2) = 104$	$f(-2)$ is a local maximum
$f(8) = -396$	$f(8)$ is the absolute minimum
$f(14) = 360$	$f(12)$ is the absolute maximum

Finally, for $-5 \leq x < \infty$, we have

$f(-5) = -58$	$f(-5)$ is a local minimum	
$f(-2) = 104$	$f(-2)$ is a local maximum	Because $x \rightarrow \infty$ on this domain and
$f(8) = -396$	$f(8)$ is the absolute minimum	

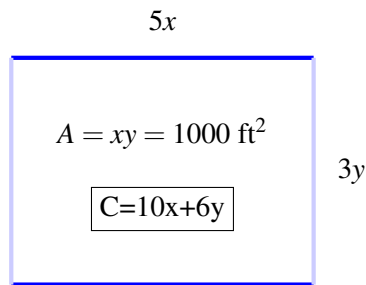
the function gets larger and larger as x gets larger and larger, there is no absolute maximum here.

11.4.1 Applications

Lets look at this idea in terms of word problems so we can see the usage in practical situations.

Example 11.4.11 *You want to fence a rectangular region of area 1000 ft^2 . You choose two different kinds of fencing to use: the parallel sides on the front and back of the region will be fenced with \$5 per foot fencing. The left and right sides will cost \$3 per foot. Minimize the cost.*

Solution We are trying to minimize the cost, so we want to find the critical point that corresponds to the smallest value of the function we need to write based on the situation. Let's start with a visualization of the situation.



Our domain is $0 < x < \infty$ because we cannot have negative length but technically there is no upper restriction as long as we have the correct area. Now, based on the picture and the situation, we set up our system of equations. We need two equations to be able to solve because there are two variables in the problem.

$$A = xy = 1000$$

$$C = 10x + 6y$$

We need to write one of the equations as a function of the other so that we can substitute to reduce an equation to one of one variable. Without loss of generality, we will solve for y here. But we had to choose this equation to rewrite because we are trying to optimize cost. So that has to be the one we substitute into afterwards.

$$xy = 1000$$

$$y = \frac{1000}{x}$$

So, we have

$$\begin{aligned}C &= 10x + 6y \\C &= 10x + \frac{6000}{x} \\C' &= 10 - \frac{6000}{x^2} = 0 \\10 &= \frac{6000}{x^2} \\x^2 &= 600 \\x &= \sqrt{600} \approx 24.49 \text{ ft}\end{aligned}$$

So, the question remains ... how do we check to see if the critical point gives a minimum as we want? We check the value at nearby points to see if this is the point that gives the smallest value.

We now use the First Derivative test.

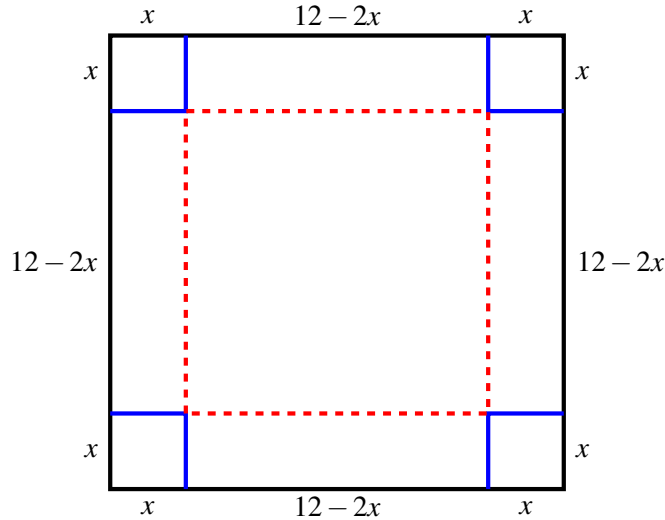
- $C'(24) = -.42 < 0$
- $C'(25) = .4 > 0$

So, we have a local minimum. Since it is the only local minimum we know this will be the absolute minimum we need. But, $x = 24.49$ feet is not the answer we are looking for. We need the cost ...

$$\begin{aligned}C &= 10(24.49) + \frac{6000}{24.49} \\&= \$489.90\end{aligned}$$

Example 11.4.12 *An open box is to be made from a square piece of material 12 inches on a side by cutting equal sized squares from each corner and turning up the sides. Find the volume, in cubic inches, of the largest box we can make in this manner.*

Solution We want to maximize the volume of the box here, so let's begin with a picture of the situation with all of the information we have filled in. Now, notice that we are cutting squares out of the corners of the rectangular sheet, but we don't know the dimensions of those squares. So let the squares be $x \times x$.



