Quantitative Tools for Economics and Business

Lecture Notes

Tim Murray, PhD
Virginia Military Institute

October 2022

Table of Contents

Part I

Solving Economic Models

Section 1

Review of Algebra

- In economics and business, we use a lot of quantitative tools to help us better understand the world we live in
  - In principles classes we do lots of graphing
  - In intermediate micro we show the algebra behind all of the graphing
  - In business finance, operations management, econometric, and data analytics you will use algebra and statistics
- This class will introduce you to many of the basic tools you will use in future and current ECBU classes
  - Algebra
  - Statistics
  - Probability
  - Future/Present Value
  - Microsoft Excel Tools
- We will spend some time on interpretation of the solutions in this class, but those will be more heavily explored in future classes
- Having been exposed to the quantitative tools one time before you use them will make it easier the next time you see them
- For some of you, you will have seen this material before and it will be a review
- For others, this may be challenging
- I encourage you to work with each other and seek help when needed.
- Some things in this class build on each other, stay on top of the in-class work and homework, otherwise this class can be difficult

1.1 Rules of Exponents

- A number raised to a power represents a product where the same number is used as a repeated factor
- In the expression b^n, b is the base and n is the exponent.
- This means the base, b is multiplied together n times

Example

$$5^3 = 5 \times 5 \times 5 = 125$$
$$10^6 = 10 \times 10 \times 10 \times 10 \times 10 \times 10 = 1,000,000$$
$$x^4 = x \times x \times x \times x$$

- When encountering exponents, there are rules that need to be followed:

The letters a and b represent any possible number and the letter x represents any possible variable.

Rule	Example
$y^1 = y$	$6^1 = 6$
$a^0 = 1$	$5^0 = 1$
$x^{-1} = \frac{1}{x}$	$K^{-2} = \frac{1}{K^2}$ *or* $\frac{1}{K^{-2}} = K^2$
$\frac{x^a}{x^b} = x^{a-b}$	$\frac{L^5}{L^2} = L^{5-2} = L^3$
$x^a x^b = x^{a+b}$	$x^2 x^3 = x^{2+3} = x^5$
$\frac{a^x}{b^x} = \left(\frac{a}{b}\right)^x$	$\frac{K^5}{L^5} = \left(\frac{K}{L}\right)^5$
$(x^a)^b = x^{a \times b}$	$(L^2)^3 = L^{2\times 3} = L^6$
$x^{a/b} = \sqrt[b]{x^a}$	$K^{1/2} = \sqrt{K}$ $L^{2/3} = \sqrt[3]{L^2}$
$\left(\frac{1}{x}\right)^a = \frac{1}{x^a}$	$\left(\frac{1}{x}\right)^2 = \frac{1}{x^2}$

1.2 Rules of the Natural Log Function

Rule	**Example**
$\ln(1) = 0$	
$\ln(0) =$undefined	
$\ln(e) = 1$	
$\ln(x^a) = a\ln(x)$	$\ln(x^2) = 2\ln(x)$
$\ln(x \cdot y) = \ln(x) + \ln(y)$	$\ln(x^2y^3) = \ln(x^2) + \ln(y^3)$
$\ln(\frac{x}{y}) = \ln(x) - \ln(y)$	$\ln(\frac{x^2}{y^3}) = \ln(x^2) - \ln(x^3)$
$\ln(e^x) = x$	
$e^{ln(x)} = x$	

1.3 Algebraic Rules and Properties

- When simplifying an algebraic expression or solving an equation, there are some rules, conventions, and "tricks" that can be helpful and necessary

Factoring

If multiple terms are added or subtracted together, a common term can be factored out. Suppose you have the following expression:

$$2x^3 + 4x^2 + 10x =$$

Each term in this equation has an x term and a term that can be multiplied by 2. Therefore, $2x$ can be factored out of this expression:

$$2x(x^2 + 2x + 5)$$

Distributive Property

When two terms are separated with parentheses, we can multiply them together by distributing the term outside of the parenthesis to each term inside the parentheses. This is the opposite of what we just did with factoring. We can use our previous example:

$$2x(x^2 + 2x + 5) = 0$$

$$2x^3 + 4x^2 + 10x = 0$$

If two expressions are multiplied together by parentheses, then we can use the "FOIL" method. Suppose we have the following expression:

$$(x+1)(x+2)$$

The "FOIL" method says we multiple the first terms of each expression together, the outside terms together, the inside terms together, and the last terms together.

$$(x+1)(x+2)$$

$$(x+1)(x+2) = x^2 + 2x + x + 3$$
$$(x+1)(x+2) = x^2 + 3x + 3$$

If an expression in parenthesis is raise to a power then you cannot distribute:

$$2x(x+1)^2 \neq (2x^2 + 2x)^2$$

This is because $(x+1)^2 = (x+1)(x+1)$, so there are technically three terms in this expression:

$$2x(x+1)^2 = 2x(x+1)(x+1)$$

Once you expand the term raised to a power, then you can use the distributive property and "FOIL" method:

$$(2x^2 + 2x)(x+1) = 2x^3 + 2x^2 + 2x^2 + 1$$

$$2x^3 + 4x^2 + 1$$

Multiplying by "1"

Anything divided by itself is equal to one: $\frac{3x}{3x} = 1$.

When trying to find a common denominator, we can use this principle. Suppose you have the following expression:

$$\frac{1}{x} + \frac{2}{x^2}$$

To simplify this expression, we need to find a common denominator, which in this example, is x^2. To make both terms have a denominator of x^2, we can multiply the first term by $\frac{x}{x}$, which equals 1.

$$\left(\frac{x}{x}\right)\frac{1}{x} + \frac{2}{x^2}$$

$$\frac{x}{x^2} + \frac{2}{x^2}$$

$$\frac{x+2}{x^2}$$

Division Property of Equality

If two expressions are equal to each other and you divide both sides by the same number that is not equal to zero, the resulting expressions will also be equivalent. So if $a = b$ then $\frac{a}{c} = \frac{b}{c}$ as long as $c \neq 0$.

$$3x = 24$$
$$\frac{3}{3}x = \frac{24}{3}$$
$$x = \frac{24}{3}$$
$$x = 8$$

Multiplication Property of Equality

If two expressions are equal to each other and you multiply both sides by the same number, the resulting expressions will also be equivalent. So, if $a = b$ then $ac = bc$.

$$\frac{10}{x} = 5$$
$$x\left(\frac{10}{x}\right) = 5(x)$$
$$10 = 5x$$
$$x = \frac{10}{5}$$
$$x = 2$$

Cross Multiplication

When a fraction is equal to another fraction, it can be solved by cross multiplication

$$\frac{3}{x} = \frac{x}{3}$$
$$\frac{3}{x} \nearrow\!\!\!\!\searrow \frac{x}{3}$$
$$9 = x^2$$
$$\sqrt{9} = x$$
$$x = 3$$

Fraction Divided by Another Fraction

When a fraction is divided by another fraction, the expression can be simplified by multiplying the numerator by the reciprocal (inverse) of the denominator

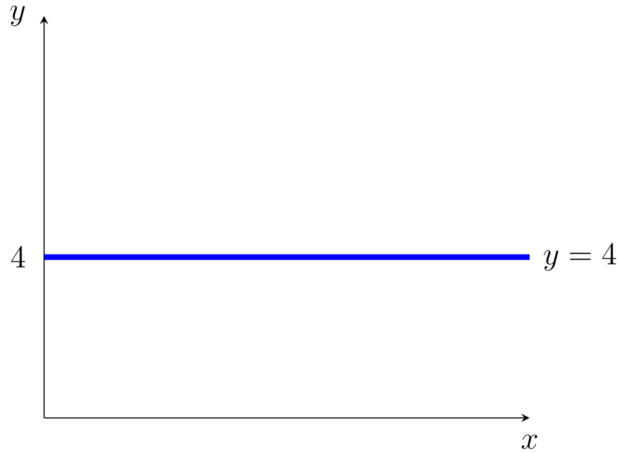
$$\frac{a/b}{c/d} = \frac{a}{b} \times \frac{d}{c}$$

$$\frac{\left(\frac{1}{2}\right)}{\left(\frac{5}{7}\right)} = \frac{1}{2} \times \frac{7}{5} = \frac{7}{10} \qquad \frac{\left(\frac{x^2}{y^2}\right)}{\left(\frac{2x}{y}\right)} = \frac{x^2}{y^2} \times \frac{y}{2x} = \frac{x}{2y}$$

1.4 Linear Equations in One Variable

- A linear equation of one variable is an equation of a straight line

Figure 1.1: Linear Equation with One Variable



- Figure 1.1 shows an example of a linear equation in one variable, $y = 4$
- In this equation, $y = 4$ for every value of x
- If we have a linear equation with one variable, we can always find the solution to that equation, even if the variable shows up multiple times

Example

$$5x + 2 = 3x - 6$$
$$2x = -8$$

$$x = -4$$

- We could have a more complex equation where we have to apply the distributive property:

Example

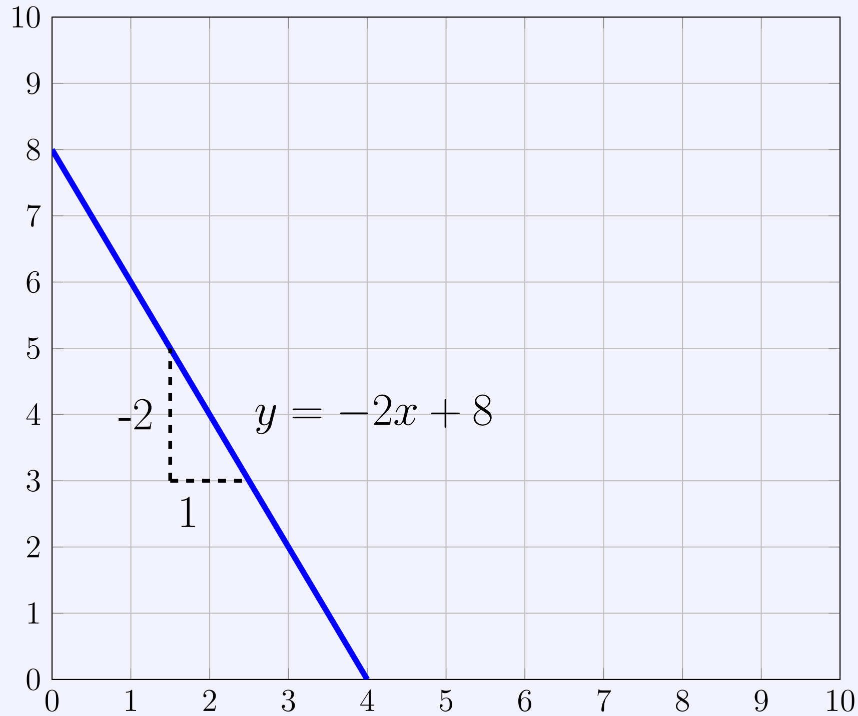
$$4(x-3)+12 = 15-5(x+6)$$
$$4x-12+12 = 15-5x-30$$
$$4x = -15-5x$$
$$9x = -15$$
$$x = -\frac{15}{9} = -\frac{5}{3}$$

1.5 Linear Equation with Two Variables

- Basic equation of a line: $y = mx + b$
  - $m =$ slope of a line
  - $b =$ the y-intercept
- Slope measures the rate of change, or, as x changes, how much does y change

Example

$$y = -2x + 8$$



- If we plug in any point for y, we can find out what the x value would be on that line, and vice versa
- Being able to plug in a value for x or y and solve for the other variable is a very important skill
- Example: If $y = 6$, what is x?

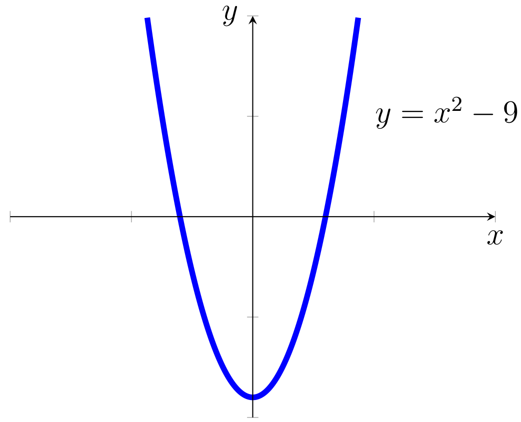
$$6 = -2x + 8$$
$$6 - 8 = -2x$$
$$-2 = -2x$$
$$x = \frac{-2}{-2} = 1$$

1.6 Solving Quadratic Equations

- A quadratic term is a variable that is raised to the power of 2 (e.g., x^2)

- When solving a quadratic equation, we want to find what the values of x when $y = 0$
  - We want to find the two x values where the graph crosses the x-axis

Figure 1.2: Graphical Example of a Quadratic Equation



- Depending on the specific equation and shape of the graph, there could be 0, 1, or 2 solutions to a quadratic equation
- If a quadratic equation does not have a linear term (e.g., $ax^2 + c = 0$), like the graph shown in Figure 1.2, then it can be solved using the square root property
  - If $x^2 = k$, then $x = \pm\sqrt{k}$

Example

Using the equation from Figure 1.2, we can see that this graph crosses the x-axis at two points, one point is negative and one point is positive. Since $y = x^2 - 9$ does not have a linear term, we can solve this using the square root property.

$$x^2 - 9 = 0$$
$$x^2 = 9$$
$$x = \pm\sqrt{9}$$
$$x = 3 \text{ and } x = -3$$

- If the quadratic equation has a linear term (e.g., $ax^2 + bx + c = 0$), then there are a few ways to solve this equation
  - Factoring
  - Grouping
  - Quadratic formula
- The quadratic formula will *always* give you the solutions to a quadratic equation
  - If $ax^2 + bx + c = 0$, then $x = \dfrac{-b \pm \sqrt{b^2 - 4ac}}{2a}$

Example

$$4x^2 - 6x + 2 = 0$$

$$a = 4 \quad b = -6 \quad c = 2$$

$$x = \frac{6 \pm \sqrt{(-6)^2 - 4(4)(2)}}{2(4)}$$

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

1.7 Solving Rational Equations

- A rational equation is an equation, which typically has fractions, that can be converted into a liner equation with some algebraic manipulation
- To add or subtract fractions, they must have the same denominator
- We can use algebraic rules from section 1.3 to find a common denominator and solve a rational equation

Example

$$\frac{7}{2x} - \frac{5}{3x} = \frac{22}{3}$$

To solve this problem, we need to find a common denominator. One trick to find this is to multiply the terms that are different, together. For example, $2 \times 3 = 6$. So, we can use $6x$ as the common denominator.

We can use the "Multiplying by '1'" rule to get a common denominator:

$$\left(\frac{3}{3}\right)\frac{7}{2x} - \left(\frac{2}{2}\right)\frac{5}{3x} = \frac{22}{3}$$

$$\frac{21}{6x} - \frac{10}{6x} = \frac{22}{3}$$

$$\frac{11}{6x} = \frac{22}{3}$$

From here, we can use cross multiplication to get a linear equation:

$$132x = 33$$

Now, we can solve for x:

$$x = \frac{33}{132} = \frac{1}{4}$$

1.8 Solving Systems of Equations

- Suppose you have multiple equations and multiple unknown variables and you want to figure out what those variables are
- This is called a system of equations
- We are going to assume that the number of equations and the number of unknown variables are the same (if they are not, it gets more complicated)
- There are three ways to solve a system of equations:
  - The substitution method
    - Rearrange one equation in terms of a single variable
    - Substitute the rearranged equation into the other equation
  - The elimination method
    - Rearrange both equations so that one variable is by itself (same variable for both equations)
    - Set them equal to each other
  - Graphing

Example

Suppose you have the following system of equations:

$$y - 2 = 5x$$

$$y = 10x - 4$$

Substitution Method

Since $y = 10x - 4$, substitute this equation into the other one:

$$10x - 4 - 2 = 5x$$

Now there is a linear equation of one variable.

$$10x - 6 = 5x$$

$$5x = 6$$

$$x = \frac{6}{5}$$

Plug $x = \frac{6}{5}$ into either equation to solve for y

$$y = 10\left(\frac{6}{5}\right) - 4$$

$$y = \frac{60}{5} - 4 = 8$$

Elimination Method

Rearrange both equations in terms of y:

$$y = 5x + 2$$

$$y = 10x - 4$$

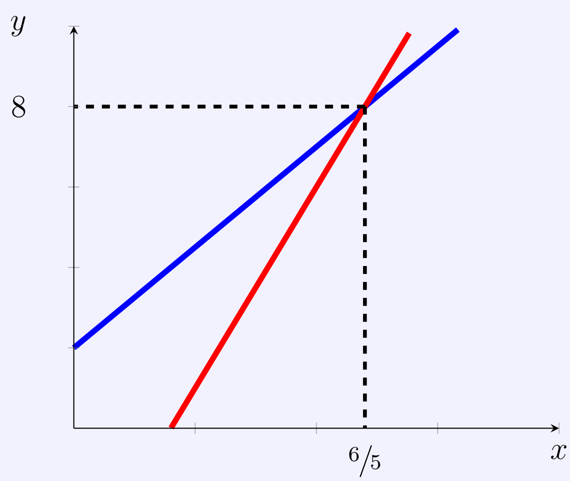
Replace y in the first equation with the y in the second equation to create an equation that is all of one variable, x, then you can solve for x:

$$10x - 4 = 5x + 2$$

$$10x - 5x = 2 + 4$$
$$5x = 6$$
$$x = \frac{6}{5}$$

$$y = \frac{60}{5} - 4 = 8$$

Graphing

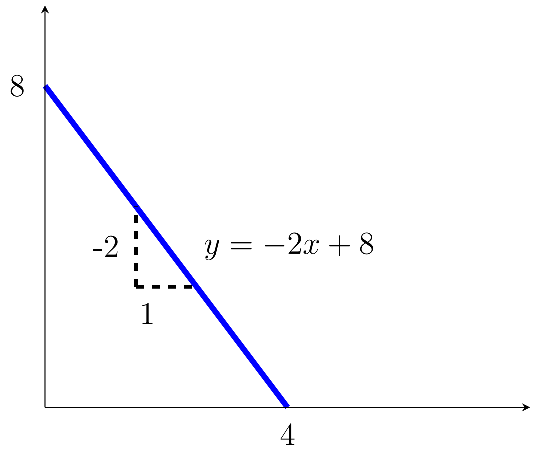


Section 2

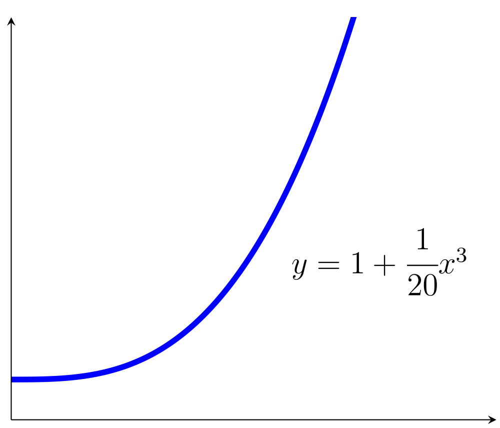
Introduction to Calculus

- There are two major branches of calculus
  - Differential calculus is about instantaneous rates of change and the slopes of curves
  - Integral calculus is about the accumulation of quantities and areas under and between curves
- In economics and business, we often use differential calculus to evaluate change and to answer important policy questions
  - What is the change in employment if minimum wage increases by $1.00?
  - If a company hires 1 more worker, how much will its profits increase?
  - What is the decrease in number of COVID-19 cases if 1% more people get vaccinated?
- These questions all revolve around what is the impact of something changing
- In economics and business, lots of the questions we ask revolve around the effect of change
- We can think of mathematical equations as relationships and math helps us evaluate change
- To evaluate the effect of change, we will use calculus in two ways
  - Evaluate change using models that approximate what we observe in the world
  - Evaluate change using statistical regression
- In this class, we will mostly focus on models, but we will briefly introduce statistical regression

2.1 Derivatives

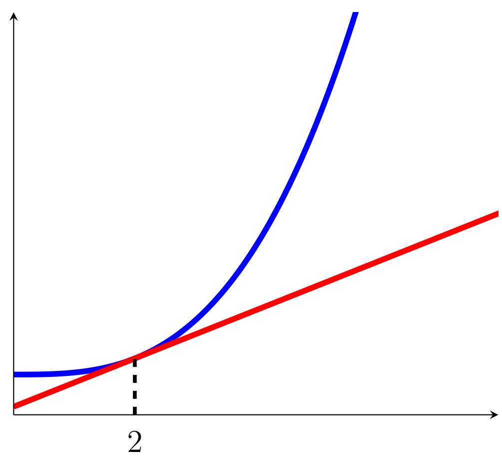


- The equation of a line is $y = mx + b$ where m is the slope
- At all points on a line, the slope is m
- This means that for any change in x, y changes by m
- In our example, a 1 unit increase in x leads to a -2 unit decrease in y, regardless of what the value of x is
- This always the case when a relationship is linear
- What if the relationship is not linear?



- In this example, the slope when $x = 2$ is not the same as when $x = 5$

- This means that a 1 unit increase in x does not have the same impact on y as x gets larger
- In order to determine the slope at each point, we need to calculate the derivative of the equation
- The derivative of an equation tells you what the slope is at any given point
- We can visually think of a derivative as the slope of a line that lies tangent to the point we are interested in



2.2 Calculating Derivatives

2.2.1 Exponents

The Basics

- The notation used for a derivative is $\frac{dy}{dx}$
- This means the derivative of function y with respect to variable x
- We will often use different letters to represent different types of relationships, so you also may see something like $\frac{dQ}{dK}$
- This notation will be important because we are it is not uncommon for equations to have more than one variable, but we will build up to that
- The "general formula" for a function with exponents is the Power Rule:

The Power Rule

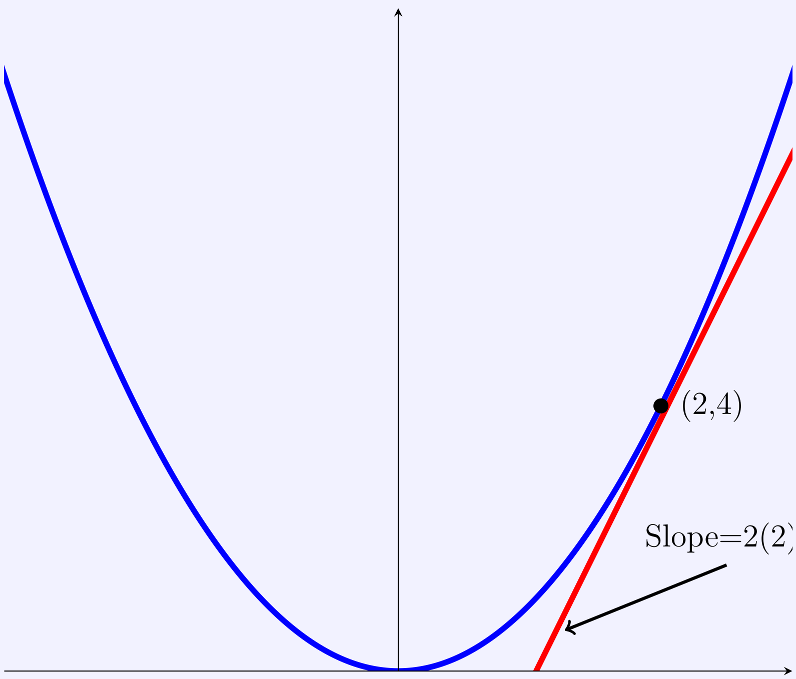
If the function is $y = x^n$ then the derivative is $\frac{dy}{dx} = nx^{n-1}$

Example 1

If $y = x^2$, then $n = 2$

$\frac{dy}{dx} = 2x^{2-1} = 2x$

So what does this mean?



The derivative gives us an equation that that tells us what the slope (rate of change) is for any value of x

For example, if $x = 2$, then the slope at that point $= 2(2) = 4$

If $x = -3$, then the slope $= 2(-3) = -6$

Example 2

If $y = x^3$, the $n = 3$

$\frac{dy}{dx} = 3x^{3-1} = 3x^2$

Example 3

If $y = x$, this is the same as $y = x^1$, so $n = 1$

$\frac{dy}{dx} = 1x^{1-1} = x^0 = 1$

Example 4

If $y = 3$, this is the same as $y = 3x^0$, so $n = 0$

$\frac{dy}{dx} = (0 \times 3)x^{0-1} = 0$

If you think of this graphically, it is a horizontal line. For each value of x, y does not change. For any "constant value" (i.e., where x^0), then $\frac{dy}{dx}$ *always* equals 0.

Functions with Exponents Multiplied by a Constant

- We can use a more generalized rule to help with this:

  If the function is $y = cx^n$, where c is any constant number, then $\frac{dy}{dx} = (c \times n)x^{n-1}$

Example

If $y = 5x^3$, then $n = 3$, and $c = 5$

So, $\frac{dy}{dx} = (5 \times 3)x^{3-1} = 15x^2$

2.2.2 The Sum and Difference Rules

- Suppose we have a function that is two functions that are added or subtracted together
- In this case, we take the derivative of each "part" independent of the addition or subtraction sign

The Sum/Difference Rule

If $y = f(x) + g(x)$ then $\frac{dy}{dx} = \frac{df(x)}{dx} + \frac{dg(x)}{dx}$

Example

Suppose $y = 5x^3 + 6x$

Part 1: $f(x) = 5x^3$, so $\frac{df(x)}{dx} = 15x^2$

Part 2: $g(x) = 6x$, so $\frac{dg(x)}{dx} = 6$

Therefore, $\frac{dy}{dx} = 15x^2 + 6$

Complete Practice Worksheet

2.2.3 Derivatives of Different Types of Functions

- Up until this point, we have been discussing functions with exponents, but there are other types of functions:
  - $y = \ln(x)$
  - $y = \sin(x)$
  - $y = \cos(x)$
  - $y = c^x$
- We rarely use trigonometric functions in economics or business, so we will not be covering them
  - If you are interested in them, a quick Google search will bring them up
- The natural log, however, is a very commonly used function in economics and business and is one you will need to know
- If $y = \ln(x)$, then $\frac{dy}{dx} = \frac{1}{x}$
- If $y = c^x$. then $\frac{dy}{dx} = c^x \ln(c)$

2.2.4 The Product Rule

- Suppose you want to take the derivative of two functions that are multiplied together: $y = x^2 \ln(x)$
- To find this derivative, we will need to use the product rule

The Product Rule

If $y = f(x)g(x)$, then $\frac{dy}{dx} = f(x)\frac{dg(x)}{dx} + g(x)\frac{df(x)}{dx}$

In English: $\frac{dy}{dx}$ = (First Function × (Derivative of the Second Function) + Second Function × Derivative of the First Function)

Example

Suppose $y = x^2 \ln(x)$

Function 1: x^2

Function 2: $\ln(x)$

Derivative of Function 1: $2x$

Derivative of Function 2: $\frac{1}{x}$

$\frac{dy}{dx} = \left(x^2 + \frac{1}{x}\right) + (\ln(x) \times 2x) = \frac{x^2}{x} + 2x\ln(x) = 3x[1 + 2\ln(x)]$

2.2.5 The Chain Rule

- Suppose you want to take the derivative of a function inside of another function: $y = (x^3 + x^2)^2$
- To find this derivative, we will need to use the chain rule

The Chain Rule

$$\text{If } y = f(g(x)), \text{ then } \frac{dy}{dx} = \frac{dg(x)}{dx} \times \frac{df(x)}{dx}$$

In English: $\frac{dy}{dx} =$ (Derivative of the Inside Function) $\times$ (Derivative of the Outside Function while Ignoring the Inside Function)

Example 1

Suppose $y = (x^3 + x^2)^2$

Step 1: Define the inside function: $g(x) = x^3 + x^2$

Step 2: Take the derivative of the inside function: $\frac{dg(x)}{dx} = 3x^2 + 2x$

Step 3: Define the outside function by replacing the inside function with x: $y = (z)^2$ where $z = x^3 + x^2$

Step 4: Take the derivative of the outside function as defined in Step 3: $2z$

Step 5: Apply the chain rule: $(3x^2 + 2x) \times 2z$

Step 6: Substitute in for z: $(3x^2 + 2x) \times 2(x^3 + x^2)$

Step 7: Simplify: $2x^3(3x^2 + 5x + 2)$

Example 2

Suppose $y = \ln(x^3)$

Step 1: Define the inside function: $g(x) = x^3$

Step 2: Take the derivative of the inside function: $\frac{dg(x)}{dx} = 3x^2$

Step 3: Define the outside function by replacing the inside function with x: $y = \ln(z)$ where $z = x^3$

Step 4: Take the derivative of the outside function as defined in Step 3: $\frac{1}{z}$

Step 5: Apply the chain rule: $\left(\frac{1}{z}\right) \times 3x^2$

Step 6: Substitute in for z: $\left(\frac{1}{x^3}\right) \times 3x^2$

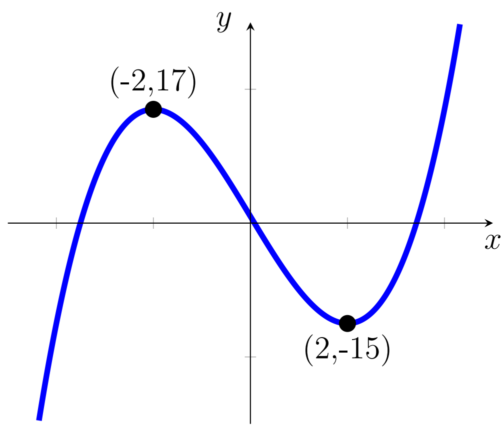
Step 7: Simplify: $\frac{3x^2}{x^3} = \frac{3}{x}$

COMPLETE PRACTICE WORKSHEET

2.2.6 Finding Minima and Maxima using Derivatives

- Suppose you have a the following function: $y = x^3 - 12x + 1$
- Graphically it looks as follows:

Figure 2.1: Finding Minima and Maxima



- We can see from the graph that this function has a maxima at the point (-2,17) and a minima at the point (2,-15)
- We can find these points using derivatives
- Recall, that a derivative tells us the slope at any given point

- At a minima or maxima, the slope at the point will be zero
- We can find the x value of all the minima and maxima of a function by setting the derivative equal to zero, and solving for x
- These points are called the critical points

Example

If $y = x^3 - 12x + 1$, then $\frac{dy}{dx} = 3x^2 - 12$

Set $\frac{dy}{dx} = 0$, and solve for x:

$$3x^2 - 12 = 0$$
$$3x^2 = 12$$
$$x^2 = 4$$
$$x = \sqrt{4}$$
$$x = 2 \text{ and } x = -2$$

- By setting the first derivative equal to zero, we can find all of the critical points, but this does not tell us if a critical point is a minima or a maxima

First Derivative Test

- To determine if a critical point is a minima or a maxima, we can use the First Derivative Test
    - We know that each critical point is an "inflection" point, where the function changes from increasing to decreasing (a maxima) or decreasing to increasing (a minima)
    - To conduct the first derivative test, we are going to split the function into segments before and after each critical point
    - In each one of these segments, we are going to determine if the function is increasing or decreasing
        - If $\frac{dy}{dx} > 0$, then the function is increasing
        - If $\frac{dy}{dx} < 0$, then the function is decreasing

Example

We found that for the function $y = x^3 - 12x + 1$, it has two critical points at $x = 2$ and $x = -2$.

Therefore, there are three segments to this function: $x = -\infty$ to $x = -2$, $x = -2$ to $x = 2$, and $x = 2$ to $x = \infty$

Pick a point in each segment, and plug it into $\frac{dy}{dx}$ to determine the sign:

Segment	**Point**	**Derivative**
$x = -\infty$ to $x = -2$	-3	$\frac{dy}{dx} = 3(-3)^2 - 12 = 15$
$x = -2$ to $x = 2$	0	$\frac{dy}{dx} = 3(0)^2 - 12 = -12$
$x = 2$ to $x = \infty$	3	$\frac{dy}{dx} = 3(3)^2 - 12 = 15$

Create the following table, with one row for each segment:

Segment	**Sign of** $\frac{dy}{dx}$	**Increasing/Decreasing**
$x = -\infty$ to $x = -2$	Positive	Increasing
$x = -2$ to $x = 2$	Negative	Decreasing
$x = 2$ to $x = \infty$	Positive	Increasing

Because the function switches from increasing to decreasing when $x = -2$, this means that critical point is a maximum.

Because the function switches from decreasing to increasing when $x = 2$, this means that critical point is a minimum.

Second Derivative Test

- Another method to determine if a critical point is a minima or a maxima is the Second Derivative Test
- The second derivative of a function is the derivative of the derivative
- The second derivative uses the following notation: $\frac{d^2y}{dx^2}$
- We can evaluate the critical points using the second derivative to determine if there is a minima or maxima

- If $\frac{d^2y}{dx^2} > 0$, then the critical point is a minimum
- If $\frac{d^2y}{dx^2} < 0$, then the critical point is a maximum
- If $\frac{d^2y}{dx^2} = 0$ or if $\frac{d^2y}{dx}$, then the second derivative test is inconclusive
  - Use the first derivative test in this instance

Example

Using our previous example $y = x^3 - 12x + 1$, we know that $\frac{dy}{dx} = 3x^2 - 12$ and that there are two critical points: $x = 2$ and $x = -2$.

The second derivative is the derivative of the derivative: $\frac{d^2y}{dx^2} = 6x$.

We can evaluate $\frac{d^2y}{dx^2}$ at each critical point:

When $x = 2$, $\frac{d^2y}{dx} = 6(2) = 12$. $\frac{d^2y}{dx^2} > 0$, therefore there is a minimum at $x = 2$.

When $x == 2$, $\frac{d^2y}{dx} = 6(= 2) == 12$. $\frac{d^2y}{dx^2} < 0$, therefore there is a maximum at $x = 2$.

2.2.7 Functions with Multiple Variables – Partial Derivatives

- As mentioned earlier, we can view equations as relationships: y changes as x changes
- However, sometimes multiple things can influence change
  - Lets consider the profit of a business
  - Companies use a combination of workers and machines to make their product
  - If we want to determine how much of a product they produce, the more workers a company hires can increase output, but also the more machines a company uses can increase output
  - So in this instance, the quantity of output depends on changes to both workers and machines
- Examples of functions like this might be: $Q = \ln(K) + \ln(L)$ or $Q = L^{1/3}K^{2/3}$
- What we want to do is determine if one variable changes, but we assume the other variable remains the same, what is the change of the variable we are interested in

- To do this, we need to take what is called a partial derivative
- Recall from derivative notation, $\dfrac{\partial[FUNCTION]}{\partial[VARIABLE\,OF\,INTEREST]}$
- If the variable in a function is *not* the variable of interest, we will treat it as if it is a constant number
    - Remember from earlier taking derivatives of a constant

Example

Suppose we have a function: $Q = L^{1/3}K^{2/3}$

K is a variable that represents quantity of capital (machinery), L is for quantity of labor (workers), and Q represents quantity of outpu.t

We want to know how much does output change if the company increases the number of workers, while the number of machines does not change (remains constant) *AND* we want to know how much output changes if the company increases the number of machines but the number of workers remains constant.

First, lets find how Q changes as L changes $\left(\dfrac{\partial Q}{\partial L}\right)$:

- We are going to treat K as if it is a constant
- To do this, define c to be the variable that has K: $c = K^{2/3}$
- Substitute c into Q: $Q = cL^{1/3}$
- Now we have a simple derivative like we did earlier
- $\dfrac{\partial Q}{\partial L} = \left(\dfrac{1}{3} \times c\right) L^{(1/3)-1} = (c)\dfrac{1}{3}L^{-2/3}$
- Substitute c back into Q and simplify: $\dfrac{\partial Q}{\partial L} = \dfrac{1}{3}L^{-2/3}K^{2/3}$

Now, lets find out how Q changes as K changes $\left(\dfrac{\partial Q}{\partial K}\right)$:

- We are going to treat L as if it is a constant
- Define c to be the variable that has L: $c = L^{1/3}$
- Substitute c into Q: $Q = cK^{2/3}$
- Now we have a simple derivative like we did earlier

- $\dfrac{\partial Q}{\partial K} = \left(\dfrac{2}{3} \times c\right) K^{(2/3)-1} = (c)\dfrac{2}{3}K^{-1/3}$
- Substitute c back into Q and simplify: $\dfrac{\partial Q}{\partial K} = \dfrac{2}{3}K^{-1/3}L^{1/3}$

COMPLETE PRACTICE WORKSHEET

Section 3

Supply, Demand, and Elasticity

3.1 Introduction to Mathematical Models

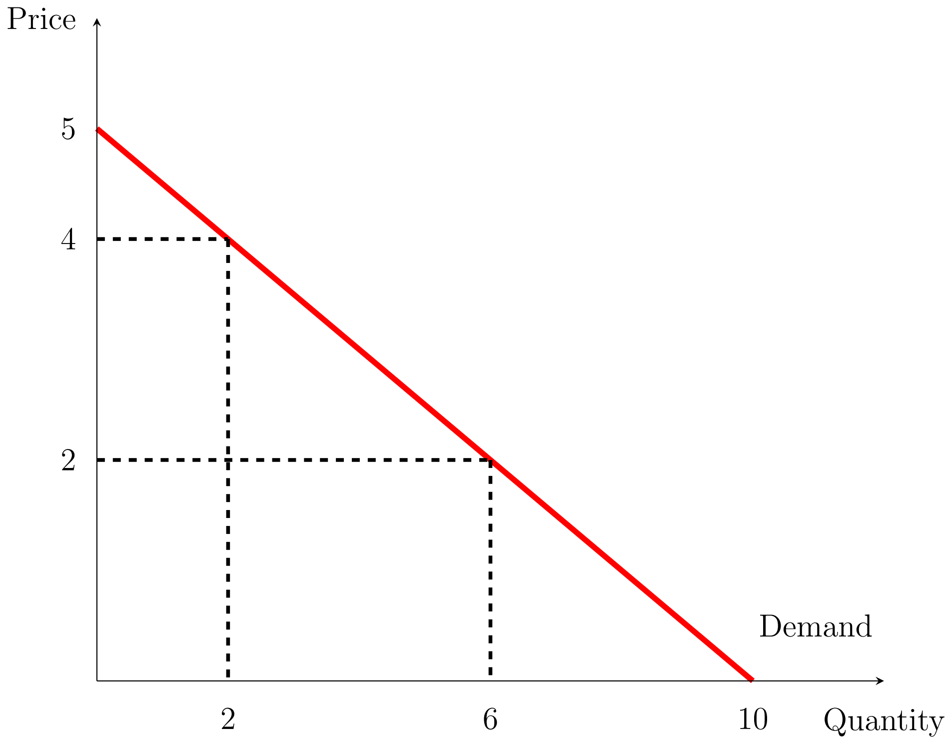
- In economics and business, we study how people and businesses make decisions with:
  - a finite amount of money
  - a finite amount of time
  - an incomplete set of information
- It is not possible to observe and study every single person, so we use mathematical models to approximate the behavior we observe
- We explore if these models are correct or not by collecting data and using statistics (which we will cover later)
- If the data shows flaws in the models, economists seek explanations for why the models are wrong and look to build extensions to the models
- In this class, we are going to start with the most basic models you will see in principles and intermediate microeconomics
- We will explore these models in a graphical nature and show the algebraic and calculus derivations that explain what we see visually

3.2 Markets – Supply and Demand

- The first models we are going to cover are supply, demand, and market equilibrium
- Together, these concepts determine markets for goods and services we buy
- These models will allow us to determine how prices are determined and the impact of government intervention (e.g., taxes, subsidies, and regulation) have on prices and output

3.2.1 Demand

- Demand is the relationship between the price of a good or service and how much of that good or service is purchased (i.e., quantity demanded)
- The Law of Demand states that as the price of a good or service increases, then the quantity demanded will decrease, and vice versa



Example

$$Q_D = 10 - 2P$$

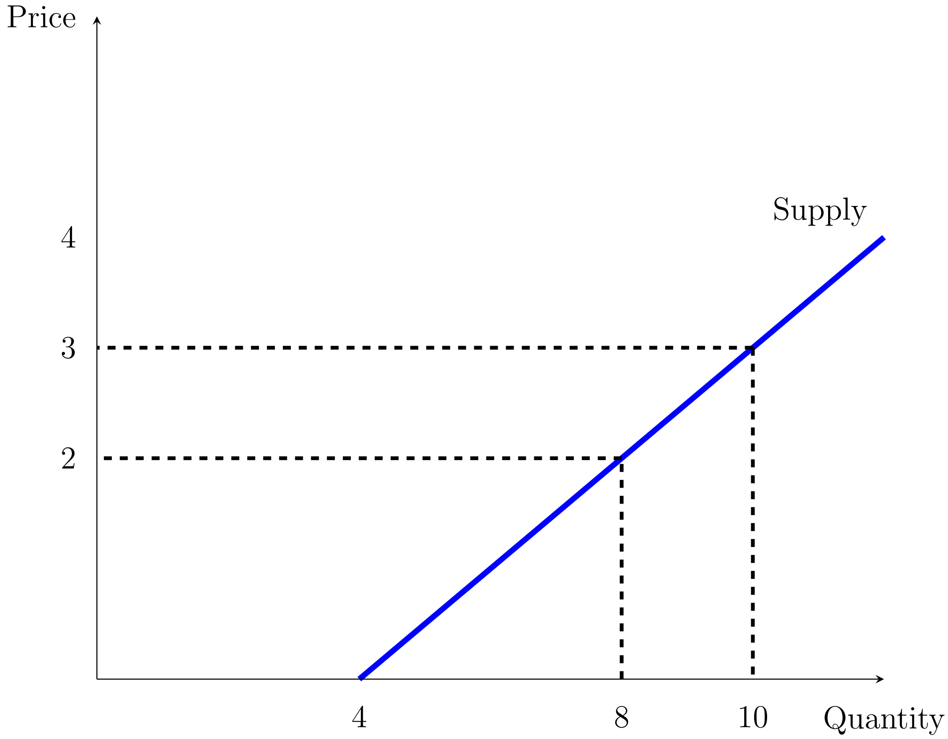
If $P = 2$, then $Q_D = 10 - 2(2) = 6$

If $P = 4$, then $Q_D = 10 - 2(4) = 2$

- If we want to use the demand equation to show it graphically, we need to rearrange it:
  - $Q_D = 10 - 2P$
  - $2P = 10 - Q_D$
  - $P = 5 - \frac{1}{2}Q_D$
- This rearranged equation is called inverse demand function
- This is because Price is on the Y-Axis and Quantity Demanded is on the X-Axis

3.2.2 Supply

- Supply is the relationship between the price of a good or service and how much of that good or service is produced by companies (i.e., quantity supplied)
- The Law of Supply states that as the price of a good or service increases, then the quantity supplied will increase



Example

$$Q_S = 4 + 2P$$

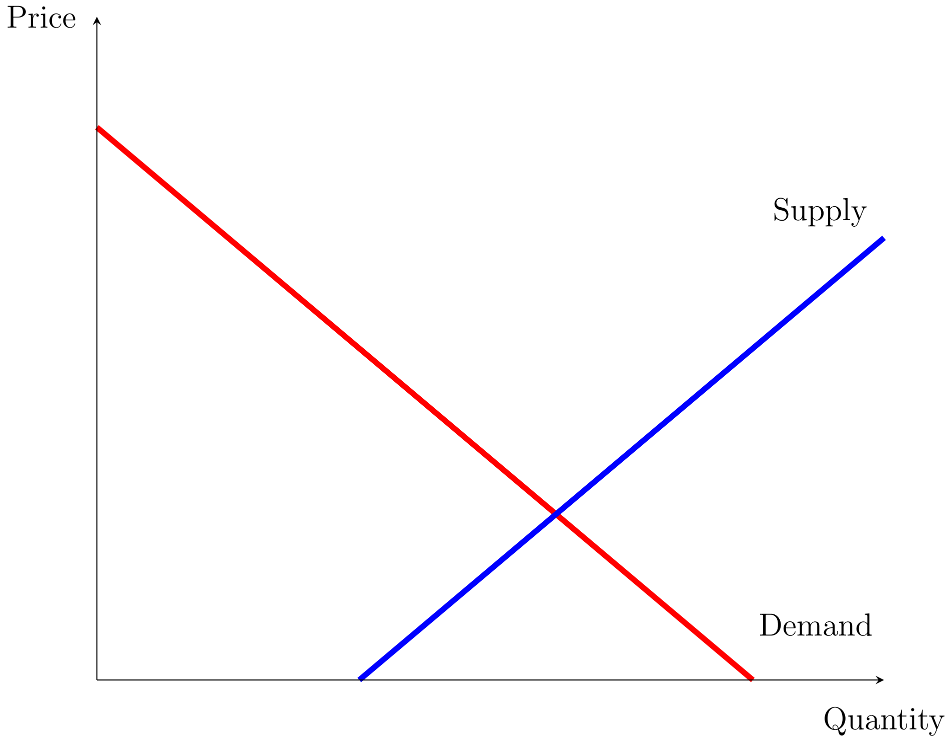
If $P = 2$, then $Q_S = 4 + 2(2) = 8$

If $P = 4$, then $Q_S = 4 + 2(4) = 12$

- If we want to use the supply equation to show it graphically, we need to rearrange it the same way we did with demand, which would be called the inverse supply function

3.2.3 Defining a Market

- A market the means by which the exchange of goods and services takes place as a result of buyers and sellers being in contact with one another, either directly or through mediating agents or institutions
- Essentially, a market is a place where things are bought and sold
- A market is characterized by a system of equations for Q_S and Q_D
- From our example, we have:
  - $Q_D = 10 - 2P$
  - $Q_S = 4 + 2P$
- Graphically:



- A market is considered to be in equilibrium when market supply and demand balance each other, and as a result prices become stable
  - When the market is in equilibrium, $Q_D = Q_S$
- The solution to the system of equations gives us the resulting stable price and the quantity of that good or service that exists in the market

- To find the price and quantity that defines the equilibrium of the market, we must solve the system of equations:

Example

$$Q_D = 10 - 2P$$
$$Q_S = 4 + 2P$$

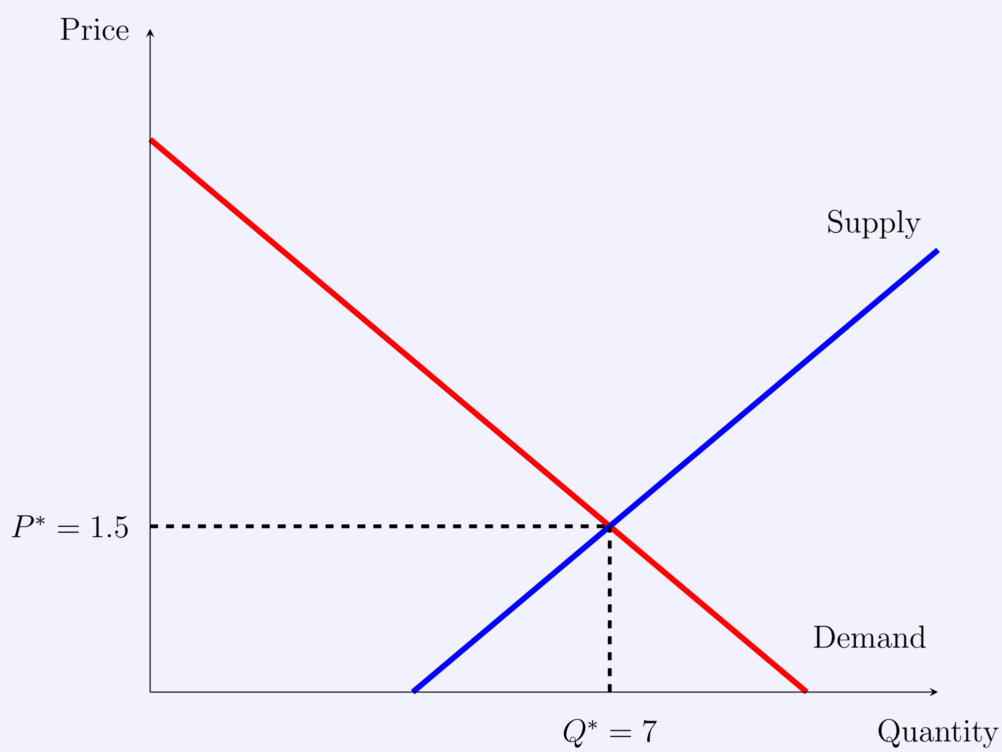
Set $Q_D = Q_S$, solve for the equilibrium price, P^*:

$$4 + 2P = 10 - 2P$$
$$4P = 6$$
$$P^* = \frac{6}{4} = 1.5$$

Substitute P^* into Q_S or Q_D to solve for the equilibrium quantity, Q^* (you will get the same answer regardless of which equation your plug it in to:

$$Q_D - 2(1.5) = 7$$
$$Q_S = 4 + 2(1.5) = 7$$
$$Q^* = 7$$

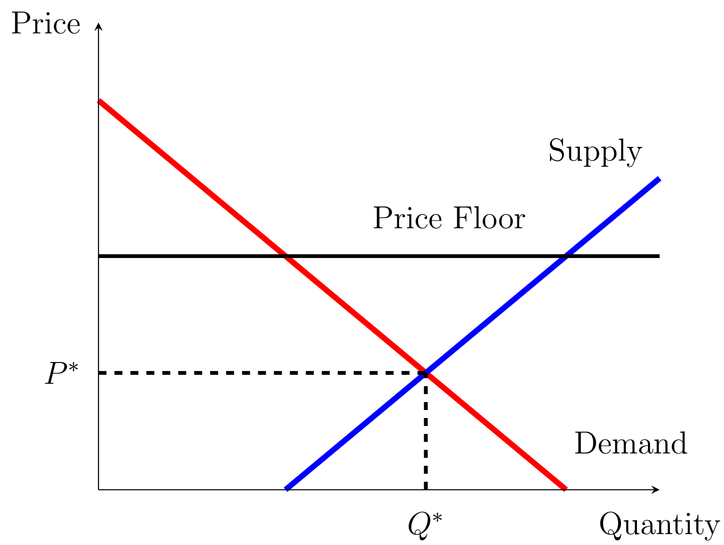
If we look at the graph, we can see visually that this is the point where the supply and demand intersect



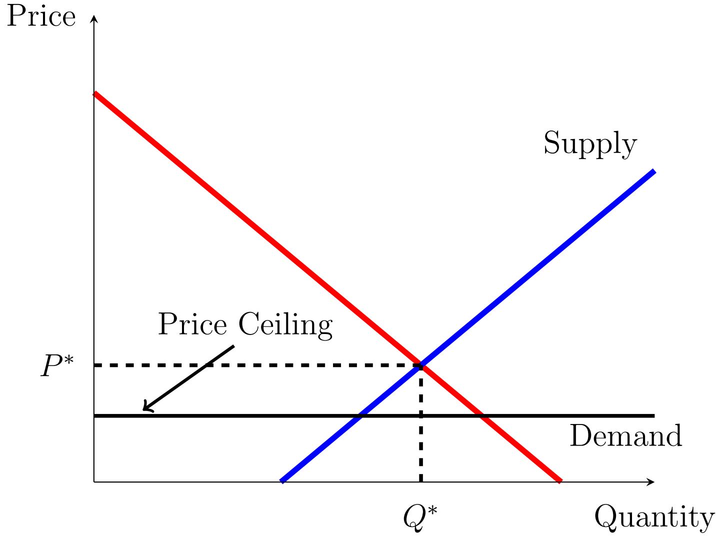
- If $Q_D \neq Q_S$, then the market is not in equilibrium
- There are two ways this can happen:
  1. A market that is influenced by government policy
     - Sales taxes
     - Subsidies
     - Price controls
  2. If the market has an externality
     - An externality is when a company is operating on the wrong supply or demand curve because they are not considering the benefits or harm to society as the equilibrium price and quantity are determined
       - A positive externality is when you consume something, you benefit from it, but so does everyone else around you (e.g., getting a vaccine)
       - A negative externality is when a company produces something and harm can be done to people who live in the vicinity (e.g., pollution)

3.2.4 Government Intervention

- There are four main ways in which the government can intervene in a market:
  1. Price Floor
     - A price floor is where the government believes the market price, P^*, is too low
     - A price floor is a type of price control that sets the minimum price a company can charge
     - The best example of a price floor in practice is minimum wage



2. Price Ceiling
   - A price ceiling is where the government believes the market price, P^*, is too high
   - A price ceiling is a type of price control that sets the maximum price a company can charge



3. Sales Taxes
   - Sales taxes can either be a tax that is levied on the total amount of a good or service that is consumed *or* as a percentage amount
   - Sales taxes can be used to raise revenue for the government or to reduce the quantity of a good or service that is consumed
     - This could be because the good is harmful to you (e.g., cigarettes)
     - This could be because the production of the good is harmful to society and used to correct an externality (e.g., reduce pollution)
4. Subsidy
   - A subsidy is a direct payment or tax credit to a company by the government
   - This is done to stimulate economic activity and can be used to correct for an externality

- For each of these cases, $Q_S \neq Q_D$

Price Controls

- The easiest way to see this is to do an example

Example

Suppose the market for renting an apartment in Lexington is defined by the following:

$$Q_D = 5000 - P$$

$$Q_S = P$$

Find the equilibrium price and quantity:

$$P = 5000 - P$$

$$2P = 5000$$
$$P^* = 2,500$$
$$Q^* = 2,500$$

Suppose the government thinks that $2,500 is too high for the price of rent and institutes rent control (or a price ceiling) stating the maximum price of rent is $2,000. How does this change Q_S and Q_D?

To find out, plug the value of the price ceiling into Q_D and Q_S and compare:

$$Q_D = 5000 - 2000 = 3000$$

$$Q_S = 2000$$

When the market is in equilibrium, $Q_D = Q_S = 2500$. With the price ceiling, $Q_D > Q_S$.

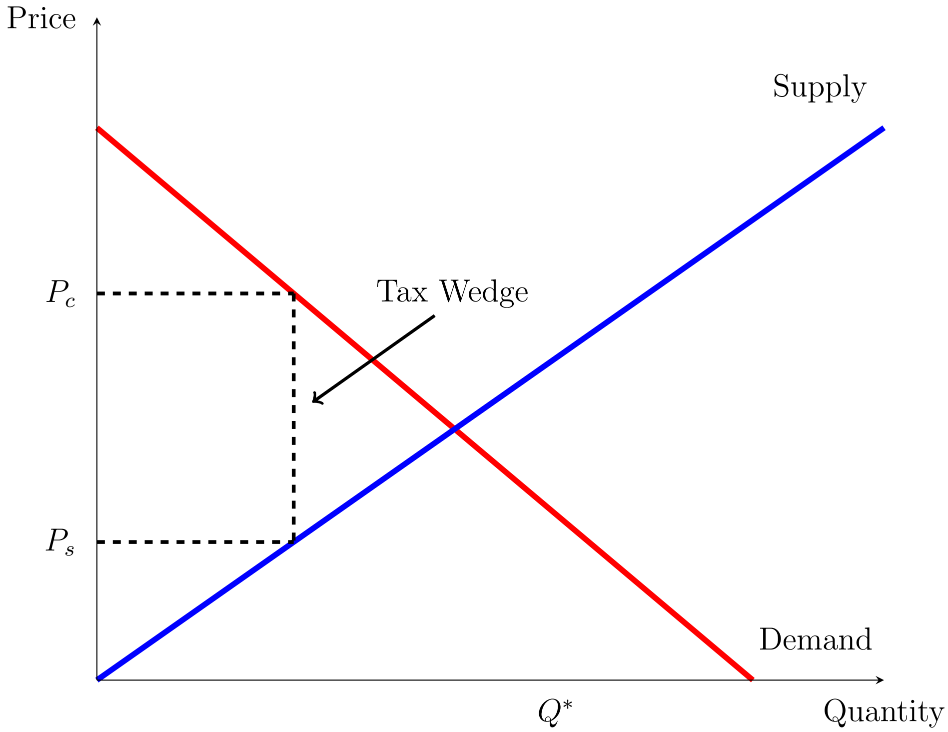
Because the price is lower, more people will want to rent apartments and landlords will be less likely to rent their property. To see the gap in Q_S and Q_D: $Q_S - Q_D = -1000$. This means that there is a <u>shortage</u> of 1,000 apartments.

- This is an example of a price ceiling
- On your homework, you will have a question with a price floor where you will see the opposite effect
  - $Q_D < Q_S$ and $Q_S - Q_D$ will be positive
  - In this instance, you will see that there is a <u>surplus</u>

Taxation

- We have all been to the store to find out that the final price that we pay for something is not the price we saw on the shelf

- This is due to a sales tax
- When the government institutes a sales tax, it puts a “wedge” between supply and demand
- This “wedge” is equal to the value of the tax.



- Based on the supply and demand for each market, the tax wedge determines the price that consumers pay at the store, P_c
- Consumers pay P_c to the store, then the store pays the tax, t, to the government and they are left with P_s
- So when a sales tax is levied, the price you (consumers) pay is different than the price companies receive

- In practice, sales taxes can be applied two ways
  1. A quantity tax is a tax levied on the amount of a good consumed
     - We can relate P_c to P_s using a simple equation:
     - $P_c = P_s + t$
  2. A value tax is a tax levied on the expenditure as a good as a percent of the price
     - The tax, t, is expressed as a percent (τ) of P_s: $t = \tau P_s$
     - So, $P_c = P_s + \tau P_s$
     - $P_c = (1 + \tau)P_s$

Example

Suppose the market for ice cream is defined by:

$$Q_D = 10 - 2P$$

$$Q_S = 1 + \frac{1}{2}P$$

First, lets find the equilibrium price, P^*

$$10 - 2P = 1 + \frac{1}{2}P$$

$$9 = 2.5P$$

$$P^* = 3.60$$

The state of Virginia imposes a 3% (τ) tax on food. Find P_c and P_s.

Step 1: Set up the tax equation relating P_c to P_s

$$P_c = (1 + .03)P_s$$

Step 2: We want to solve this system of equations where there is only one price. We use P_c in the demand equation and P_s in the supply equation (*Note*: In equilibrium, $P_c = P_s$).

Substitute the tax equation into the demand equation to get one equation in terms of P_s:

$$10 - 2[(1 + .03)P_s] = 1 + \frac{1}{2}P_s$$

$$10 - 2.06P_s = 1 + \frac{1}{2}P_s$$

$$9 = 2.56P_s$$

$$P_s = 3.5156$$

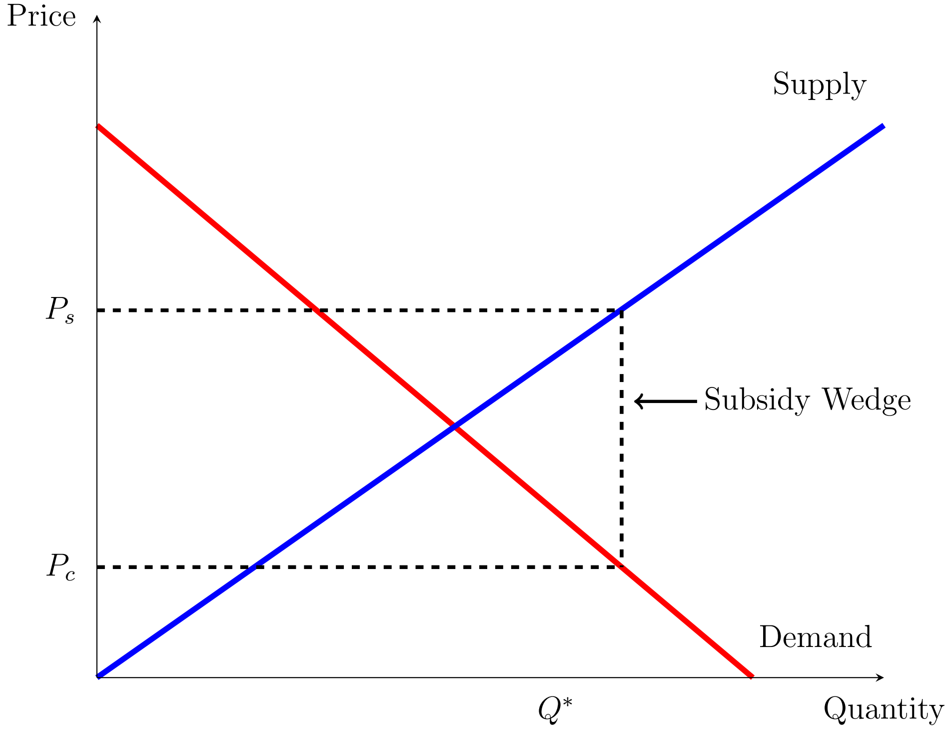
To get P_c, plug P_s into the tax equation:

$$P_c = (1.03)(3.5156) = 3.62109$$

You will notice that $P_c > P_s$ and that P^* falls in between the two, just like we see graphically.

Government Subsidies

- A subsidy is a payment by the government to a company to incentivize production
- A subsidy works similar to a tax, except that with a subsidy, the company receives a higher price than the consumer pays (the subsidy reduces the amount the consumer pays by s dollars per unit)



Example

Suppose we have the same market for ice cream, but the government decides that it wants to encourage more people to eat ice cream by lowering the price by providing a \$1 subsidy per unit. Find P_c and P_s.

Step 1: Set up the subsidy equation relating P_c to P_s

$$P_c = P_s - 1$$

Step 2: We want to solve this system of equations where there is only one price. We use P_c in the demand equation and P_s in the supply equation (*Note*: In equilibrium, $P_c = P_s$).

Substitute the subsidy equation into the demand equation:

$$10 - 2[P_s - 1] = 1 + \frac{1}{2}P_s$$

$$11 = 2.5P_s$$

$$P_s = 4.40$$

To get P_c, plug P_s into the subsidy equation:

$$P_c = 2 - 1 = 3.40$$

You will notice that $P_c < P_s$ and that P^* falls in between the two, just like we see graphically.

3.2.5 Solving Nonlinear Systems of Equations

- In practice, supply and demand are not always linear relationships
- We can easily solve systems of equations that are linear and quadratic in nature
- Higher powered relationships exist but are more easily solved using computer software
- In order to solve systems of equations that are quadratic in nature, we *may* need to use the quadratic formula if the relationship takes the following form: $ax^2 + bx + c = 0$

$$\text{Quadratic Formula: } \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Example

$$Q_D = 5 - P^2$$

$$Q_S = P^2 + 2P$$

Solve for P^*:

$$5 - P^2 = P^2 + 2P$$

$$2P^2 + 2P - 5 = 0$$

Using the quadratic formula:

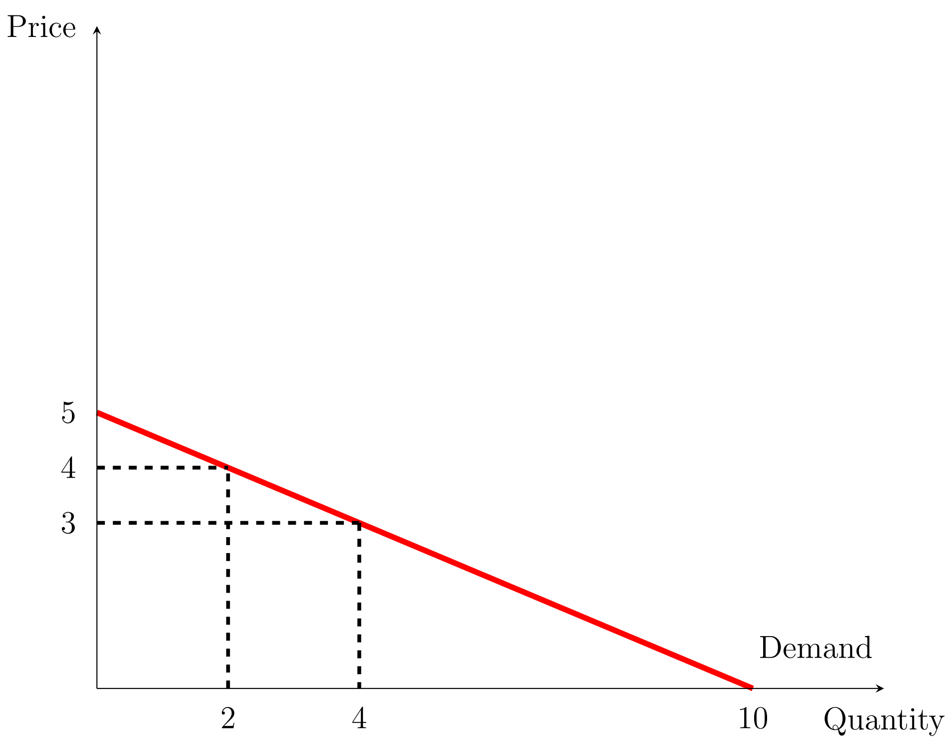
$$\frac{-2 \pm \sqrt{2^2 - 4(2)(-5)}}{2(2)}$$

We get that $P = 1.16$ and $P = -2.16$. It is not possible to have a negative price, so $P^* = 1.16$

- We only need to use the quadratic formula if there is a P^2 and a P in the equation
- If there is a P^2 without a P, then the answer can be solved using a square root

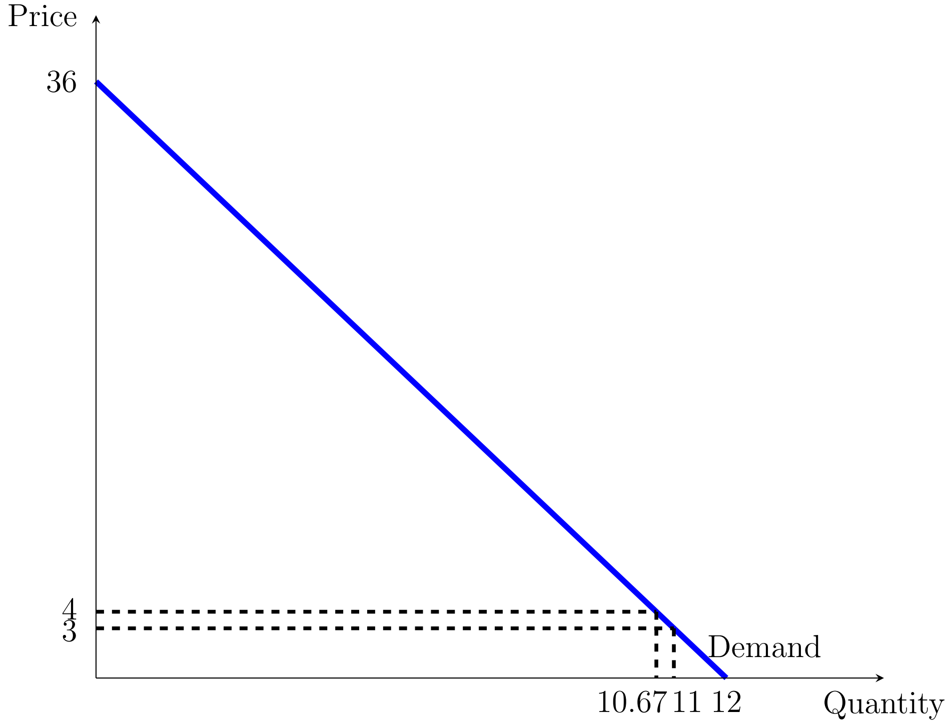
3.3 Elasticity of Demand

- We are often interested in how consumers will respond if a price of a good or service changes
  - A manager may want to know how a price increase or a sale will affect the quantity that is sold
  - The government may be interested in how a sales tax will affect the quantity that is sold
    - How will it affect tax revenue
    - Will it reduce demand if it is trying to disincentivize purchase
- Suppose we use our first example, $Q_D = 10 - 2P$
  - If $P = 4$, then $Q_D = 2$
  - If P decreases to 3, then $Q_D = 4$



- Suppose we instead have the following demand curve: $Q_D = 12 - \frac{1}{3}P$

- If $P = 4$, then $Q_D = 2.67$
- If P decreases to 3, then $Q_D = 3$



- With this demand curve, Q_D is less sensitive to a price change than in the previous example
  - In the first example, a \$1 decrease in price led to a 2 unit increase in Q_D
  - In the second example, a \$1 decrease in price led to a 0.33 unit increase in Q_D
- We can formally measure this for any demand equation using the elasticity of demand
- Elasticity of demand is a measure of how responsive consumers are to a change in the price
- There are three different types of demand curves:
  - Demand curves that are more "flat" are called elastic
    - This means that consumers are sensitive to changes in price
    - These are often luxury goods or goods that have lots of alternatives (substitutes)
  - Demand curves that are more "steep" are called inelastic
    - This means that consumers are not very sensitive to changes in price
    - These are often necessities, addictive products, or good with not alternative

  - If Q_D and P change by the same amount, these demand curve are called unit elastic

- The formula to calculate elasticity is:

$$E_D = \frac{dQ}{dP} \times \frac{P}{Q}$$

- The result you will get will always be a negative number because the demand curve is downward sloping

- We can interpret the elasticity of demand as follows:
  - If $|E_D| > 1$, the good or service is elastic
    - For every 1% increase in P, Q_D decreases by more than 1%
  - If $|E_D| < 1$, the good or service is inelastic
    - For every 1% increase in P, Q_D decreases by less than 1%
  - If $|E_D| = 1$, the good or service is unit elastic
    - For every 1% increase in P, Q_D decreases by exactly 1%

Example

If the demand for paper cups is given by: $Q_D = 10 - 2P$, what is the elasticity of demand if the price of paper cups is $2?