To find the volume of the box, we need to find the volume of the right rectangular prism formed by folding the box with these corners cut out. The dimensions are $12 - 2x \times 12 - 2x \times x$, so the volume we want to maximize is $V = (12 - 2x)^2 x$. Since it is a polynomial, we will multiply it out so that we can differentiate.

$$V = 144x - 48x^2 + 4x^3$$

The domain of the function is $(0, 6)$. Since we are cutting corners out of the rectangular sheet, we cannot cut two on the same side if each is longer than half the length of the side. And, we cannot cut nothing, so the length of x must be positive. We now differentiate to find the critical point(s). Notice that since V is a polynomial, it has a domain of all real numbers, as will its derivative, so there will be no critical points for the derivative being undefined.

$$\begin{aligned} V' &= 144 - 96x + 12x^2 = 0 \\ 12(12 - 8x + x^2) &= 0 \\ x^2 - 8x + 12 &= 0 \\ (x - 6)(x - 2) &= 0 \\ x &= 2, 6 \end{aligned}$$

But, since $x = 6$ is not in our domain, we get that the only critical point is $x = 2$. Now we classify. To check to if this gives maximum, we will use the First Derivative Test.

- $V'(1) = 60 > 0$
- $V'(3) = -36 < 0$

We can see that $x = 2$ is a local maximum, so the maximum volume is $V(2) = 128 \text{ in}^3$.

An important concept that relates to derivatives is the idea of marginality.

Definition 11.4.13 The marginal cost is the cost of producing one more item. It is the rate of change of the cost function, so we can define this as the derivative of the cost function.

The idea is that the slope of the cost curve gives us the change in value with respect to the variable, which is the number of units of goods. So, we can use the same ideas as before with optimization derivatives to optimize here as well.

Example 11.4.14 *A small manufacturing company reports that their production costs per week are given by the function*

$$C(x) = 250 + 100x - .1x^2$$

for $0 \leq x \leq 500$, where x is the number of items produced.

1. What is the cost to produce the 201th item?
2. What is the marginal cost at $x = 200$?
3. What is the marginal cost at $x = 201$?

Solution First, to find the cost to produce the 201st item, we can find the difference between the cost at 200 and the cost at 201.

$$\begin{aligned} C(201) - C(200) &= 250 + 100(201) - .1(201)^2 - (250 + 100(200) - .1(200)^2) \\ &= 16309.90 - 16250 \\ &= \$59.90 \end{aligned}$$

Now, we will find the marginal costs, where we first need to find the derivative.

$$\begin{aligned} C'(x) &= 100 - .2x \\ C'(200) &= \$60 \\ C'(201) &= \$59.80 \end{aligned}$$

Notice that the marginal cost is not the same as the actual cost. This is because the marginal cost is based on the rate of change at a point, but is actually giving the additional cost to produce the next item. So, when $x = 200$, we are actually approximating that it would cost \$60 to make the 201st item and when $x = 201$, we are approximating the additional cost of producing the 202nd item.

Example 11.4.15 *If the Providence Grays report that their cost for producing x baseball bats is given by*

$$C(x) = 200 + 150x - .25x^2$$

What is the marginal cost at $x = 100$?

Solution We differentiate as before and then substitute $x = 100$ into the result.

$$\begin{aligned} C'(x) &= 150 - .5x \\ C'(100) &= 150 - .5(100) \\ &= 100 \end{aligned}$$

So, it would cost \$100 to produce the 101st bat.

The marginal revenue and marginal profit work the same way as marginal cost.

Definition 11.4.16 The marginal revenue is the amount of revenue generated by selling one more item. It is the rate of change of the revenue function, so we can define this as the derivative of the revenue function.

So, the marginal revenue tells us how much more money we would make on selling the next item. We can think of the revenue function as

$$R(x) = x \cdot p(x)$$

where x is again the number of items and $p(x)$ is the demand function. This function is literally the result of multiplying the demand by how many items we sold.

Example 11.4.17 The demand for a certain new phone is given by

$$p(x) = 300 - .01x$$

Find the marginal revenue from selling 500 phones.