Step 1: Find the derivative:

$$\frac{dQ}{dP} = -2$$

Step 2: We know that $P = 2$. We can plug this into the demand equation to find Q_D:

$$Q_D = 10 - 2(2) = 6$$

Step 3: Calculate the elasticity:

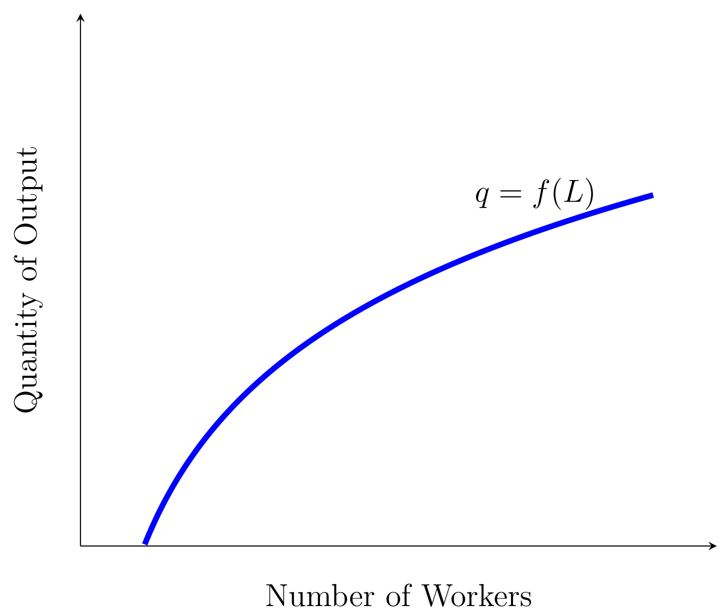
$$E_D = -2 \times \frac{2}{6} = -\frac{2}{3} = -0.67$$

The interpretation of this would be that for every 1% increase in P, Q_D decreases by 0.67%. Because $|E_D| < 1$ when $P = 2$, paper cups are inelastic.

Section 4

The Production Function

- Another model that is commonly used in economics is called the production function
- The production function shows the relationship between the inputs a company uses in production and the quantity of output it produces
    - An input is anything that a company uses to produce goods and services
        - Land
        - Labor
        - Capital
    - Output is the product a company sells
- Here is a graphical example of a simple production function that relates output to the number of workers a company has:



- You will notice the shape of the production function is concave, or that it does not increase until infinity
- Each additional input that is added has a smaller impact on output that the unit before it
- This is called diminishing marginal returns
- This means that a company cannot hire an infinite number of workers and produce an infinite amount of output
- Because of this, companies need to determine what is the appropriate number of workers to hire and capital to use

4.1 Marginal Product

- A marginal product tells us what the additional, or incremental, output each additional worker adds to the total output (q)
- We can calculate the marginal product for each input
- We are going to focus on two inputs: labor (L) and capital (K)

4.1.1 Marginal Product of Labor

- The marginal product of labor (MP_L is the additional output that is added if the company were to hire one additional worker

$$MP_L = \frac{\text{Change in Quantity}}{\text{Change in the Number of Workers}}$$

- It is first helpful to visualize the MP_L through an example:
  - Suppose a pizza company is trying to evaluate how productive each of its employees is and has the following information on its output:

Number of Workers	**Quantity of Pizza**	**Marginal Product of Labor**
0	0	–
1	10	$10 - 0 = 10$
2	15	$15 - 10 = 5$
3	17	$17 - 15 = 2$

- We can formally express the MP_L as follows:
  - If $q = f(K, L)$ then $MP_L = \frac{dq}{dL}$

Example

Lay's produces potato chips using workers and heavy machinery. It's output per day is defined by the following production function:

$$q = K^{1/2}L^{1/2}$$

Find the marginal product of labor.

$$q = K^{1/2}L^{1/2}$$

$$MP_L = \frac{dq}{dL} = \frac{1}{2}L^{-1/2}K^{1/2}$$

$$MP_L = \frac{1}{2}\left(\frac{K}{L}\right)^{1/2}$$

- We can use the MP_L to determine how many workers a company should hire if we know the current level of capital that is not changing and the number of workers a company is considering hiring

4.1.2 Marginal Revenue Product of Labor

- Each unit of output that is produced is sold at the market price
- We can build on the MP_L by not just looking at the amount of output each worker generates for a company, but the amount of revenue it generates as well
- We do this by calculating the Marginal Revenue Product of Labor (MRP_L):

$$MRP_L = MP_L \times Price$$

Example

Continuing with the previous example, if the price of a bag of Lay's Potato Chips is $5, then:

$$MRP_L = \frac{1}{2}\left(\frac{K}{L}\right)^{1/2} \times 5$$

$$MRP_L = 2.5\left(\frac{K}{L}\right)^{1/2}$$

- If we have the number of units of capital and the number of workers, we can calculate the MRP_L for the last workers hired

4.1.3 Marginal Profit

- While all companies want to generate as much revenue as possible, what they are ultimately interested in is profit
- Profit = Revenue - Costs
- In this instance, the cost of hiring a worker is the wage they must be paid
- The marginal profit shows how much profit each worker adds to the company

$$Marginal\ Profit = MRP_L - Wage$$

- If $Marginal\ Profit > 0$, then the company can hire that worker
- If the $Marginal\ Profit < 0$, that means the last worker hired would cost more to pay than they would generate in revenue and the company should not hire them
- A company should hire the number of workers up until $Marginal\ Profit = 0$
- When $Marginal\ Profit = 0$, then $MRP_L = wage$
- We can use this relationship to determine the number of workers a company should hire

Example

Continuing with your example, suppose Lay's uses 1,000 units of capital and pays its employees $50 per day, how many workers should Lay's hire?

$$2.5\left(\frac{1000}{L}\right)^{1/2} = 50$$

$$\left(\frac{1000}{L}\right)^{1/2} = 20$$

$$\frac{1000}{L} = 400$$

$$L^* = \frac{1000}{400} = 2.5$$

4.1.4 Marginal Product of Capital

- Just like with labor, a company needs to know how many units of capital it should use
- We can follow the same steps as we did with labor:
  - Marginal Product of Capital: $MP_K = \frac{dq}{dL}$
  - Marginal Revenue Produce of Capital: $MRP_K = MR_K \times price$
  - Marginal Profit of Capital: $Marginal\ Profit = MRP_K - rental\ price$
    - We will assume that companies must pay rent to use capital
    - If we get into ownership, we have to deal with other costs such as maintenance and depreciation

Section 5

Profit Maximization

5.1 Profit

- In the previous section, we discussed the relationship between inputs (K and L) used in production and the output (q) of a firm using the Production Function
- Each product a firm produces will have its own unique production function
- The firm will then sell the products it produces for a certain price, which will generate revenue for the company
  - Revenue is defined as $price \times quantity$, or pq
- The firm must also spend money on inputs used to produce these products (costs)
  - Firms have two costs: labor and capital
  - The cost of labor is the $wage \times number\ of\ workers$, or wL
  - The cost of capital is the $rental\ price \times units\ of\ capital$, or rK
  - Therefore, the cost of a firm is defined as $wL + rK$
- Therefore, we can define the profit, π, of a firm as the sum of its revenue minus the sum of its costs:

$$\pi = pq - (wL + rK) \tag{5.1}$$

Where,

p is the price of the product

q is the output of the firm

L is the units of labor (number of workers)

w is the wage rate

K is the units of capital

r is the rental rate of capital

- We know from the previous section that the quantity of goods a firm produces can be obtained from the firms production function: $q = f(K, L)$
- Therefore, we can write revenue as $p \cdot f(K, L)$
- Therefore, we can rewrite Equation 5.1 as:

$$\pi = pf(K, L) - wL - rK \tag{5.2}$$

5.2 Profit Maximization

- The goal of any firm is to generate the most profit possible
- To do that, a firm needs to find the optimal level of inputs, L^* and K^*
- To find the optimal level of inputs, we need to maximize the profit function with respect to each input by taking partial derivatives
  - Recall from earlier, that we can set the derivative equal to zero to find maximum and minimum points
- The profit maximization problem is set up as follows:

$$\max \pi = pf(K, L) - wL - rK \tag{5.3}$$

Which has the following first derivatives:

$$\frac{\partial \pi}{\partial L} = p\frac{\partial q}{\partial L} - w = 0 \tag{5.4}$$

$$\frac{\partial \pi}{\partial K} = p\frac{\partial q}{\partial K} - r = 0 \tag{5.5}$$

- Recall that $\frac{\partial q}{\partial L}$ is the Marginal Product of Labor and $\frac{\partial q}{\partial K}$ is the Marginal Product of Capital
- In order to find the profit maximizing quantities of K and L, we need to find the Marginal Rate of Technical Substitution (MRTS)
  - The MRTS is the amount by which capital and labor can be substituted between each other such that output does not change

$$MRTS = \frac{MP_L}{MP_K} = \frac{\partial q/\partial L}{\partial q/\partial K} \tag{5.6}$$

- We then set the MRTS equal to the ratio of the factor prices:

$$\frac{MP_L}{MP_K} = \frac{w}{r} \tag{5.7}$$

- From here, we rearrange this equation for either K or L
- To solve for the profit maximizing quantity of labor, L^*, substitute K from the above step into $\frac{\partial q}{\partial L}$ and solve
- To solve for K^*, substitute L into $\frac{\partial q}{\partial K}$ and solve
- Once you have K^* or L^*, you can find the other by either substituting it into $MRTS = \frac{w}{r}$ or doing both the previous steps.
- To ensure that K^* or L^* are profit maximizing quantities, we will check to make sure the sign of the second derivatives is negative

Example

Suppose you have the following production function for a firm: $q = 10K^{1/3}L^{1/3}$. This firm pays its workers \$15 per hour and rents its capital at \$15 per hour. It sells is product for \$20. Find the values of K and L that maximize the firms profit.

$$\max \pi = pq - wL - rK$$
$$\text{s.t. } q = f(K, L)$$
$$\max \pi = 200K^{1/3}L^{1/3} - 15L - 15K$$

Find the first derivatives:

$$\frac{\partial \pi}{\partial L} = \frac{200}{3}K^{1/3}L^{-2/3} - 15 = 0$$
$$\frac{\partial \pi}{\partial K} = \frac{200}{3}K^{-2/3}L^{1/3} - 15 = 0$$

Find the Marginal Rate of Technical Substitution:

$$MRTS = \frac{MP_L}{MP_K} = \frac{\frac{1}{3}K^{1/3}L^{-2/3}}{\frac{1}{3}K^{-2/3}L^{1/3}} = \frac{K}{L}$$

Set MRTS equal to the ratio of the factor prices, $\frac{w}{r}$:

$$\frac{K}{L} = \frac{15}{15}$$

$$K = L$$

Substitute this into the first derivatives, solve for L^*:

$$\frac{200}{3}(L)^{1/3}L^{-2/3} = 15$$
$$\frac{200}{45} = L^{1/3}$$
$$L^* = 87.79$$

Plug L* into $K = L$ to solve for K^*:

$$K^* = 87.79$$

Plug K^* and L^* into the production function to find q^*:

$$q^* = 10(87.79)^{1/3}(87.79)^{1/3} = 197.53$$

Plug q^*, K^* and L^* into the profit function to find π^*

$$\pi = 50(197.53) - 15(87.79) - 15(87.79) = 1316.87$$

Check the second derivatives to ensure both are a maxima:

$$\frac{\partial^2\pi}{\partial L^2} = -\frac{400}{9}(87.79)^{1/3}(87.79)^{-5/3} = -0.114$$
$$\frac{\partial^2\pi}{\partial K^2} = -\frac{400}{9}(87.79)^{-5/3}(87.79)^{1/3} = -0.114$$

Since both second derivatives are negative, K^* and L^* are profit maximizing values.

5.3 Marginal Revenue and Marginal Cost Approach

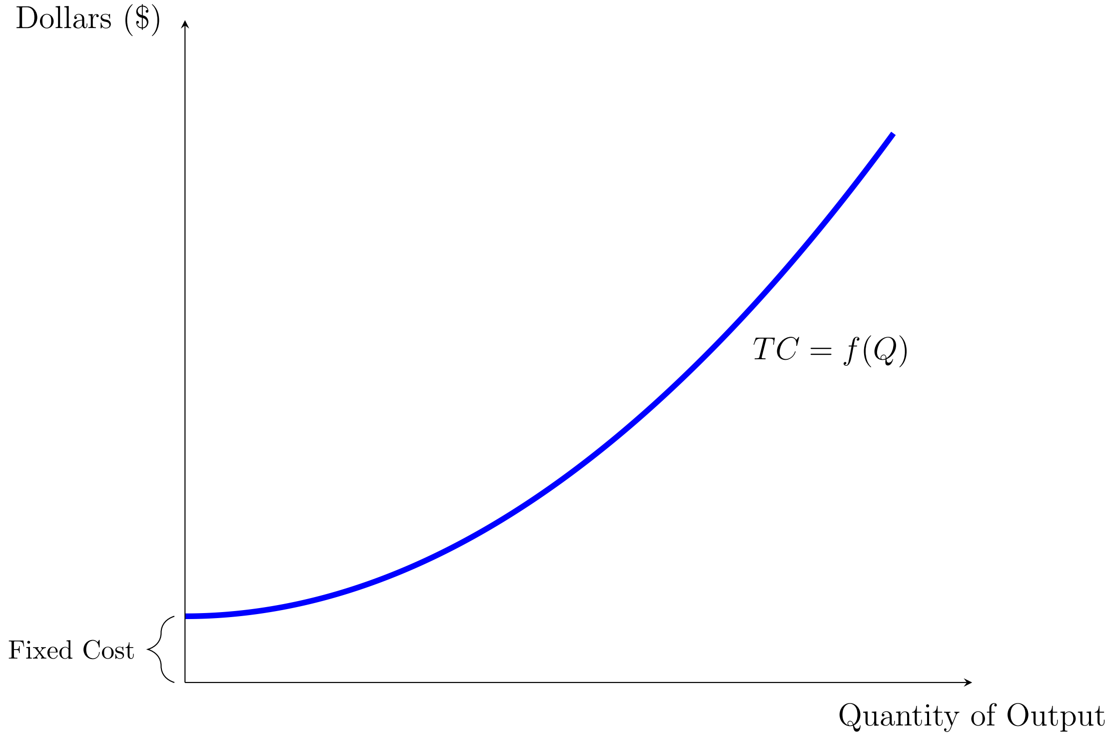
5.3.1 Costs Structure of a Firm

- Firms have two types of costs:

  1. Fixed Costs: Costs that a company has to pay regardless of how much output it produces
     - Rent/mortgage payments
     - Electricity
  2. Variable Costs: Costs that are dependent on the quantity of output
     - Labor costs

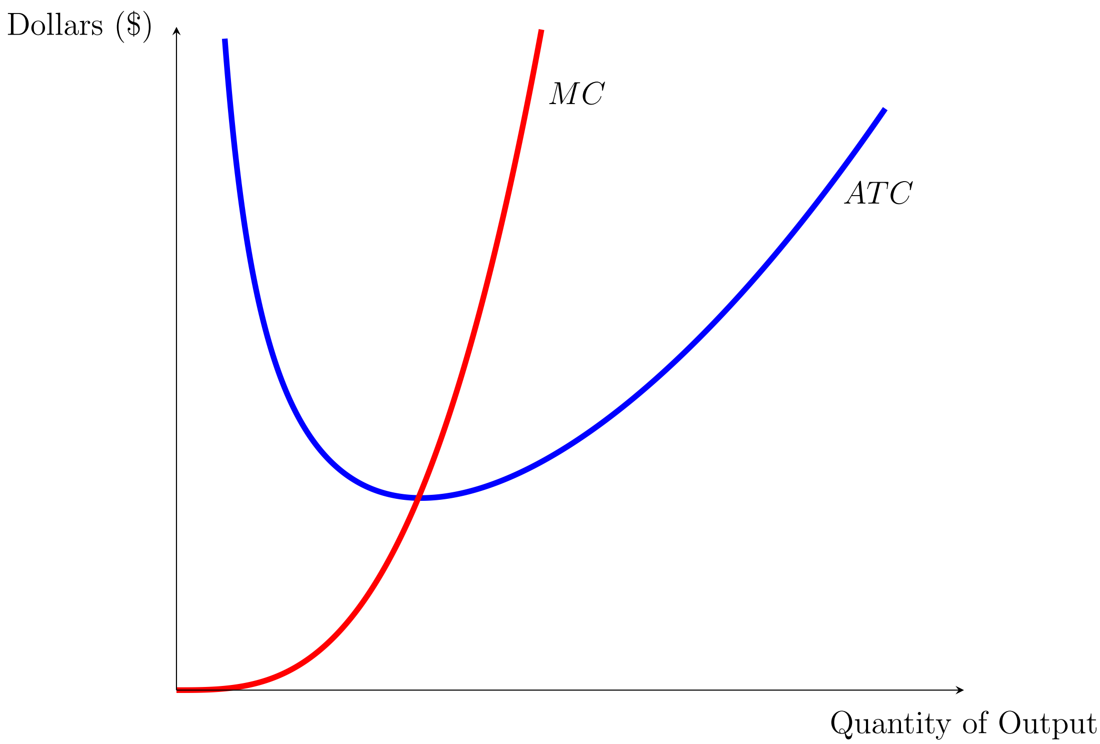
  - Raw materials

- So a firms **Total Cost** = *Fixed Costs* + *Variable Costs*



- An example of a total cost function would look like this: $TC = 50 + 3Q^2$
- There are two other cost metrics that we care about in economics:
  - Average Total Cost: Based on the total number of units produced, what is the average cost of each unit
    - $ATC = \frac{TC}{Q}$
    - From our example: $ATC = \frac{50 + 30Q^2}{Q} = \frac{50}{Q} + 3Q$
  - Marginal Cost: The additional cost incurred by the firm if one extra unit is produced
    - $MC = \frac{dTC}{dQ}$
    - From our example: $MC = 6Q$

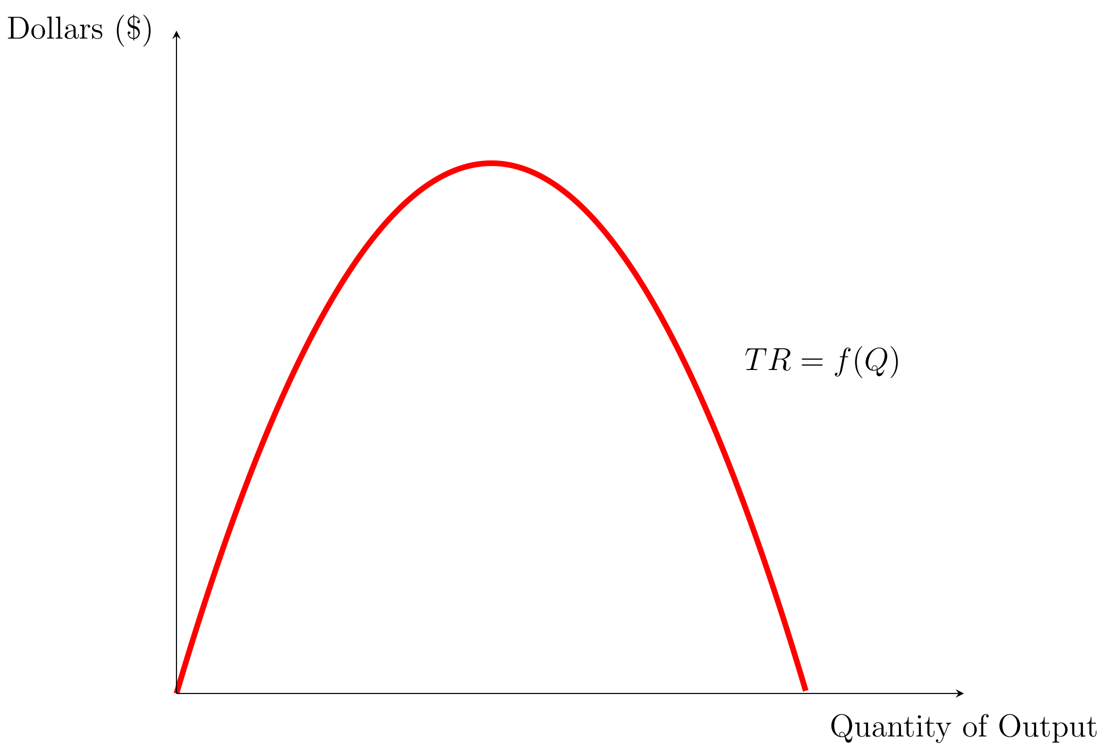
- Graphically, $MC = ATC$ when ATC is at its minimum



5.3.2 Revenue

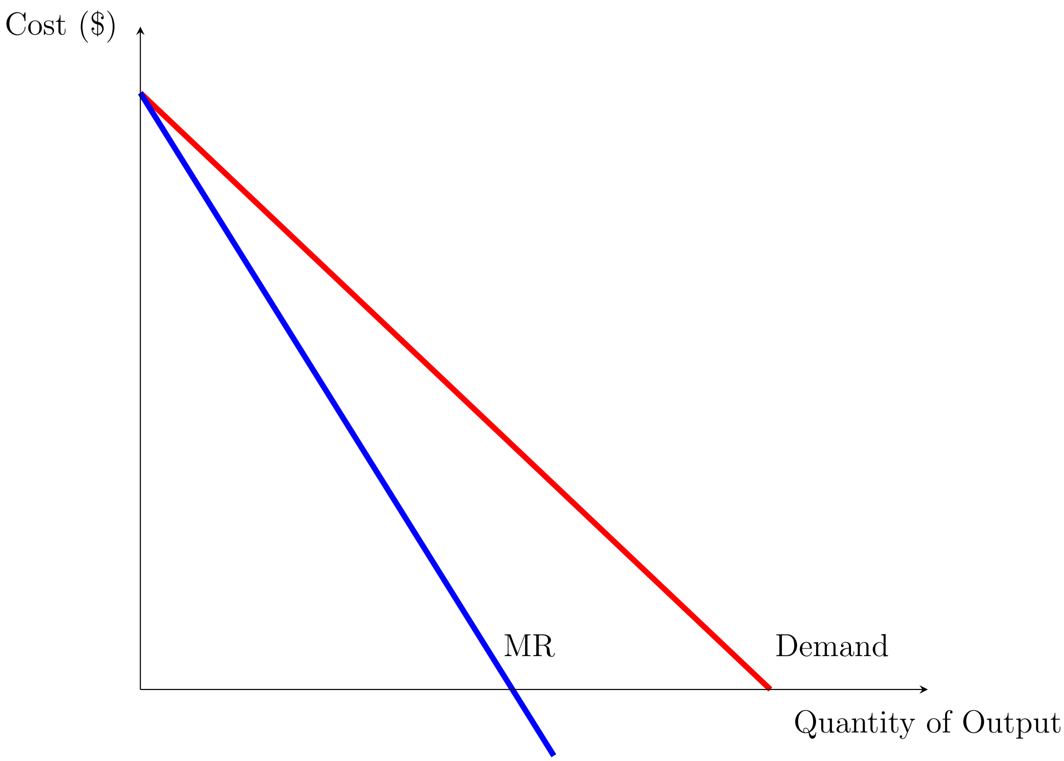
- The revenue a firm generates is determined by the price of the product it sells and the price of that product
- Total Revenue is equal to price times quantity: $TR = P \times Q$
- To calculate TR, we will use the inverse demand function for P

- Example
- $Q_D = 100 - 2P$
- Inverse Demand: $P = 50 - \frac{1}{2}Q_D$
- $TR = \left(50 - \frac{1}{2}Q\right)Q = 50Q - \frac{1}{2}Q^2$



- TR starts to decrease when the price gets high enough that demand becomes inelastic
- TR tells us for the quantity of output that that is produced, what is the revenue that is generated
- Economists are also interested in the Marginal Revenue, or the additional revenue that is generated for each unit of output that is produced
  - $MR = \frac{dTR}{dQ}$
  - From our example: $MR = 50 - Q$

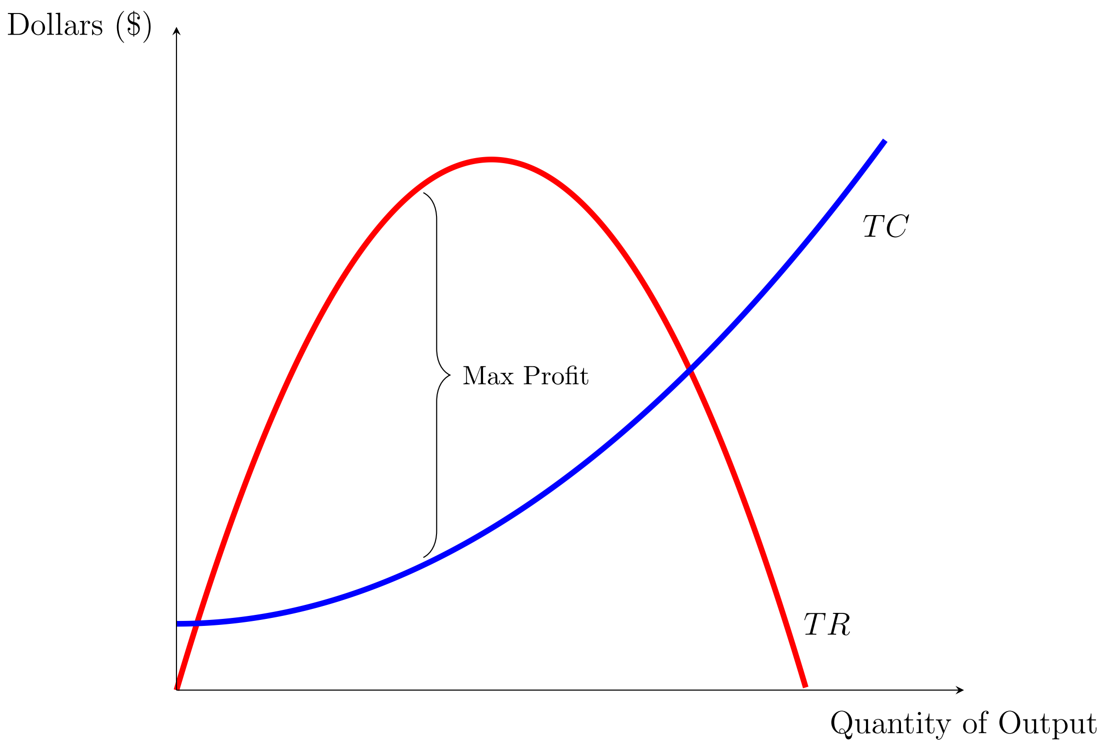
- Graphically, Q_D and MR look as follows:



- MR will be negative after a certain quantity, after TR starts to decrease
- Again, this is a reflection of where the demand for that good or service becomes inelastic

5.3.3 Maximizing Profit

- For any Q, we can determine what the profit a company will make
  - $Profit = TR - TC$
- We can substitute in TR and TC to get an equation for profit
  - $Profit = \left(50Q - \frac{1}{2}Q^2\right) - (50 + 3Q^2) = 50Q - \frac{7}{2}Q^2 - 50$
- The goal of each firm is to maximize its profit
- This occurs when the distance between TR and TC is at its greatest
- We can see that graphically:



- To find this point mathematically, we want to think about marginal profit
- We discussed marginal profit in the previous section
- $Marginal\ Profit = Marginal\ Revenue - Marginal\ Cost$
- If $Marginal\ Profit > 0$, then that unit adds profit to the company and the firm should produce it
- If $Marginal\ Profit < 0$, then that unit costs more to produce than the revenue it generates and the firm should not produce it
- The firm can determine the "last unit" that generates profit, and therefore the optimal quantity it should produce, when $Marginal\ Profit = 0$
- When $Marginal\ Profit = 0$, $MR = MC$
- Therefore, we can find the Q that maximizes profit by setting $MR = MC$ and solving for Q
- From our example:
    - $50 - Q = 6Q$
    - $50 = 7Q$
    - $Q^* = \frac{50}{7} = 7.14$

- To find out the price that maximizes profit, we plug Q^* into Q_D
  - $Q_D = 100 - 2P$
  - $\frac{50}{7} = 100 - 2P$
  - $P^* = \frac{650}{14} = 46.43$
- To find the profit, plug Q^* into the profit function:
  - $Profit = 50Q - \frac{7}{2}Q^2 - 50$
  - $Profit = 50(7.14) - \frac{7}{2}(7.14)^2 - 50$
  - $Profit = 128.57$

Section 6

Utility Maximization

- As we have discussed earlier, economics is the study of how people make choices when faced with constraints
  - Money
  - Time
  - Information
- The information constraint is something that explored in the context of game theory
- However, we can model how people make decisions given that they have a limited amount of time and money through utility maximization

6.1 Utility

- Every person has their own unique set of preferences. These preferences determine how we choose to spend our time and money.
- In general, the more we consume the things we like, the more "happy" we are
- In economics, utility is a measure of the level of "happiness" we get from consuming goods and services
- People generally make choices that make them as happy as possible, therefore, we try to obtain the highest level of utility possible
- Since each person has different preferences, so the level of "happiness" one person gets consuming something may be very different than the level of satisfaction someone else gets
- We can model the relationship between our individual preferences and level of utility with a utility function
- A utility function must have certain properties for it to represent preferences:

1. **Preferences are complete**
   - It can't be true that item A is preferred to item B *and* that you are indifferent between item A and item B, only one can be true
2. **Preferences are transitive**
   - If A preferred to B and B is preferred to C, then A must be preferred to C
   - This says that consumers are consistent in their preferences
3. **Preferences are continuous**
   - If the function is not continuous, we cannot take a derivative and we will have a demand curve that not well-behaved
4. **Preferences are monotonically increasing**
   - This means that we always prefer more utility to less utility

Commonly Used Utility Functions

- **Cobb-Douglass**

$$U = X^{\alpha}Y^{1-\alpha} \qquad \text{Example: } U = X^{1/2}Y^{1/2}$$

- **Constant Elasticity of Substitution (CES)**

$$U = (\alpha X^{\rho} + (1-\alpha)Y^{\rho})^{\frac{1}{\rho}} \qquad \text{Example: } U = \left({}^1\!/\!{}_3X^2 + {}^2\!/\!{}_3Y^2\right)^{\frac{1}{2}}$$

- **Constant Relative Risk Aversion (CRRA)**

$$U = \begin{cases} \dfrac{X_i^{1-\theta}}{1-\theta} & \text{if } \theta > 0,\ \theta \neq 1 \\ \ln(X_i) & \text{if } \theta = 1 \end{cases}$$

- Utility can be shown graphically using indifference curves
  - An indifference curve is a graph showing combination of two goods that give the consumer equal satisfaction and utility
  - A consumer would be indifferent between each point on an indifference curve because and all points give him the same utility
- We can get the equation for an indifference curve by rearranging the utility function
  - If $U = X^{1/2}Y^{1/2}$, then the indifference curve is $Y = \dfrac{U^2}{X}$, where U is the "fixed" level of utility

Example

Suppose a person's preferences are defined by $U = X^{1/2}Y^{1/2}$. In order to graph this function, we need to rearrange it:

$$Y^{1/2} = \frac{U}{X^{1/2}}$$

$$Y = \frac{U^2}{X}$$

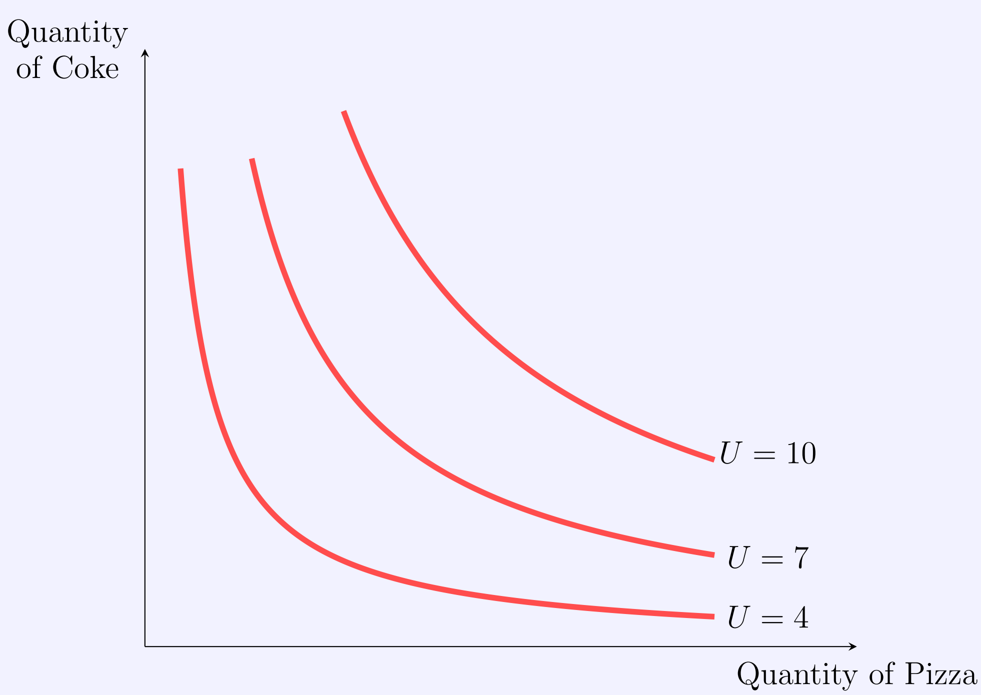
U represents the level of utility. Therefore, this will be a number. For each level of utility, there will be a different indifference curve:

$$\text{If } U = 4, \text{ then } Y = \frac{16}{X}$$

$$\text{If } U = 7, \text{ then } Y = \frac{49}{X}$$

$$\text{If } U = 10, \text{ then } Y = \frac{100}{X}$$

We can graph each of these indifference curves:



- If all of the properties of the utility function are satisfied, then the indifference curves will never intersect and will be parallel to each other
- Since each indifference curve represents a fixed level of utility and preferences are

monotonically increasing, we would prefer to be on a higher indifference curve

- From the indifference curve, we can also see the marginal rate of substitution (MRS)
  - The MRS is the amount of a good that a consumer is willing to give up of one good to consume another good, but keeping utility constant

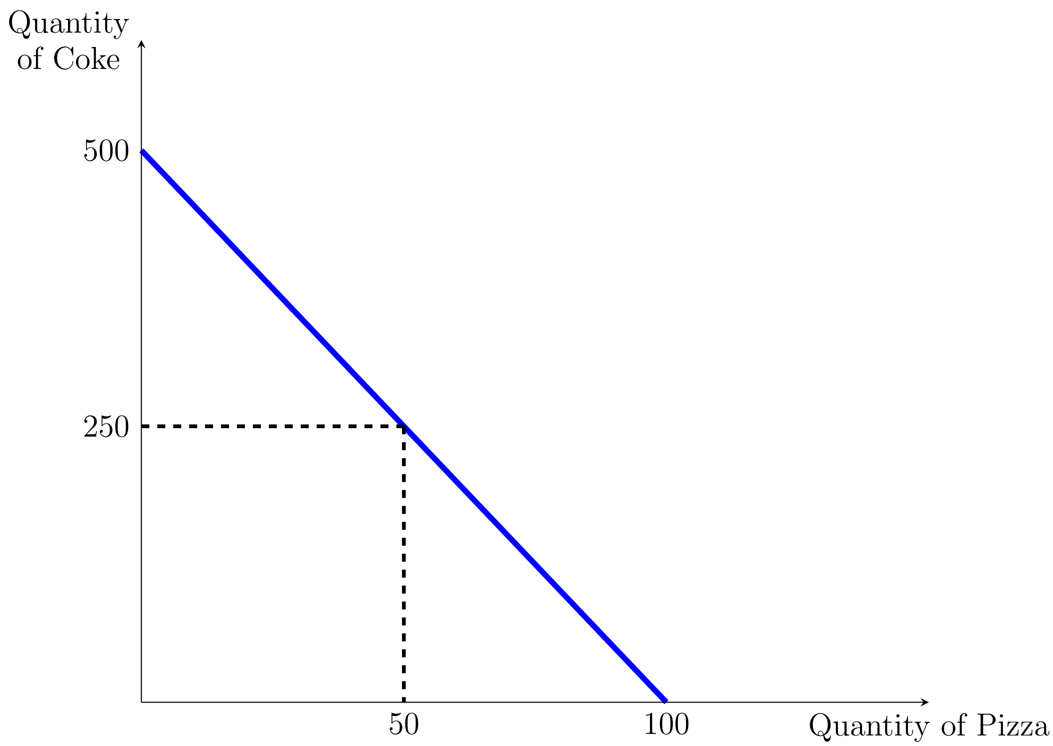
$$MRS = \frac{\text{Marginal Utility of Good X}}{\text{Marginal Utility of Good Y}}$$

$$Marginal\,Utility_x = \frac{dU}{dX} \qquad Marginal\,Utility_y = \frac{dU}{dY}$$

$$MRS = \frac{\left(\frac{dU}{dX}\right)}{\left(\frac{dU}{dY}\right)}$$

6.2 The Budget Constraint

- The budget constraint can show two things:
  - All the combinations of goods and services you can buy with your available income
  - All the combinations of things you can do given the amount of time you have available
- Lets think this through with an example:
  - Suppose you have an income of $1,000
  - You are at the store and want to buy some frozen pizza and bottles of Coke for a party you are going to have
  - The price of pizza is $10 and the price of Coke is $2
    - If you spend all of your money on pizza, then you could purchase $\frac{1000}{10} = 100$ pizzas
    - If you spend all of your money on Coke, then you could purchase $\frac{1000}{2} = 500$ bottles of coke
    - Alternatively, you could buy some pizza and some Coke.
- We can see the budget constraint graphically:



- We can also express this as a mathematical relationship:

  $(Price\, of\, Coke \times Quantity\, of\, Coke) + (Price\, of\, Pizza \times Quantity\, of\, Pizza) = Income$

- The left hand side must equal the right hand side
- This equation will give you every possible combination of Coke and pizza you could buy for the given prices and income
- So from our example:
  - $(\$2 \times Quantity\ of\ Coke) + (\$10 \times Quantity\ of\ Pizza) = \$1000$
- More generally, we can write the budget constraint as follows:

$$P_1X + P_2Y = I$$

  - Where,

    P_1 = Price of good 1

    P_2 = Price of good 2

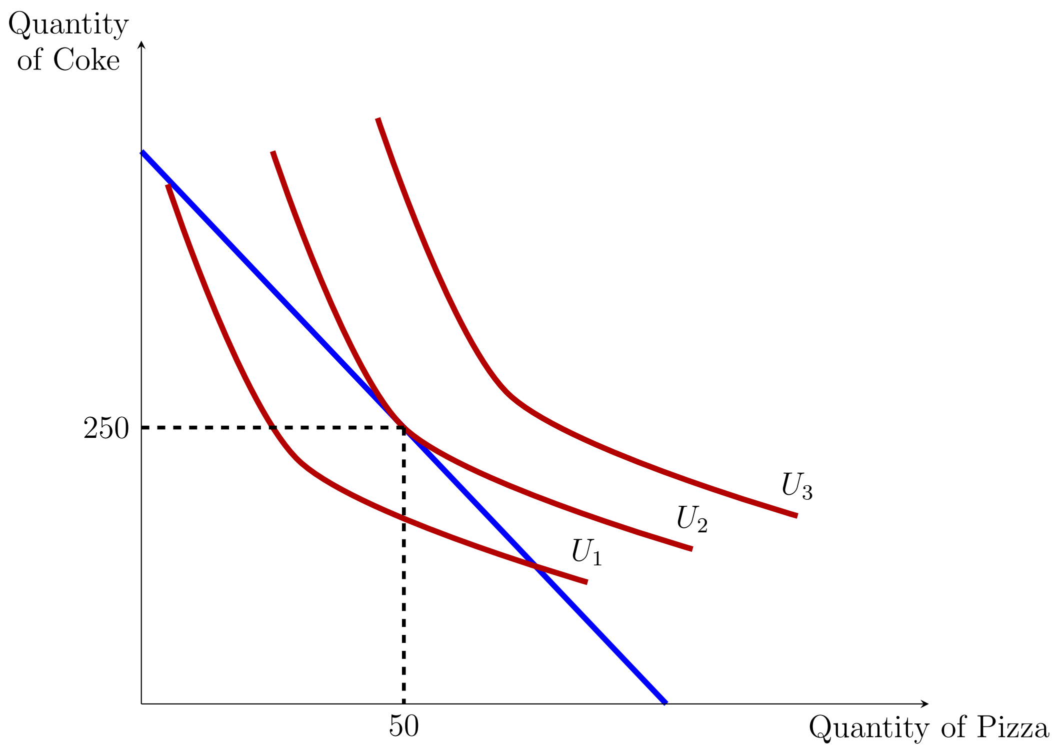
    X = Quantity of good 1

    Y = Quantity of good 2

    I = Income

6.3 Utility Maximization

- Because preferences are monotonically increasing, this means that we always want to have the highest utility possible
- However, we are constrained by how high our utility can go by our budget constraint
- Utility maximization is the process of finding what are the optimal quantities for goods given our individual preferences and our budget constraint



- We can solve quantity of each good that maximizes utility mathematically by the following process:

$$\begin{aligned} \max \quad & U(X,Y) \\ \text{s.t.} \quad & P_1X + P_2Y = I \end{aligned}$$

- Steps to maximize utility:
  - **Step 1**: Find the Marginal Utility for each good
    - $MU_x = \dfrac{dU}{dX}$
    - $MU_y = \dfrac{dU}{dY}$
  - **Step 2**: Find the Marginal Rate of Substitution

  - $MRS = \frac{MU_1}{MI_2} = \frac{dU}{dX} \Big/ \frac{dU}{dY}$
- **Step 3**: Apply the Utility Maximization Rule
  - $MRS = \frac{P_1}{P_2}$
- **Step 4**: Rearrange the Utility Maximization Rule and plug into budget constraint to solve for X^* and Y^*
- **Step 5**: Plug X^* and Y^* into the Utility Function to solve for the level of Utility

Example

Becca has an income of \$100. She is trying to decide between buying pizza and fries. Pizza costs \$10 and fries cost \$5. Becca's preferences are defined by:

$$U = \ln(X) + \ln(Y)$$

Where X is the quantity of pizza and Y is the quantity of fries. What combination of pizza and fries maximizes her utility?

- Set up the budget constraint: $10X + 5Y = 100$
- Find MU_1:
  - $\frac{dU}{dX} = \frac{1}{X}$
- Find MU_2:
  - $\frac{dU}{dY} = \frac{1}{Y}$
- Find MRS:
  - $MRS = \frac{1}{X} \Big/ \frac{1}{Y} = \frac{Y}{X}$
- Apply Utility Maximization Rule:
  - $\frac{Y}{X} = \frac{10}{5}$
- Rearrange the equation in terms of either X or Y:
  - $Y = \frac{10}{5}X$
  - $X = \frac{5}{10}Y$
- Substitute into the budget constraint:

  - $10\left(\frac{5}{10}Y\right) + 5Y = 100$

- Solve for Y^*:
  - $10Y = 100$
  - $Y^* = \frac{100}{10} = 10$
- Plug Y^* into rearranged Utility Maximization Rule and solve for X^*:
  - $10 = \frac{10}{5}X$
  - $X^* = 5$
- Plug X^* and Y^* into U, solve for U^*
  - $U^* = \ln(5) + \ln(10) = 3.91$

Part II

Probability and Statistics

Section 7

Probability and Counting Rules

- Probability is a measure of how likely something is to happen
- Every event has a probability assigned to it
- The more likely an event is to occur, the higher the probability

7.1 Single Events

- The probability that an event can occur can be calculated by:

$$P(event) = \frac{\text{Number of Favorable Outcomes}}{\text{Number of Total Possible Outcomes}} \tag{7.1}$$

- Where, $0 \leq P(event) \leq 1$
- Therefore, the probability that an event does not occur can be calculated by:

$$Prob(event\ does\ not\ occur) = 1 - Prob(event) \tag{7.2}$$

Example 1

If you flip a coin, what is the probability of getting heads?

Number of favorable outcomes: 1
Number of total possible outcomes: 2

$$P(heads) = \frac{1}{2} = 0.50$$

Example 2

If you roll a single die, what is the probability of getting a 4?

Number of favorable outcomes: 1

Number of total possible outcomes: 6

$$P(4) = \frac{1}{6} = 0.167$$

Example 3

There are 5 marbles in a bag. 3 are blue, 1 is red, and 1 is green. What is the probability of picking a blue marble?

Number of favorable outcomes: 3
Number of total possible outcomes: 5

$$P(Blue) = \frac{3}{5} = 0.60$$

7.2 Counting with Multiple Events

7.2.1 The Counting Rule

- The Counting Rule is a way to figure out the number of outcomes in a probability problem

- You multiply the outcomes of the events together to get the total number of outcomes

Example 1

You take a survey with five "yes" or "no" answers. How many different ways could you complete the survey?

There are 5 questions and each question has 2 choices.

Number of total choices: $2 \times 2 \times 2 \times 2 \times 2 = 32$

Example 2

A company puts a code on each different product they sell. The code is made up of 3 numbers and 2 letters. How many different codes are possible?

There are 10 possible numbers: $0 - 9$
There are 26 possible letters: A – Z

Number of total choices: $10 \times 10 \times 10 \times 26 \times 26 = 676,000$

7.2.2 Combinations and Permutations

- When thinking about multiple events, we need to think about how to count the number of possible outcomes and the number of favorable outcomes
- In English, we use the word "combination" anytime we are talking about a group of things together
- However, with statistics, we can be more precise because sometimes the order of things matter
- For example, a combination on a lock requires a specific order
- However, for the lottery, you just need to match the same numbers, they do not need to be in the same order
- When the order *does not* matter, we are dealing with a combination
- When the order *does* matter, we are dealing with a permutation
- We also need to think about whether or not events are allowed to repeat or not

Permutations with Repetition

- A permutation with repetition is dealing with counting a group where order does matter but events are allowed to repeat themselves
- Lets think of an example with a combination lock

Example

A combination lock has 4 dials that each have the numbers 0-9, so each dial has 10 options. Since each dial is independent of each other, we can get the number of possible combinations by:

$10 \times 10 \times 10 \times 10 \ or \ 10^4 = 10,000$

This means that there are 10,000 total possible outcomes for the lock

- More generally, we can calculate the number of total possible outcomes for a permutation with repetition by: n^r
  - $n =$ the number of possible outcomes from the first choice
  - $r =$ the number of times the first choice is repeated
- Anytime we are trying to figure out the number of outcomes where order matters and repetition is allowed, we can use this formula

Permutation without repetition

- A permutation without repetition is dealing with counting a group where order does matter but events are not allowed to repeat themselves

Example

You have three classes you need to take, math, science, and history. How many possible ways are there to arrange these classes?