Solution The revenue function is given by

$$R(x) = x \cdot p(x) = x(300 - .01x) = 300x - .01x^2$$

So, we can find the marginal revenue by differentiating.

$$\begin{aligned} R'(x) &= 300 - .02x \\ R'(500) &= 300 - .02(500) \\ &= \$290 \end{aligned}$$

So, the company would make \$290 off of selling the 501st phone.

This is how much they would sell the phone for, but it is not the profit. To be able to find that, we need to consider the cost and the demand.

Definition 11.4.18 The profit is the difference between the revenue and the cost. The profit function $P(x)$, is given by

$$P(x) = R(x) - C(x) = x \cdot p(x) - C(x)$$

where x is the number of units.

Example 11.4.19 Suppose a small electronics accessory company can produce a maximum of 5000 docking stations per month. The demand is governed by the equation

$$p(x) = 125 - .003x$$

And the cost function is given by

$$C(x) = 10,000 + 50x - .05x^2 + .00002x^3$$

Find the marginal cost, marginal revenue and marginal profit when 2000 docking stations are sold, assuming that the company sells exactly the same number of items that are produced.

Before we proceed, note the last line of the question. Why is this important? By making this assumption, we don't have to worry about real-life concerns of businesses, like overhead costs, warehouse space, etc.

Solution We will first use the information we have to find the revenue and profit functions.

$$\begin{aligned}R(x) &= x \cdot p(x) = x(125 - .003x) = 125x - .003x^2 \\P(x) &= R(x) - C(x) = 125x - .003x^2 - (10,000 + 50x - .05x^2 + .00002x^3) \\&= -10,000 + 75x + .047x^2 - .00002x^3\end{aligned}$$

Now that we have all of the functions we need, we can find the derivatives and corresponding marginal values.

$$\begin{aligned}C'(x) &= 50 - .1x + .00006x^2 \\C'(2000) &= 90\end{aligned}$$

This tells us that the cost for producing the 2001st docking station is \$90.

$$\begin{aligned}R'(x) &= 125 - .006x \\R'(2000) &= 113\end{aligned}$$

The revenue generated from selling the 2001st docking station will be \$113.

$$\begin{aligned}P'(x) &= 75 + .094x - .00006x^2 \\P'(2000) &= 23\end{aligned}$$

Finally, the profit made by selling the 2001st docking station is \$23.

How many items should be produced in order to maximize the profit? We can find this by setting the marginal cost equal to the marginal revenue.

$$\begin{aligned}C'(x) &= R'(x) \\50 - .1x + .00006x^2 &= 125 - .006x \\-0006x^2 - .0094x - 75 &= 0\end{aligned}$$

Due to the coefficients having so many decimal places, this is a prime candidate for the quadratic formula, which yields $x = -581.8$ and $x = 2148.47$ but we know we cannot produce a negative number of docking stations. So, we see that profit will be maximized at approximately 2148 docking stations manufactured.

11.4.2 Exercises

1. Find and classify any extrema points of the function $f(x) = x^2 - 4x$.
2. Find and classify any extrema points of the function $f(x) = x^3 - 6x^2 + 9x$.
3. Find and classify any extrema points of the function $f(x) = 4x^2 - 8x$.
4. Find and classify any extrema points of the function $f(x) = x^4 - 8x^2$.
5. It is desired to fence a rectangular pasture of area 9600 yds² and divide it into two parts by one more fence across the shorter dimension of the pasture. Find the minimum length (in yards) of the fencing required.
6. A landscape architect plans to enclose a 3000 square foot rectangular region in a botanical garden. She will use shrubs costing \$25 per foot along three sides and fencing costing \$10 per foot along the fourth side. Find the minimum total cost.
7. A tire company reports that their profit per tire is given by

$$P(t) = 0.04t^2 + 40t - 640$$

- (a) Find the profit on 75 tires.
 - (b) Find the marginal profit at 75 tires.
8. Let $C(x) = 900 + 5x + .01x^2$ be the cost of producing x pairs of hockey skates.
 - (a) Find the marginal cost function.
 - (b) Find $C'(300)$ and explain what this answer means.
 - (c) Find the actual cost of producing the 301st pair of hockey skates and explain why this answer is not the same as $C'(300)$.
 9. Suppose the revenue generated from selling t-shirts is given by the function $R(x) = -.003x^2 + 75x + 300$. Find the marginal revenue function and use it to find the number of t-shirts that would maximize revenue. (Hint: how do we find extrema points of functions?)
 10. The Newquist guitar manufacturing company has a demand given by $p(g) = 3400 - 14g$ where g represents the number of stratocasters. The cost function of manufacturing g guitars is given by $c(g) = 32000 + 1000g - 3.2g^2 + .008g^3$.
 - (a) Find the marginal cost function.
 - (b) Find the marginal revenue function.
 - (c) Find the marginal profit function.
 - (d) Find the number of guitars that should be produced in order to maximize profit.