- You start out with three choices: Math, Science, and History
- If you pick science to be your first class, then your are left to select between math and history for your second class
- If you pick history as your second class, then you only have one choice left for your third class: math

Using the counting rule, the number of possible ways to arrange these classes $= 3 \times 2 \times 1 = 6$

- We can express the number of possible outcomes for a permutation using a math tool called a factorial, which is designated by "!"
  - $10! = 10 \times 9 \times 8 \times 7 \times 6 \times 5 \times 4 \times 3 \times 2 \times 1$
  - On the TI-83/84, this can be found on the MATH button on the PRB tab
- Sometimes, we may not want to use *all* of the possible outcomes, just some

$$_nP_r = \frac{n!}{(n-r)!}$$

Where,

$n =$ the number of possible choices

$r =$ the number of those choices that are selected

NOTE: if $n = r$, then the denominator $= 0! \rightarrow 0! = 1$

Example

Suppose you have 10 classes to choose from and you only need to take 6. How many ways could you take those 6 classes?

There are 10 classes to choose from (n), you are choosing 6 of them (r)

$$_{10}P_6 = \frac{10!}{(10-6)!} = 151,200$$

If you were to choose all 10 classes:

$$_{10}P_{10} = \frac{10!}{(10-10)!} = \frac{10!}{0!} = 10! = 3,628,800$$

If $n = r$, then we have our first scenario, $_nP_r = n!$

Combinations without Repetition

- A combination without repetition is dealing with counting a group where order does not matter, but events are not allowed to repeat
- To calculate the number of possible outcomes for a combination without repetition:

$$_nC_r = \frac{n!}{r!(n-r)!}$$

- $n =$ the number of possible choices
- $r =$ the number of those choices that are selected
- The best example of this is a lottery
  - Suppose a lottery has 48 ping pong balls numbered 1-48. 6 of them are drawn at random
  - If the 6 numbers drawn match the 6 numbers you picked, you win, even if the numbers are not in the same order
    - There are 48 ping pong balls to choose from (n)
    - 6 of them are chosen (r)
    - $_{48}C_6 = \frac{48!}{6!(48-6)!} = 12,271,512$ possible choices

Example

Suppose a lottery has 48 ping pong balls numbered 1-48. 6 of them are drawn at random. If the 6 numbers drawn match the 6 numbers you picked, you win, even if the numbers are not in the same order. How many different combinations of ping pong balls are there?

There are 48 ping pong balls to choose from (n), 6 of them are chosen (r).

$$_{48}C_6 = \frac{48!}{6!(48-6)!} = 12,271,512$$

There are 12,271,512 possible choices.

7.3 Probability with Multiple Events

- It is helpful to remember that $P(event) = \dfrac{\text{Number of Favorable Outcomes}}{\text{Number of Total Possible Outcomes}}$
- These problems require some thought as to what the total number of possible and favorable outcomes are
- For each problem we need to determine how to count these outcomes
- Lets start with continuing with our lottery example

Example 1

What is the probability of winning the lottery that has 48 ping pong balls where 6 numbers are drawn?

Number of possible outcomes: This is a combination, ${}_{48}C_6 = 12,271,512$
Number of favorable outcomes: This is a simple event, 1

$$P(Win\ Lottery) = \frac{1}{{}_{48}C_6} = 0.0000000815 = 0.00000815\%$$

Example 2

If you throw two dice, what is the probability you do not throw any 4, 5, or 6's?

To count the number of possible outcomes, we will use the counting rule:

Number of outcomes for one die: 6
Number of die: 2
Total possible outcomes: $6 \times 6 = 36$

To count the number of favorable outcomes, we need to use the counting rule:

Each die has 3 favorable outcomes: 1,2, or 3
Since there are 2 dice, number of favorable outcomes is $3 \times 3 = 9$

$$P(\text{no 4, 5, or 6}) = \frac{9}{36} = \frac{1}{4} = 0.25$$

Example 3

Suppose you throw three dice, what is the probability you do not throw any 4, 5, or 6's?

To count the number of possible outcomes, we will use the counting rule:

Number of outcomes for one die: 6
Number of die: 3
Total possible outcomes: $6 \times 6 \times 6 = 216$

To count the number of favorable outcomes, we need to use the counting rule:

Each die has 3 favorable outcomes: 1,2, or 3
Since there are 3 dice, number of favorable outcomes is $3 \times 3 \times 3 = 27$

$$P(\text{no 4, 5, or 6}) = \frac{27}{216} = \frac{1}{8} = 0.125$$

Example 4

You have to select a 4 digit PIN for your ATM/debit card. What is the probability that if randomly selected, there are no repeated digits? There are 10 digits to choose from: 0-9

To count the total number of possible outcomes, this is is permutation with repetition because each pin number can be repeated:

Number of possible outcomes for first choice (n): 10
Number of times this choice is repeated (r): 4
Number of possible outcomes: $10^4 = 10,000$

To count the number of events where there are no repeated digits, we need to use a permutation without repetition:

Number of possible choices (n): 10
Number of times this choice is repeated (r): 4
Number of favorable outcomes: ${}_{10}P_4 = \frac{10!}{(10-4)!} = 5,040$

$$P(\text{4 Digit PIN without repeated numbers}) = \frac{5,040}{10,000} = 0.504$$

Example 5

Suppose there are 10 people in a room. What is the probability that at least 2 of them have a birthday that falls on the same day?

First we want to calculate the probability there are no repeated birthdays, then we want to calculate the opposite where there is a repeated birthday. This is a permutation because the order of days matters.

For the number of possible outcomes, people *could* have the same birthday, so this is a permutation with repetition:

Number of possible outcomes (number of days in a year) (n): 365
Number of times first choice is repeated (number of birthdays) (r): 10
Number of possible outcomes: 365^{10}

For the number of favorable outcomes, people all have to have different birthdays, so this is a permutation without repititon:

Number of possible choices (n): 365
Number of choices selected (r): 10

Number of Favorable Outcomes: ${}_{365}P_{10} = \frac{365!}{(365-10)!} = 3.71 \times 10^{25}$

$$P(No\ Repeated\ Birthdays) = \frac{{}_{365}P_{10}}{365^{10}} = 0.883$$

$$P(Repeated\ Birthday) = 1 - P(No\ Repeated\ Birthdays)$$

$$P(Repeated\ Birthday) = 1 - 0.883 = 0.117$$

7.4 Conditional Probability

- Conditional probability means the probability that event A and event B occur at the same time

- Two events can be conditional on each other in two ways:
  - Two events can occur but one event does not depend on another event
  - Two events can occur but the second event depends on the outcome of the first event

7.4.1 Independent Events

- Multiple events that occur, but neither event occurring is dependent on the other event, are called independent events

- To calculate the probability of two events occurring together, we can use the counting rule:

$$P(A \text{ and } B) = P(A) \times P(B) \tag{7.3}$$

Example

If you roll three dice, what is the probability of getting three fours?

$$P(four) = \frac{1}{6}$$

$$P(three\ fours) = \frac{1}{6} \times \frac{1}{6} \times \frac{1}{6} = \frac{1}{216} = 0.00463$$

7.4.2 Dependent Events

- Multiple events that occur, but an event occurring depends on the outcome of a previous event, are called dependent events

- The notation for this is $P(B|A)$, which means the probability that event A occurs conditional on event B having occurred

- To calculate the probability of event A occurring, then event B occurring conditional on event A have occurring, we can use the counting rule as follows:

$$P(A \text{ and } B) = P(A) \times P(B|A) \tag{7.4}$$

Example

Suppose you pick two cards out of a deck of cards. What is the probability the first card is an ace and the second card is a king?

$$P(Ace) = \frac{4}{52}$$

Once the ace has been removed, there are only 51 cards left, so the total possible outcomes changes:

$$P(King|1^{st}\ Card\ Ace) = \frac{4}{51}$$

$$P(Ace \text{ and } King) = \frac{4}{52} \times \frac{4}{51} = \frac{4}{663} = 0.006$$

- When looking at conditional probabilities, it is really important to make sure the statistic you are looking at is what you are actually interested in because

$$P(A|B) \neq P(B|A)$$

- Let's consider the following example:

Example

Suppose you have 100 people. 50 are vaccinated. 50 are not. 14 people are sick, 2 of them are vaccinated.

You see the following headline: "14.3% of all those who are sick are vaccinated." What is this conditional probability?

This the the probability of being vaccinated if you are sick: $P(vaccinated|sick) = \frac{2}{14} = 0.1428$

However, there are more useful probabilities for this issue:

The probability of being sick if you are vaccinated:

$$P(sick|vaccinated) = \frac{2}{50} = 0.04$$

The probability of being sick if you are not vaccinated:

$$P(sick|not\ vaccinated) = \frac{12}{50} = 0.24$$

7.4.3 Bayes Rule

$$P(A|B) = \frac{P(A) \times P(B|A)}{P(B)}$$

- Where,

  $P(A)$ = The probability of event A occurring

  $P(B)$ = The probability of event B occurring

  $P(A|B)$ = The probability of event A occurring given event B having occurred

  $P(B|A)$ = The probability of event B occurring given event A having occurred

- Bayes rule provides us with a way to find the relationship between $P(A|B)$ and $P(B|A)$
  - Sometimes we know $P(A|B)$ or $P(B|A)$ but not the other
  - Using Bayes Rule, as long as we know one, we can find the other
  - Lets look at our previous example:

Example

$$P(sick|vaccinated) = \frac{P(sick)P(vaccinated|sick)}{P(vaccinated)} = \frac{\frac{14}{100} \times \frac{2}{14}}{\frac{50}{100}} = 0.04$$

- Bayes rules also proves us with a way to update our beliefs based on the arrival of new, relevant pieces of evidence
- For example, if we were trying to provide the probability that a given person has cancer, we would initially just say it is whatever percent of the population has cancer
- However, given additional evidence such as the fact that the person is a smoker, we can update our probability, since the probability of having cancer is higher given that the person is a smoker
- This allows us to utilize prior knowledge to improve our probability estimations

Example

You have the following data from a hospital:

10% of your patients have liver disease ($P(A)$)

5% of your patients are an alcoholic ($P(B)$)

7% of those with liver disease are alcoholics ($P(B|A)$)

What is the probability of having liver disease if they are an alcoholic?

$$P(\text{Liver Disease}|\text{Alcoholic}) = \frac{P(\text{Liver Disease})P(\text{Alcoholic}|\text{Liver Disease})}{P(\text{Alcoholic})}$$

$$Prob(\text{Liver Disease}|\text{Alcoholic}) = .10 \times \frac{.07}{.05} = 0.14$$

- Baye's rule can be extended if there are multiple groups:

$$P(A_1|B) = \frac{(A_1)P(B|A_1)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_n)P(B|A_n)}$$

$$P(A_1|B) = \frac{P(A_1)P(B|A_1)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$$

- Where,

$P(A_i)$ = The probability of event A for group i occurring

$P(B|A_i)$ = The probability of event B conditional on being in group i occurring

Example

There is a test for an allergy to cats, but this test is not always right. For people who have an allergy to cats, the test correctly says "yes" 80% of the time. For people who that do not have an allergy to cats, the test says "yes" 10% of the time (false positive). If 1% of the population has an allergy to cats and you take a test that says you have a cat allergy, what is the probability you actually have a cat allergy?

To solve this, we need to use Bayes Rule:

$$P(A_1|B) = \frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2)}$$

Group 1: Has allergy

Group 2: Does not have allergy

$$P(\text{Allergy}|\text{Pos}) = \frac{P(\text{Allergy})P(\text{Pos}|\text{Allergy})}{P(\text{Allergy})P(\text{Pos}|\text{Allergy}) + P(\text{No Allergy})P(\text{Pos}|\text{No Allergy})}$$

$P(\text{Allergy}) = 0.01$

$P(\text{Positive}|\text{Allergy}) = 0.80$

$P(\text{Not Allergy}) = 1 - P(\text{Allergy}) = 0.99$

$P(\text{Positive}|\text{No Allergy}) = 0.10$

$$P(\text{Allergy}|\text{Positive}) = \frac{(0.01)(0.80)}{(0.01)(0.80) + (0.99)(0.10)} = 0.0748$$

This means that if you test positive, there is a 7.48% change you are allergic to cats

PROBABILITY PRACTICE WORKSHEET

Section 8

Describing Data and Distributions

- When analyzing data, there are two types of variables
  - **Quantitative variables** take numerical values for which adding and averaging make sense
  - There are two types of quantitative variables: discrete and continuous
    - A discrete variable can be counted and there are no in-between values (e.g., 0, 1, 2, 3, 4, etc)
      - The number of people at an event
      - The number of people who are sick
      - The number of of shots a player takes in a game
    - A continuous variable is a variable that can take any number of values
      - Height
      - Weight
      - Speed
  - **Categorical variables** are variables where something falls into different categories and the number of times something occurs in each category can be counted
    - Men vs women
    - State people live in
    - Marital status (e.g., married, single, divorced, widowed)

8.1 Visualizing Data using Graphs

- Whenever you have a set of data, one of the first things you should always do is plot the data graphically
- Depending on the type of data that you have, different graphs may be appropriate
- **Categorical variables**
  - Bar graph where each bar is a category

  - Pie chart

- **Quantitative variables**
  - Histogram
    - Breaks the values of a variable into "bins" or "ranges" and displays only the count or percent of the observations that fall into each "bin"
  - Line graph
    - Use when you want to plot a variable that occurs at different points in time
- Graphing the data helps us see the shape of the data, which we will discuss more later
- You will have an Excel assignment to get you more comfortable with graphing data

8.2 Descriptive Statistics

- Economists and data analysts gather and analyze data to test the predictions of these models and to asses the impact of policy changes
- It is often impossible to measure the true value of a population
- So we collect data on samples of the population to make inference on the entire population
  - The Law of Large Numbers states that as a sample size grows, the mean of the sample will converge to the mean of the population
  - There is no guarantee that a single sample will reflect true population characteristics, particularly if the sample is small
- This section is going to go through how how to describe and understand samples of data
  - We are not going to focus on analyzing statistics of an entire population because we rarely have data on an entire population

8.2.1 Measures of Central Tendency

- Measures of central tendency help find the middle or "typical value" of a data set
- There are two common measures of central tendency
  - Mean
  - Median

Mean

- The mean (or average) of a data set is the sum of the value of each observation divided by the number of observations
- This is the most common measure of central tendency
- The mean is usually expressed using the greek letter "mu"

$$\mu = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

  Where, $n =$ the number of observations in the data (the number of data points)

- The are two instances where the mean can be misleading as a measure of central tendency
  - If the data has outliers
    - An outlier is a value that is much larger or smaller than the rest of the data
  - If the data is skewed (we will discuss this more soon)

Median

- The median is the middle value of a data set that has been arranged in order of magnitude
- If there are an even number of observations, then we can calculate the median by taking the mean of the two middle scores
- The median is less likely to be impacted by outliers or skewed data
- If the median and mean are not similar, it likely means either the data is skewed or there are outliers (or both)

8.2.2 Measures of Variability

- Measures of central tendency tell us where, on average, most of the data is
- Measures of variability tell us how spread apart the data is
- Two data sets can have the same central tendency but have various levels of variability
- Combining these two measures can help paint a complete picture of the data
- There are two main measures of variability
  - Variance
  - Standard Deviation

Variance

- The variance is the average of squared deviations from the mean
- A deviation from the mean is how far a score lies from the mean
- Variance reflects the degree of spread of the data set, the more spread the data, the larger the variance

$$\sigma^2 = \frac{(x_1-\mu)^2+(x_2-\mu)^2+\cdots+(x_n-\mu)^2}{n-1} = \frac{\sum_{i=1}^{n}(x_n-\mu)^2}{n-1}$$

Standard Deviation

- The standard deviation is the average amount of variability in the data set
- It tells you, on average, how far each score is from the mean
- The larger the standard deviation, the more spread out the data is
- To calculate the standard deviation, it is the square root of the variance

$$\sigma = \sqrt{\frac{(x_1-\mu)^2+(x_2-\mu)^2+\cdots+(x_n-\mu)^2}{n-1}} = \sqrt{\frac{\sum_{i=1}^{n}(x_n-\mu)^2}{n-1}}$$

Example

Suppose you have the following data on returns from the S&P 500 for the last five years:

2020	31.10%
2019	-4.41%
2018	21.94%
2017	11.93%
2016	1.31%

Find the Mean, Median, Variance, and Standard Deviation.

Mean:

$$\mu = \frac{31.10-4.41+21.94+11.93+1.31}{5} = 12.37$$

Median: (Arrange from smallest to largest)

-4.41 1.31 11.93 21.94 31.10 → 11.93 is the median

Variance:

$$\sigma^2 = \frac{(31.10 - 12.37)^2 + (-4.41 - 12.37)^2 + (21.94 - 12.37)^2 + (1.31 - 12.37)^2}{4}$$

$\sigma^2 = 211.7$

Standard Deviation:

$$\sigma = \sqrt{\frac{(31.10 - 12.37)^2 + (-4.41 - 12.37)^2 + (21.94 - 12.37)^2 + (1.31 - 12.37)^2}{4}}$$

$\sigma = 14.55$

8.3 Relationship between two Variables

- Sometimes we may have multiple variables in a data set and we want to understand the relationship between those two variables
- This can be very useful when doing forecasting, machine learning, and data analysis
- There are two measures that measure the relationship and dependency between two variables:
    - Covariance
    - Correlation

8.3.1 Covariance

- Covariance, $cov(x, y)$, measures the directional relationship between two variables
- If covariance is positive, the variables move together (i.e., if one increases the other increases)
- If covariance is negative, this means the variables move in opposite directions (i.e., if one increases, the other decreases)
- If covariance is zero, the two variables are unrelated

$$cov(x, y) = \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y})}{n - 1} \tag{8.1}$$

$$cov(x, y) = \frac{1}{n - 1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \tag{8.2}$$

- Covariace cannot be compared between two variables because they are not standardized, it can only tell us the direction two variables move in

8.3.2 Correlation

- Correlation, ρ, measures the directional relationship and the strength of that relationship between two variables
- Correlation is standardized, so its values fall between -1 and 1
- If $\rho > 0$, then the two variables have a positive correlation and move in the same direction
- If $\rho < 0$, then the two variables have a negative correlation and move in the opposite direction
- The closer ρ gets to 1 or -1, the stronger the relationship is
  - If $\rho = 1$, then both variables move in exactly the same direction in the same magnitudes
  - If $\rho = -1$, then both variables move in exactly the opposite direction in the same magnitudes

$$\rho = \frac{(x_1 - \bar{x})(y_1 - \bar{y}) + (x_2 - \bar{x})(y_2 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y})}{(n-1)\sigma_x \sigma_y} \tag{8.3}$$

$$\rho = \frac{1}{(n-1)\sigma_x \sigma_y} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \tag{8.4}$$

- Where,

  σ_x is the standard deviation of x

  σ_y is the standard deviation of y

$$\rho = \frac{cov(x, y)}{\sigma_x \sigma_y}$$

8.4 Probability Distributions

- It is often helpful to graph a histogram for a continuous variable or a bar graph for a discrete variable
- The shape of the graph will help determine the type of distribution if the data
- Some distributions are well known with well defined mathematical properties, some are not
- A distribution is symmetrical if the *mean* and *median* are similar in value

- If the the *mean* $>$ *median*, then a distribution has a positive skew

  - A distribution with a positive skew is where the values are clustered towards the left side with few values further out to the right side
  - Examples of data with a positive skew:
    - Income
    - House values
    - Number of pets in a household

- If the *mean* $<$ *median*. then a distribution has a negative skew

  - A distribution with a negative skew is where the values are clustered towards the right side with few values closer in to the left side
  - Examples of data with a negative skew:
    - Age of death
    - GPA values
    - Stock market returns

8.4.1 Probability Density and Probability Mass Functions

- For a continuous variable, each distribution has a probability density function (PDF)

  - A PDF of a continuous variable will graphically show the distribution of the data
  - A PDF also shows that for variable X, what is the probability that the outcome is k, or $P(X = k)$
  - A PDF shows the probability of every possible outcome for a continuous variable, therefore the area under the PDF curve must sum to 1 or 100

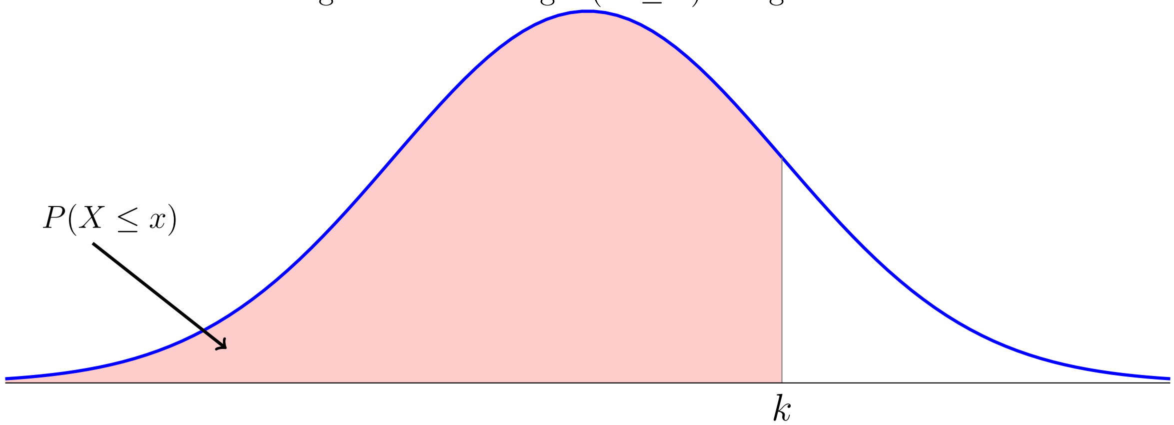
- For a discrete variable, each distribution has a probability mass function (PMF)

  - A PMF of a discrete variable behaves similarly to the PDF of a continuous variable, and shows $P(X = k)$

- Every distribution has a unique PDF or PMF

8.4.2 Cumulative Density Function

Figure 8.1: Finding $P(X \leq k)$ using a PDF



- If we want to find the probability that variable X is less than or equal to k, it is the area under the PDF to the left of that value

- To find this area, or probability, we can use what is called the Cumulative Distribution Function (CDF)

  - In order to get the CDF of a continuous variable, it is the integral of the PDF (we do not cover integrals in ths class, but the CDF can be evaluated using a graphing calculator or Microsoft Excel)

- A CDF evaluated at a specific value of x, gives the probability that the value of the sample takes a value less than or equal to x, or $P(X \leq k)$

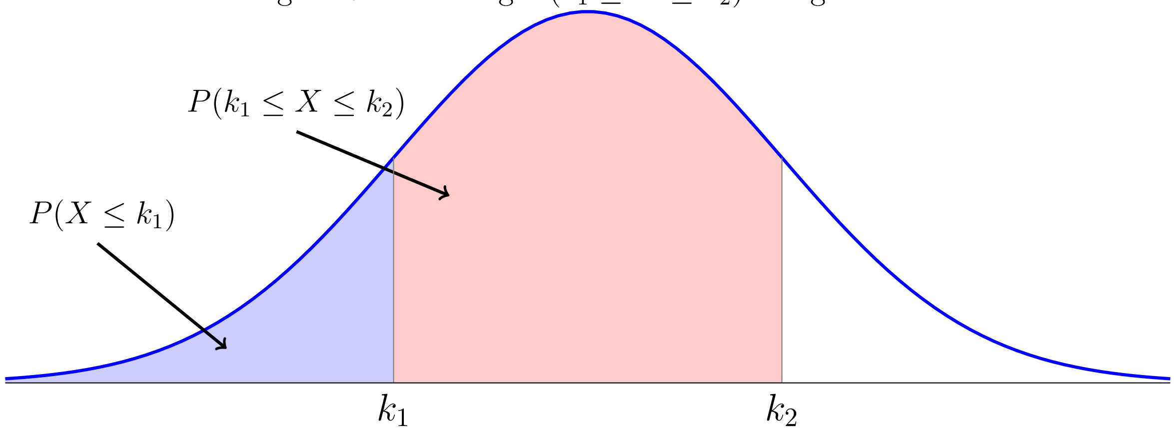
- Because the area under the PDF must sum to 1, if we want the the probability that variable X is greater than k, we can find it by: $P(X > k) = 1 - P(X \leq k)$, or 1 minus the CDF evaluated at the value of k

- Suppose we want to find the probability that variable X is between two values, k_1 and k_2

  - First, evaluate the CDF at k_2 to find the probability that $P(X \leq k_2)$
  - Second, evaluate the CDF at k_1 to find the probability that $P(X \leq k_1)$
  - Last, subtract: $P(k_1 \leq X \leq k_2) = P(X \leq k_2) - P(X \leq k_1)$

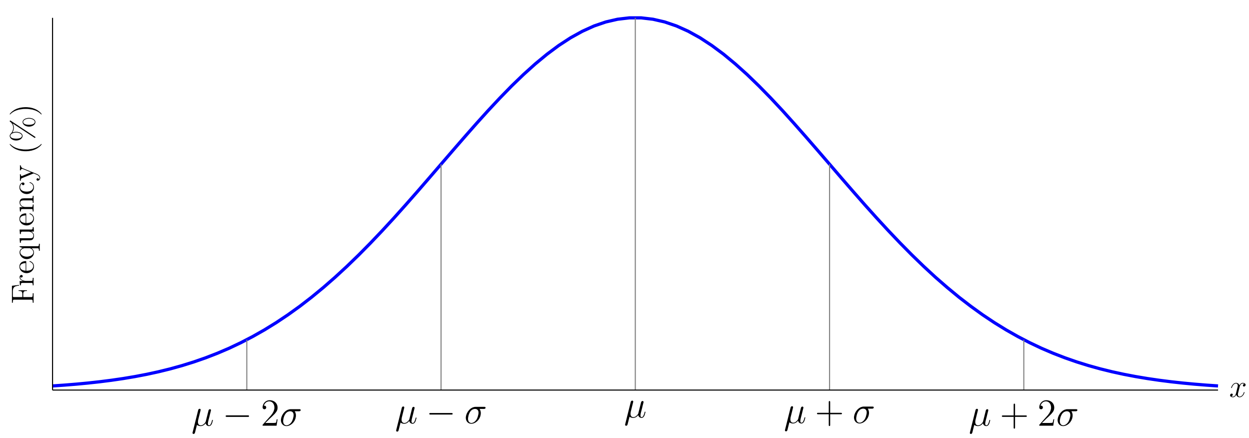
- Every distribution has its own unique CDF

Figure 8.2: Finding $P(k_1 \leq X \leq k_2)$ using a PDF



8.5 Normal Distribution

- On of the most common distributions seen in real-world data is the normal distribution
- The normal distribution is often used in the natural and social sciences to represent real-valued random variables
  - This is because of the Central Limit Theorem (CLT)
    - The CLT states that if you have a population with mean μ and standard deviation σ and take sufficiently large samples of the population, then the distribution of the sample means will have a normal distribution, regardless if the true population is normal or not.
- The normal distribution has a bell shaped curve



- For a given variable X, if X is distributed normally, we use the following notation: $X \sim N(\mu, \sigma)$
- The normal distribution has several nice properties that help us with visualizing the spread of the data

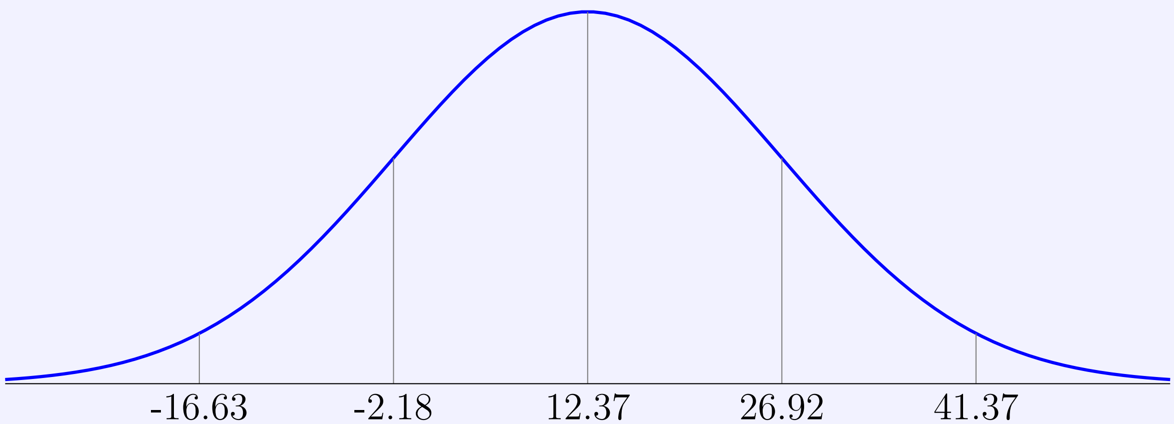
- If data is distributed "normally," then 68.3% of all the data lies between ±1 standard deviations from the mean and 95.5% of all the data lies between ±2 standard deviations from the mean

Example

Let's use data from our previous example using stock returns, where we found that $\mu = 12.37$ and $\sigma = 14.55$

$$\mu + \sigma = 12.37 + 14.55 = 26.92 \qquad \mu - \sigma = 12.37 - 14.55 = -2.18$$

$$\mu + 2\sigma = 12.37 + 14.55 = 41.37 \qquad \mu - 2\sigma = 12.37 - 14.55 = -16.63$$



This means that we can expect 68.3% of stock returns to fall between -2.18% and 26.92% and that 95.5% of stock returns will fall between -16.63% and 41.37%.

- When data is distributed normally, the mean and the median should have similar values
- However that is not always the case and sometimes the data can be skewed and there may be a different distribution for that data set, which is why its important to graph a histogram of the data
- If we wanted to graph the PDF of the normal distribution, we would use the following equation with using the mean, μ, and standard deviation, σ for the sample:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-(x-\mu)^2/2\sigma^2} \tag{8.5}$$

- We can use the PDF to find the probability that the variable X takes the value of x

Example

Suppose you have the height of 1,000 third graders where the mean height is 37.0 inches with a standard deviation of 1.99. What is the probability of a first grader being 39 inches?

$\mu = 37.0$
$\sigma = 1.99$
$x = 39.0$

$$f(x) = \frac{1}{1.99\sqrt{2\pi}} \cdot e^{-(39-37)^2/2(1.99)^2} = 0.12098$$

- We can use the CDF of the normal distribution to find the probability that the variable X is less than or equal to k, or $P(X \leq k)$
- The CDF of the normal distribution is as follows:

$$\Phi(k) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{k} e^{-(k-\mu)^2/2\sigma^2} \tag{8.6}$$

- If we want to know the probability that X is greater than k, then: $P(X > k) = 1 - \Phi(k)$
- For the normal distribution, evaluating the CDF requires integration, which is a mathematical technique we do not require you to know
- So we can calculate the value of the CDF using a calculator or Excel

Normal CDF with TI-83/84

Step 1: Press the 2nd key
Step 2: Press the VARS key
Step 3: Select option 2 or "normalcdf("
The format you type in the calculator is: `normalcdf(lower bound, upper bound,` μ, σ`)`

Note: If you want $P(X \leq k$, then the lower bound would be $-\infty$, which can be approximated on the calculator by using "–1e99" for the lower bound, which is the default option. If you want to find the $k_1 \leq X \leq k_2$, then k_1 would be the lower bound and k_2 the upper bound

Normal CDF with Microsoft Excel

`=NORM.DIST(`k, μ, σ`, cumulative)`

Where,

- x is the value of x
- mean is the mean of the distribution
- std. dev. is the standard deviation of the distribution
- cumulative is set to "TRUE" to return the value of the CDF, if set to "FALSE," it returns the value of the PDF for a given value of x

Example

Using the same information from the previous example, what is the probability the height of a third grader is less than or equal to 39, or $P(height \leq 39)$?

For TI-83/84

```
normalcdf(−1e99, 39, 37, 1.99)
```

For Microsoft Excel

```
=NORM.DIST(39,37,39,TRUE)
```

$$\Phi(39) = 0.84256$$

There is a 84% chance that the height of a third grader is less than or equal to 39 inches.

8.6 Poisson Distribution

- The Poisson probability distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event
- The Poisson distribution helps to predict the probability of certain events happening when you know how often the event has occurred
- The Poisson distribution is often used by companies to make forecasts about the number of customers or sales on certain days, months, or seasons of the year
- This is important to ensure that companies have the appropriate level of inventory on hand
- Because the Poisson distribution is a discrete probability distribution, it has what is called a Probability Mass Function (PMF), which operates similar to the PDF of a continuous probability distribution
- The PMF for the Poisson distribution is as follows:

$$f(k, \lambda) = \frac{\lambda^k \cdot e^{-\lambda}}{k!} \tag{8.7}$$

Where,

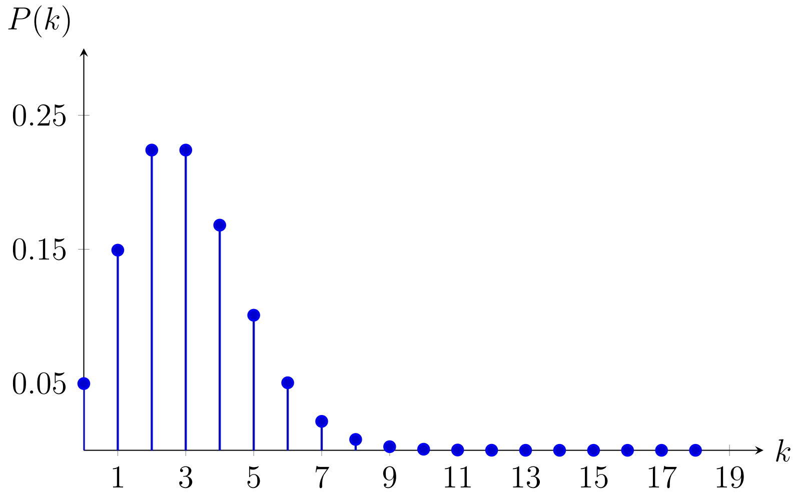
k is the number of occurrences

λ is the average rate

- The PMF of the Poisson distributions tells us the probability that the variable X occurs k times $\longrightarrow P(X = k)$

- The notation for the Poisson distribution is $X \sim Poisson(\lambda)$

Figure 8.3: Poisson Distribution with $\lambda = 3$



Example

Leah's answering machine receives about six telephone calls between 8 a.m. and 10 a.m. What is the probability that Leah receives 3 calls in the next 15 minutes?

On average, Leah receives 6 telephone calls in two hours and we split each hour into 15-minute intervals.

There are 8 15-minute intervals in 2 hours, so that means Leah receives an average of $^6/_8 = 0.75$ calls every 15 minutes.

For this problem then, $k = 3$ and $\lambda = 0.75$

$$f(3, 0.75) = \frac{0.75^3 \cdot e^{-0.75}}{3!} = 0.03321$$

There is a 3.3% chance that Leah gets exactly 3 calls in the next 15 minutes.

- If we wanted to find the probability that the discrete variable X is less than or equal to k, or $P(X \leq k)$, we can use the CDF of the Poisson distribution to calculate this:

$$F(k, \lambda) = \sum_{i=1}^{k} \frac{\lambda^k \cdot e^{-\lambda}}{k!} \tag{8.8}$$

- The CDF of the Poisson function essentially sums the probability of each k in the specified range
- If we wanted to find the probability that X is greater that k, $P(X > k) = 1 - F(k, \lambda)$

Normal CDF with TI-83/84

Step 1: Press the 2nd key

Step 2: Press the VARS key

Step 3: Select option D or "poissoncdf("

The format you type in the calculator is: `poissoncdf(`λ`, `x`)`

Normal CDF with Microsoft Excel

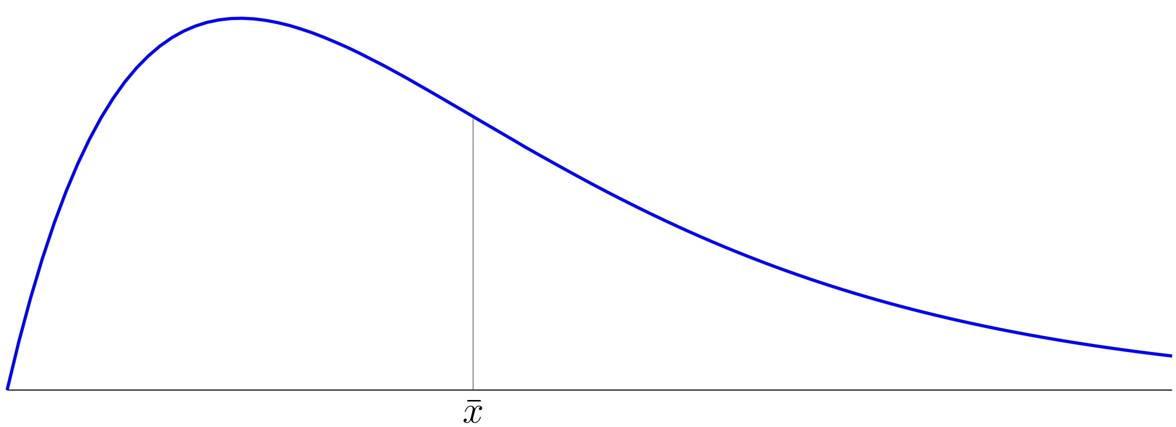
`=POISSON.DIST(x, `λ`, cumulative)` Where,

- x is the value of x
- λ is the mean
- cumulative is set to "TRUE" to return the value of the CDF, if set to "FALSE," it returns the value of the PDF for a given value of x

8.7 Other Distributions

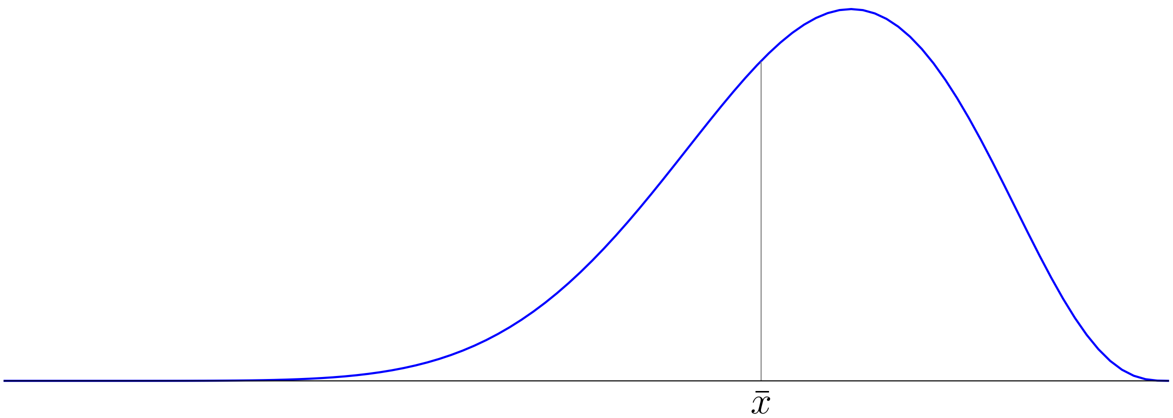
- There are numerous other distributions, but the PDF/PMF and CDF of all of them behave the same way and can be used to calculate probabilities associated with variable X that follows a given distribution

8.7.1 Gamma Distribution



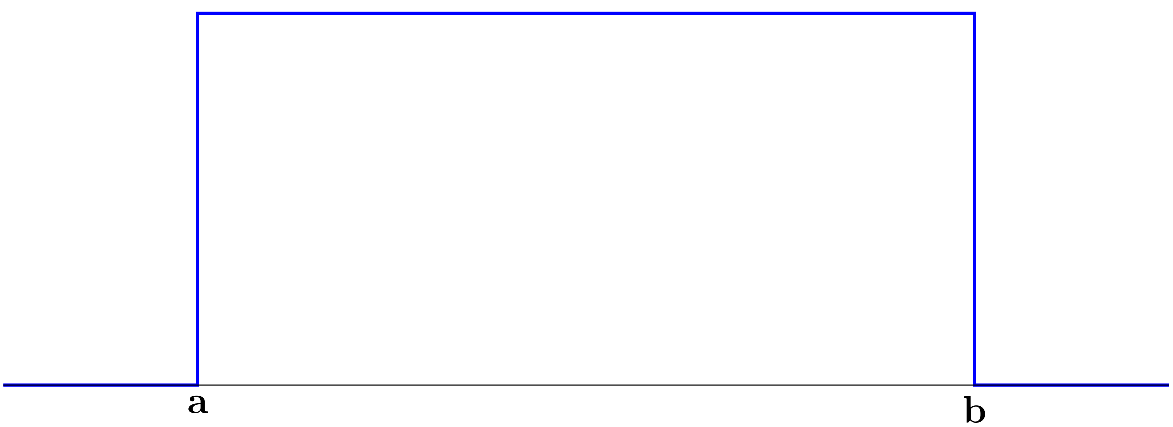
- The Gamma Distribution is continuous probability distribution with a positive skew

8.7.2 Beta Distribution



- The Beta Distribution is a continuous probability distribution with a negative skew

8.7.3 Uniform Distribution

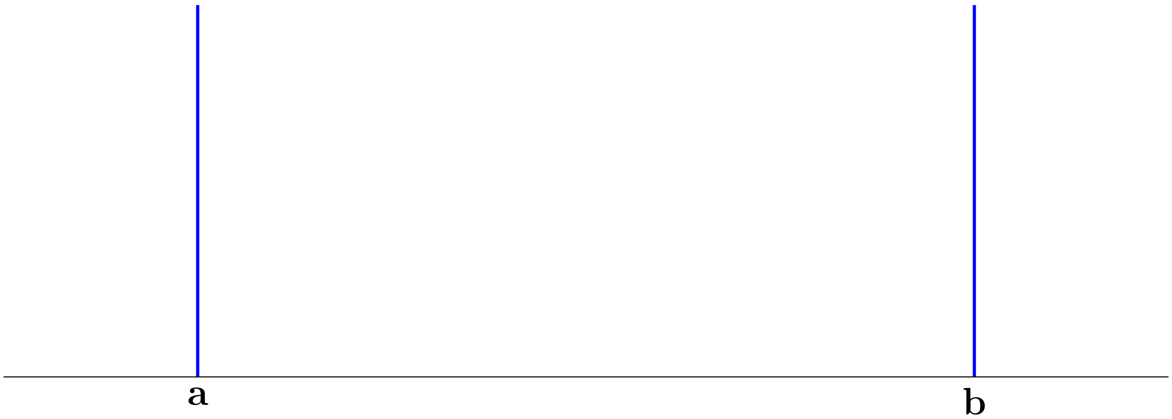


- The Uniform Distribution is a distribution where there is equal likelihood of any value between values a and b

- Examples of a uniform distribution

  - The probability of drawing a specific suit from a deck of cards
  - The probability of rolling a specific number on a die
  - The probability of winning a raffle

8.7.4 Bernoulli Distribution



- The Bernoulli Distribution is a distribution where an event only has two outcomes, a and b, and how likely those outcomes are to occur
- Examples of a Bernoulli Distribution
  - Flipping a coin
  - The number of boys and girls born each day
  - Winning or losing an event

8.8 Discrete Probability and Expected Value

- A discrete probability distribution function (PDF) shows all of the possible outcomes of a discrete variable along with the probability of those outcomes occurring
- Each probability is between 0 and 1, inclusive
- The sum of all the probabilities must equal one
- Let's start with a simple example of rolling a single die:

Example

Outcome of Roll	1	2	3	4	5	6
Probability of Outcome	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

- There are six possible outcomes
  - There is a finite amount of outcomes
  - You can't have an "in-between" outcome

- If you roll a die, one of these outcomes if guaranteed to happen, we just don't know which one
- Because of this, the sum of all the probabilities must be 1
  - $\frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$

Expected Value

- We can use the discrete PDF to calculate the expected value of an event
- Expected value is the anticipated value of a random variable at some point in the future
- We can use a discrete probability distribution to calculate the expected value (or expected outcome) of an event

$$E[Outcome] = p_1x_1 + p_2x_2 + \cdots + p_nx_n = \sum_{i=1}^{n} p_ix_i$$

Where,

p_i = the probability of a possible outcome

x_i = the result from the possible outcome

- Let's use the Powerball example we saw from the worksheet:

Example 1

Previously, we showed that the probability of winning the Powerball lottery is 0.000000342%.

Suppose the jackpot is $10 million and a Powerball ticket costs $2. What would your expected earnings be?

There are two outcomes: you win, or you don't win. Using this, we can set up a probability table knowing that regardless of if you win or don't, you still must pay $2 for the lottery ticket:

	Win	**Don't Win**
Payoff	$\$10,000,000 - \$2 = \$9,999,998$	$\$0 - \$2 = -\$2$
Probability	0.000000342%	99.99999966%

$$E[Lottery\ Return] = .00000000342(\$9,999,998) + 0.9999999966(-\$2) = -\$1.97$$

This means that your expected return from buying a lottery ticket is to lose $1.97.