11.4.3 Solutions

1. Find and classify any extrema points of the function $f(x) = x^2 - 4x$.

$$f'(x) = 2x - 4 = 0$$

$$2x = 4$$

$$x = 2$$

$$f'(1) = -2 < 0$$

$$f'(3) = 2 > 0$$

Since the sign of f' changes from negative to positive at $x = 2$, we have a local minimum at $x = 2$.

2. Find and classify any extrema points of the function $f(x) = x^3 - 6x^2 + 9x$.

$$f'(x) = 3x^2 - 12x + 9 = 0$$

$$x^2 - 4x + 3 = 0$$

$$(x - 3)(x - 1) = 0$$

$$x = 1, 3$$

$$f'(0) = 9 > 0$$

$$f'(2) = -3 < 0$$

$$f'(4) = 9 > 0$$

The sign of f' changes from positive to negative at $x = 1$ and from negative to positive at $x = 3$, we can classify $x = 1$ as a local maximum and $x = 3$ as a local minimum.

3. Find and classify any extrema points of the function $f(x) = 4x^2 - 8x$.

$$f'(x) = 8x - 8 = 0$$

$$8x = 8$$

$$x = 1$$

$$f'(0) = -8 < 0$$

$$f'(2) = 8 > 0$$

Since the sign of f' changes from negative to positive at $x = 1$, we can classify $x = 1$ as a local minimum.

4. Find and classify any extrema points of the function $f(x) = x^4 - 8x^2$.

$$f'(x) = 4x^3 - 16x = 0$$

$$4x(x^2 - 4) = 0$$

$$4x(x + 2)(x - 2) = 0$$

$$x = 0, \pm 2$$

$$f'(-3) = -60 < 0$$

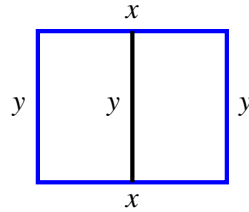
$$f'(-1) = 12 > 0$$

$$f'(1) = -12 < 0$$

$$f'(3) = 60 > 0$$

Since f' changes from negative to positive at $x = \pm 2$ and changes from positive to negative at $x = 0$, we can classify $x = \pm 2$ as local minimums and $x = 0$ as a local maximum.

5. It is desired to fence a rectangular pasture of area 9600 yds² and divide it into two parts by one more fence across the shorter dimension of the pasture. Find the minimum length (in yards) of the fencing required.



We want to minimize the amount of fencing, which is given by $F = 2x + 3y$. Since we have two variables, we need another equation that relates x and y . We are given the area of 9600 yds² and since the region is rectangular, we have $xy = 9600$. We need to rewrite this to solve for one variable so that we can write F in terms of one variable.

$$y = \frac{9600}{x}$$

$$F = 2x + 3\left(\frac{9600}{x}\right) = 2x + \frac{28800}{x}$$

$$F' = 2 - \frac{28800}{x^2} = 0$$

$$2x^2 - 28800 = 0$$

$$x^2 - 14400 = 0$$

$$x^2 = 14400$$

$$x = \pm 120$$

Since we cannot have a negative length, the only critical point is $x = 120$. Also, notice that the domain would be $[0, \infty)$ since we cannot have a negative length but we technically could make as long a fence as we want as long as the area is at the set amount. We of course know that we would get no area for $x = 0$ for the length, however.

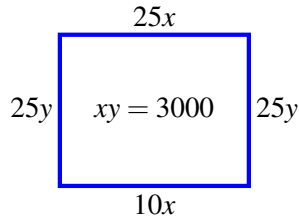
$$F'(100) = -\frac{22}{25}$$

$$F'(150) = \frac{18}{25}$$

Since the sign of F' changes from negative to positive, we have a local minimum at $x = 120$. So, substituting this into our fencing equation, we get our total amount of fence needed.

$$F = 2(120) + \frac{28800}{120} = 9600 \text{ yds}^2$$

6. A landscape architect plans to enclose a 3000 square foot rectangular region in a botanical garden. She will use shrubs costing \$25 per foot along three sides and fencing costing \$10 per foot along the fourth side. Find the minimum total cost.



Here, we are looking to minimize the cost, which is given by $C = 50y + 35x$. Since there are two variables in the cost equation, we will rewrite the second equation to solve for one variable and then substitute into our object function.

$$\begin{aligned}
 xy = 3000 &\Rightarrow y = \frac{3000}{x} \\
 C = 50\left(\frac{3000}{x}\right) + 35x &= \frac{150000}{x} + 35x \\
 C' &= 35 - \frac{150000}{x^2} = 0 \\
 35x^2 - 150000 &= 0 \\
 35x^2 &= 150000 \\
 x^2 &\approx 4285.71 \\
 x &= \pm 65.47
 \end{aligned}$$

Since the length cannot be negative, it must be that the critical point here is $x = 65.47$. And, with out domain being $[0, \infty)$, we would get no area in either extreme case. This means we only need consider the critical point.

$$\begin{aligned}
 C'(60) &= -\frac{20}{3} < 0 \\
 C'(70) &= \frac{215}{49} > 0
 \end{aligned}$$

Since the sign of F' changes from negative to positive, we have a local minimum at $x = 65.47$. Substituting this into cost equation gives

$$C = \frac{150000}{65.47} + 35(65.47) = \$4582.58$$

7. A tire company reports that their profit per tire is given by

$$P(t) = 0.04t^2 + 40t - 640$$

(a) Find the profit on 75 tires.

$$P(75) = .04(75)^2 + 40(75) - 640 = \$2585$$

(b) Find the marginal profit at 75 tires.

$$\begin{aligned}
 P'(t) &= .08t + 40 \\
 P'(75) &= \$46
 \end{aligned}$$

8. Let $C(x) = 900 + 5x + .01x^2$ be the cost of producing x pairs of hockey skates.

(a) Find the marginal cost function.

$$C'(x) = 5 + .02x$$

(b) Find $C'(300)$ and explain what this answer means.

$$C'(300) = 5 + .02(300) = 11$$

The cost to produce the 301st pair of hockey skates is \$11.

(c) Find the actual cost of producing the 301st pair of hockey skates and explain why this answer is not the same as $C'(300)$.

$$C(300) = \$3300$$

$$C(301) = \$3311.01$$

The actual cost to make the 301st pair of hockey skates is \$11.01. The marginal cost is based on the price at the time of producing the 300th pair and projecting that the cost for the next pair will be at the same rate. Unless the function is linear, the rate of change is not constant and will therefore not be the same.

9. Suppose the revenue generated from selling t-shirts is given by the function $R(x) = -.003x^2 + 75x + 300$. Find the marginal revenue function and use it to find the number of t-shirts that would maximize revenue. (Hint: how do we find extrema points of functions?)

$$R'(x) = -.006x + 75 = 0$$

$$.006x = 75$$

$$x = 12500$$

$$R'(12000) = 3 > 0$$

$$R'(13000) = -3 < 0$$

Since the sign of R' changes from positive to negative, we can classify $x = 12500$ as a local maximum. So, we would maximize revenue by selling 12500 t-shirts.

10. The Newquist guitar manufacturing company has a demand given by $p(g) = 3400 - 14g$ where g represents the number of stratocasters. The cost function of manufacturing g guitars is given by $c(g) = 32000 + 1000g - 3.2g^2 + .008g^3$.

(a) Find the marginal cost function.

$$C'(g) = 1000 - 6.4g + .024g^2$$

(b) Find the marginal revenue function.

$$R(g) = g(3400 - 14g) = 3400g - 14g^2$$

$$R'(g) = 3400 - 28g$$

(c) Find the marginal profit function.

$$P(g) = 3400 - 14g^2 - (32000 + 1000g - 3.2g^2 + .008g^3)$$
$$P'(g) = 3400 - 28g - 1000 + 6.4g - .024g^2 = 2400 - 21.6g - .024g^2$$

(d) Find the number of guitars that should be produced in order to maximize profit.

$$2400 - 21.6g - .024g^2 = 0$$
$$g = 100, -1000$$

We certainly cannot produce a negative number of guitars, so we have only one critical point to consider.

$$P'(99) = 26.376 > 0 \quad P'(101) = -26.424 < 0$$

The sign of $P'(g)$ changes from positive to negative at $g = 100$, so we will maximize profit by selling 100 guitars.