Example 2

You plan to invest $1,000 in the stock market, but you don't know what the market will do over the next year. Industry analysts have derived the following discrete PDF for stock market returns:

Stock Market Returns	**Strong**	**Average**	**Weak**
Return	8%	2%	−2%
Probability	10%	70%	20%

$$E[return] = 0.10(0.08) + 0.70(0.02) + 0.20(-0.02) = 0.018 = 1.8\%$$

If you invested $1,000 in the stock market today, you could expect that on average, you will have a 1.8% return.

$$\$1,000 + \$1,000(0.018) = \$1,018$$

Section 9

Hypothesis Testing

- Hypothesis testing in statistics is a way to test if the results of an experiment or research yields meaningful results
- You want to know if the results you have just happened by random chance
- A valid experiment or research design should be able to be repeated and get similar results
- Meaningful results are called statistically significant
- Before we can conduct a statistical test, we must first define the null and alternative hypotheses
- The null hypothesis, H_0, proposes that there is no difference between certain characteristics of a population
- The alternative hypothesis, H_1, is a claim about the population that is contradictory to the null hypothesis
- Since the null and alternative hypotheses are contradictory, we must examine evidence to decide if you have enough evidence to reject the null hypothesis or not
- The evidence comes from sample data
- The statistical test that you use to test the alternative hypothesis will depend on the distribution of your data and the question you are asking
- For this class, we are going to deal with normally distributed data

9.1 The T-Test

- We can use the t-test in two ways:
    - A one-sample test is used to determine if the mean of a population has a value specified in the null hypothesis

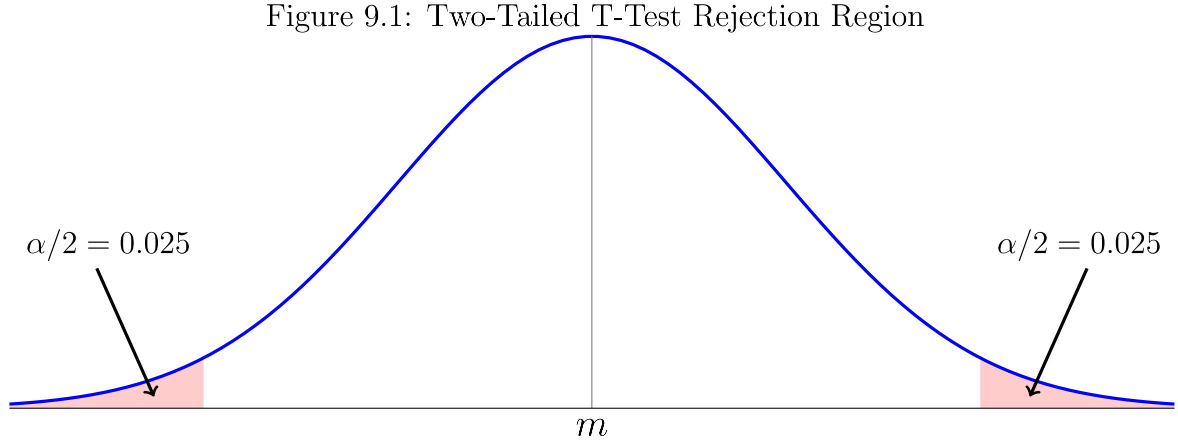
- A two-sample test is used two test if the means of two separate populations are equal

9.1.1 One Sample T-Test

- If you have a sample of data, we can use the t-test to determine if the mean of your sample is equal to another value
- Often, we are testing to see if the mean is different from zero
  - Suppose you collect data on the prices of cigarettes before and after a change in the cigarette tax
  - You want to know if the tax caused an increase in prices
  - You don't have data on every store that sells cigarettes, only a few in the local area
  - So you need to know if the price change you calculated is statistically different from zero
  - If it is, then you likely captured the actual price change
  - If it is not, then you do not have enough data (or statistical power) to say one way or the other

Two-Tailed Test

Figure 9.1: Two-Tailed T-Test Rejection Region



- For a two-tailed T-Test, we are testing to determine if the mean of our sample, μ, is statistically different from another value, m
- First, we need to define the null and alternative hypotheses for the t-test:
  - Null hypothesis: $H_0 : \mu = m$
  - Alternative hypothesis: $H_1 : \mu \neq m$

- For the T-Test, we are trying to determine if we have collected enough data to determine that the null hypothesis, H_0, is false, so that we can accept the alternative hypothesis, H_1, as true

- Because we only have data on a sample of the population, we need to determine how sensitive our test is going to be to the possibility of coming to an incorrect conclusion

- We do this by determining a significance level, which is defined at α

  - The significance level is the probability of rejecting the null hypothesis when it is true
  - The significance level also defines how far the sample mean, μ, must be from m before we can reject the null
  - The most common value of for a significance level is = 0.05, which means that we are conducting a hypothesis test where there is a 5% chance we are going to incorrectly reject the null hypothesis
    - It is not uncommon to see a problem expressed as a confidence level, which is $(1 - \alpha)$
    - If there is a significance level of $\alpha = 0.05$, then you could say that the test has as $(1 - 0.05) = 0.95$, or 95% confidence level
  - The lower the value of α, the less likely the test is to incorrectly reject the null hypothesis

- We use α to determine the rejection region in the normal distribution

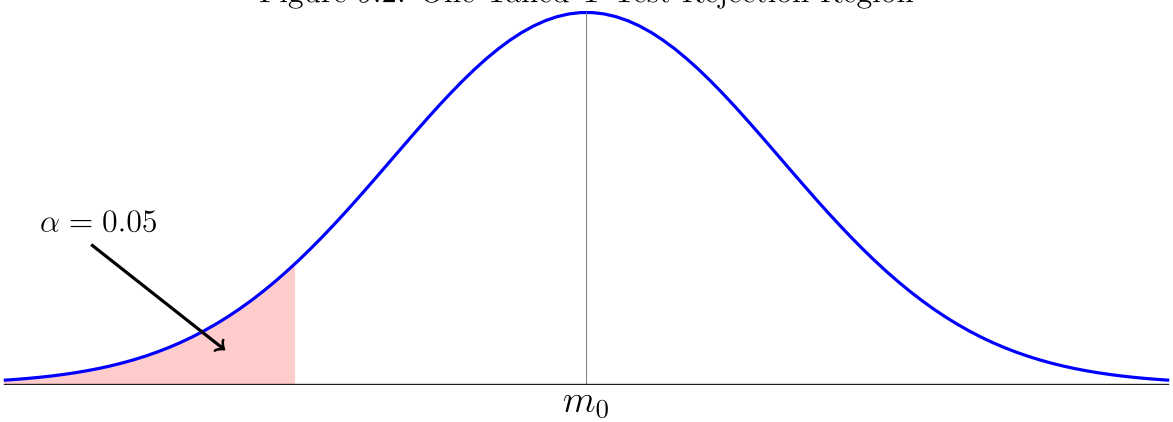
  - The rejection region is the percentage of the area under the curve of the normal distribution that the sample mean, μ will fall in if we reject the null hypothesis
  - For a two-tailed test, the rejection region lies on both tails of the distribution, where each tail has /2 percent of the total area

- We can see this on the normal distribution for a two-tailed test in Figure 9.1

- If the mean of the sample lies in the rejection region, then that means there is enough evidence to reject the null hypothesis and accept the alternative hypothesis

One-Tailed T-Test

Figure 9.2: One-Tailed T-Test Rejection Region



- A one-tailed T-Test is similar to a two-tailed T-Test, except that instead of testing to determine if the sample mean, μ, is equal to m, we are testing to determine if the sample mean is greater than or less than m
- We need to define the null and alterhative hypotheses:
  - Null hypothesis: $H_0 : \mu \leq m$
  - Alternative hypothesis: $H_1 : \mu > m$

  OR
  - Null hypothesis: $H_0 : \mu \geq m$
  - Alternative hypothesis: $H_1 : \mu < m$
- For a one-tailed test, the rejection region lies only on one tail of the normal distribution, accounting for α percent of the area under the curve
- If the sample mean, μ, lies in the rejection region, then there is enough evidence to reject the null hypothesis and accept the alternative hypothesis

Steps to Conduct a One-Sample T-Test

- Step 1: Determine the value of comparison, m
- Step 2: State the null hypothesis, H_0
  - H_0: $mean = m$ or H_0: $mean \geq m$ or H_0: $mean \leq m$
- Step 3: State the alternative hypothesis, H_1
  - H_1: $mean \neq m$ or H_1: $mean > m$ or H_1: $mean < m$
- Step 4: Determine the significance level, α

- Step 5: Determine the degrees of freedom
  - Degrees of freedom is an estimate of the number of independent pieces of information used to calculate the test
  - for the T-Test, degrees of freedom $= n - 1$
- Step 6: Calculate the T-Statistic
  - $t = \dfrac{\mu - m}{\sigma/\sqrt{n}} = \mu - m\left(\dfrac{\sqrt{n}}{\sigma}\right)$

    μ is the mean of the sample

    n is the number of observations/data points

    σ is the standard deviation of the sample
- Step 7: Determine the critical value from the table
- Step 8: Compare the T-Statistic to the critical value
  - If |T-Statistic| > critical value, then there is enough evidence to reject the null hypothesis and accept the alternative hypothesis
  - If |T-Statistic| $\leq$ critical value, the we do not have enough evidence to reject the null hypothesis
    - We cannot claim that the null hypothesis is true, we did not test for that

Example

In the population, the average IQ is 100. A team of scientists is testing a new medication to see what kind of effect it has on intelligence. A sample of 30 people participate in a trial where the mean IQ is 140 with a standard deviation of 20. Did the medication of a significant effect on IQ with a 95% confidence level ($\alpha = 0.05$)?

- From the problem, we know that $\bar{x} = 140$, $n = 30$, and $\sigma = 20$
- Step 1: We are comparing the sample IQ to the national average, so $m = 100$
- Step 2: $H_0 : mean = 100$
- Step 3: $H_1 : mean \neq 100$
- Step 4: $\alpha = 0.05$
- Step 5: $n - 1 = 30 - 1 = 29$
- Step 6: $t = (140 - 100)\left(\dfrac{\sqrt{30}}{20}\right) = 10.95$

- Step 7: We are doing a two-tailed test, $\alpha = 0.05$, and degrees of freedom = 29. From the table, the critical value = 2.045
- Step 8: $|10.95| > 2.045$. Since $|t| >$ critical value, there is enough evidence to reject H_0 and say that the IQ of the sample is statistically different from the IQ of the national average

Example2

The mean body weight from men in the United States is 195lb. A survey of 50 male cadets at VMI found that the average weight was 170lb with a standard deviation of 10lb. Is the weight of a male VMI cadet less than the population average? Assume $\alpha = 0.05$

- From the problem, we know that $\bar{x} = 170$, $n = 50$, and $\sigma = 10$
- Step 1: $m = 195$
- Step 2: $H_0 : mean \leq 195$
- Step 3: $H_1 : mean > 100$
- Step 4: $\alpha = 0.05$
- Step 5: $n - 1 = 50 - 1 = 49$
- Step 6: $t = (170 - 195)\left(\frac{\sqrt{50}}{10}\right) = -17.68$
- Step 7: We are doing a one-tailed test, $\alpha = 0.05$, and degrees of freedom = 49. From the table, the critical value = 1.684
- Step 8: $|17.68| > 1.684$. Since $|t| >$ critical value, there is enough evidence to reject H_0 and say that the mean weight of male cadets at VMI is statistically less than the population

- Because we are using a sample of data from the population, there is a possibility we reach the wrong conclusion
- When doing hypothesis tests, there are two possible errors we can make
  - Type I Error
    - This is where we reject the null hypothesis and support the alternative hypothesis, we should not have
    - The significance level, α, is how likely a Type I Error is going to occur
  - Type II Error

- This is where we fail to reject the null hypothesis but the alternative hypothesis is actually true

- The significance level is the probability of committing a Type I Error
- In statistics, alpha, α, represents the significance level
  - As α increases, the probability of committing a Type I Error increases
  - As α decreases, the probability of committing a Type II Error increases
  - To balance between the two, the most common alpha is 0.05, but sometimes 0.10 and 0.01 are also used
  - If $\alpha = .05$, then there is a 5% chance of a Type I Error, or we are 95% confident we are right

9.1.2 Two-Sample T-Test

- Sometimes we collect data for different groups of people and we want to know are the means for a variable different for the different groups?
  - Men and women
  - Young and old
  - Republicans and Democrats
  - Between races
- The steps to conduct a two-sample test are the same as a one sample test
- The difference is in how the null and alternative hypotheses are stated, the test statistic, and degrees of freedom
  - Null hypothesis
    - H_0: mean group 1 = mean group 2
  - Alternative hypothesis
    - H_1: mean group 1 $\neq$ mean group 2
  - Degrees of freedom
    - $n_1 + n_2 - 2$
- For a two-sample test, there will be two different samples, each sample will have its own mean, standard deviation, and number of observations
- The find the t-statistic:
  - $t = \dfrac{\mu_1 - \mu_2}{\sqrt{Z\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}}$

- $Z = \dfrac{(n_1 - 1)\sigma_1^2 + (n_2 - 1)\sigma_2^2}{n_1 + n_2 - 2}$
- Where,

  $\mu_1, \mu_2 =$ are the mean of each groups

  $\sigma_1^2, \sigma_2^2 =$ are the variance of each group

  $n_1, n_2 =$ are the number of observations of each group

- The Z term is what is called the pooled standard deviation
- Once we calculate the t-statistic, we compare it to the critical value to determine if we should reject the null hypothesis

9.2 P-Values

- The p-value (probability value) is the probability that the null hypothesis is correct
- It is expressed in a decimal from between 0 and 1
- Once the p-value has been calculated, we compare it to α
- If the p-value $\leq \alpha$, then that means there is enough evidence to reject the null hypothesis
- P-values are typically calculated on a calculator or computer using the t-statistic, but can be approximated manually using the t-table
  - On the row for the degrees of freedom, find the where the calculated t-statistic would be
  - See what the two α values it falls between, take the average, and that is an approximation of the p-value

Part III

Introduction to Finance

Section 10

The Time Value of Money

- There is a time component to money
  - Example: You have \$100 and are at the bar where you can buy drinks and food today
  - I come along and say you should hold your \$100 and save it for the summer where you can use it for something else
  - You likely would say no, you are going to spend it now because that is what is most fun and "valuable" to you in the present
  - From your economics classes, there is an "opportunity cost" associated with waiting to spend your money → If you wait to spend it, you are forgoing current utility ("satisfaction")
  - That \$100 is worth more to you today than it would be in the future
- If I wanted you to hold your money for the future, I would need to compensate your for the opportunity cost or lost utility
  - I could offer you the opportunity to put your money in an investment that earns interest with the promise that you will have more money in the future than you do today
- **Present Value (PV)** is stating a future dollar value in today's terms when accounting for the interest you could earn
- **Future Value (FV)** is stating today's dollar value in terms of a future date when accounting for the interest you could earn

10.1 Future Value

- If we were to state dollars today in terms of a future period, we call this future value
- Money in an interest bearing account will accumulate interest for the length of time that it is stored in the account

10.1.1 Simple Interest

- Simple interest is where the interest payment is calculated on the PV and then applied each period and does not cosider that the amount in the account is changing as a result of the interest
- **Example:** Suppose I put \$100 in a bank account today that will earn 5% interest annually, how much money will I have after three years?

$$Interest = PV \times interest\,rate \tag{10.1}$$

$$Interest = (\$100)(.05) = \$5$$

$$Simple\ Interest = PV_0 \times i \times n \tag{10.2}$$

Where,

i = the interest rate

n = the number of years

$$Simple\ Interest = (\$100)(.05)(3) = \$15$$

$$FV_n = PV_0 + Simple\ Interest \tag{10.3}$$

$$FV_n = PV_0 + PV_0 in \tag{10.4}$$

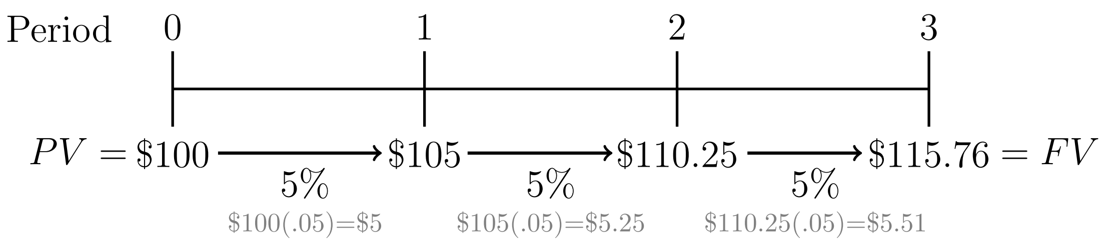
$$FV_n = PV_0(1 + in) \tag{10.5}$$

$$FV_3 = \$100[1 + (.05)(3)] = \$115$$

- Simple interest is most commonly used in auto loans

10.1.2 Compound Interest

- Compound interest is when the interest payment is updated to consider that the money stored in the account changes each period
- **Example:** Suppose I put \$100 in a bank account today that will earn 5% interest compounded annually, how much money will I have after three years?



$$FV_1 = \$100 + \$100(.05) = \$100(1 + .05) = \$105$$
$$FV_2 = \$105(1 + .05) = \$110.25$$
$$FV_3 = \$110.25(1 + .05) = \$115.76$$

From here, we can see that:

$$FV_1 = PV_0(1 + i) \quad (10.6)$$

$$FV_2 = FV_1(1 + i) \quad (10.7)$$

$$FV_3 = FV_2(1 + i) \quad (10.8)$$

If we substitute in FV_2 into Equation (8), we get:

$$FV_3 = FV_1(1 + i)(1 + i) \quad (10.9)$$

Then if we substitute FV_1 into Equation (9), we get:

$$FV_3 = PV_0(1 + i)(1 + i)(1 + i) \quad (10.10)$$

From this, we get a nice Equation that we can use to calculate Future Value when interest is compounded:

$$FV_n = PV_0(1 + i)^n \quad (10.11)$$

From out example:

$$FV_3 = \$100(1 + .05)^3 = \$115.76$$

10.2 Present Value

- Now that we have established how to calculate the future value of a present dollar amount, what if we are discussing a dollar amount in the future?
- Suppose you are asked the question would you prefer \$500 in 4 years or \$400 today?
  - We cannot directly compare money in the future to money today
  - So we must find the present value of the future dollar amount
    - This is called discounting a future value into toady's terms
- We can get the Present Value formula by rearranging the Future Value formula:

$$PV = \frac{FV_n}{(1+i)^n} \tag{10.12}$$

- If we want to compare \$500 in 4 years to \$400 today, need to find the Present Value of the \$500
- If we assume the interest rate is 5% compounded annually:

$$PV = \frac{\$500}{(1+.05)^4} = \$411.35$$

- This means that \$500 in 4 years is worth \$411.35 today, so that means we would be better of taking the \$411.35 in 4 years than \$400 today
- Why? Because if we were to take the \$400 and invest it at 5% compounded annually for 4 years, we would only have:

$$FV_4 = \$400(1.05)^4 = \$486.20$$

Another Example

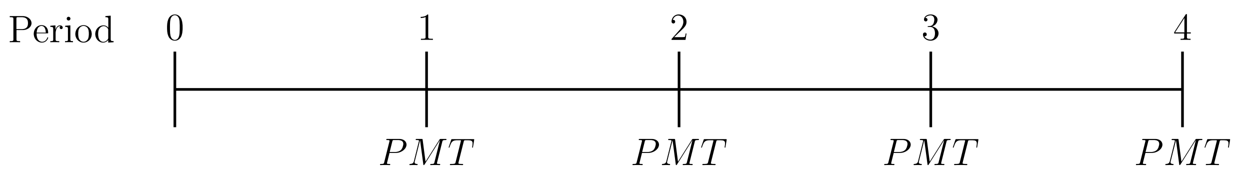
- Another way we can use present value is in the following situation:
  - Suppose you want to buy a car in 3 years that costs \$12,000
  - If you had an account that paid 7% interest compounded annually, how much would you need to put in that account today to have \$12,000 in 3 years?

$$PV = \frac{\$12,000}{(1+.07)^3} = \$9,795.57$$

  - This means that if you were to put \$9,795.57 in an account today at 7% interest compounded annually, you would have \$12,000 in 3 years

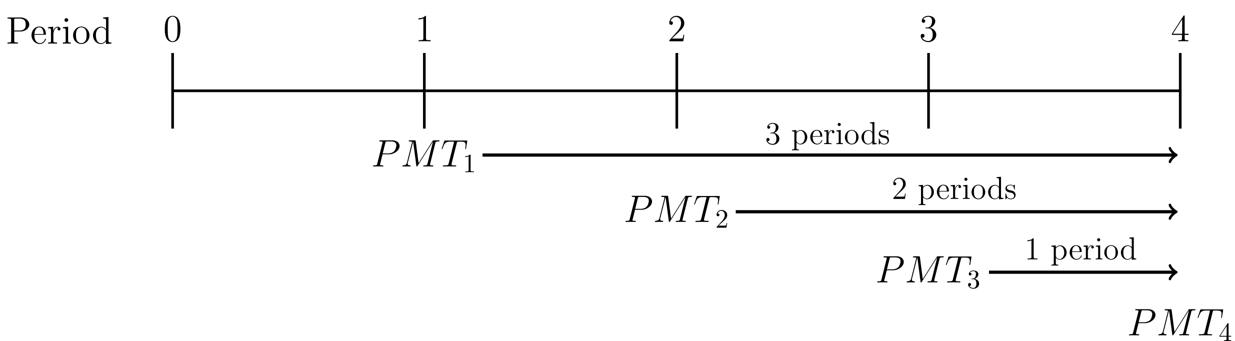
10.3 Annuities

- The previous section focus on lump-sum payments
- While these are interesting problems and sometimes encountered in the real-world, it is more common where payments are made over time while collecting compound interest
- If the payments made are identical, equal dollar amounts, each period, this is called an annuity
- There are two types of annuities:
  - An **ordinary annuity** is an annuity where the payment is made at the end of each period
    - Examples: mortgages, student loans



10.3.1 Future Value of an Ordinary Annuity

- Each deposit will collect interest that is compounded



- PMT_1 collects interest for 3 periods

  $FV = PMT_1(1+i)^3$

- PMT_2 collects interest for 2 periods

  $FV = PMT_2(1+i)^2$

- PMT_3 collects interest for 1 periods

  $FV = PMT_1(1+i)$

- PMT_4 collects interest for 0 periods

  $FV = PMT_4$

$$FVA_{ord} = PMT(1+i)^3 + PMT(1+i)^2 + PMT(1+i) + PMT \tag{10.13}$$

Fortunately for us, there is a simplified equation to calculate the future value on an ordinary annuity:

$$FVA_{ord} = \frac{PMT}{i}[(1+i)^n - 1] \tag{10.14}$$

10.3.2 Present Value of an Ordinary Annuity

- Suppose you win the Powerball lottery worth \$500,000,000
- The \$500,000,000 payment assumes that you will take payments over time
- Typically you will be offered the option to take the winnings as an annuity (say \$7,525,717 per year for 30 years, if $i = .05$), where you will receive the full value, or as a lump sum payment where you will receive a smaller amount
- The lump sum (say it is \$110,500,000) is money you will get today, so it is expressed in the present value
- The \$500,000,000 you will get through an annuity is expressed as a future value
- If we want to compare the two, we will need to discount the value of the annuity back to today (e.g., calculate the present value of the annuity)

$$PVA_{ord} = \frac{PMT}{1+i} + \frac{PMT}{(1+i)^2} + \frac{PMT}{(1+i)^3} + ... + \frac{PMT}{(1+i)^n} \tag{10.15}$$

Fortunatley, this simplifies into a nice expression:

$$PVA_{ord} = \frac{PMT}{i}\left[1 - \frac{1}{(1+i)^n}\right] \tag{10.16}$$

From our example:

- Lump sum payment: $PV = \$110,500,000$
- Annuity: $PVA_{ord} = \frac{7,525,717}{.05}\left[1 - \frac{1}{(1.05)^{30}}\right] = \$115,688,716$
- This means that the annuity is worth more than the lump sum

10.4 Amortized Loans

- When you borrow money (e.g., student loans, mortgages, business loans, etc.), you pay it back in installments over time
- We can use the concept of annuities to understand how your monthly payment is calculated
- When you borrow money, the initial amount you borrow is called the principal balance
- As you make regular loan payments, you are paying down the principal and paying interest
    - Interest is usually expressed as an Annual Percentage Rate (APR)
- To find out what the regular payment is, we use the PVA_{ord} formula where the $PVA =$ Principal, $n =$ number of years, and $i =$ the APR. We then solve for PMT
- We can identify how much of each regular payment goes towards the principal and to interest by constructing an amortization schedule
- We can use this as a "what-if" analysis to determine what happens if you want to pay back your loan early

Example:

Suppose you want to borrow $15,000 to build an addition on your house. The loan is going to be repaid in three equal payments at the end of the next three years. The bank is offering you an APR of 8%.

To find the payment:

$$15,000 = \frac{PMT}{.08}\left[1 - \frac{1}{(1.08)^3}\right]$$

$$PMT = \frac{15,000}{2.5771} = \$5,820.50$$

Once we have the payment, we can construct the amortization schedule. For each payment, the interest gets paid first, then the remaining amount goes towards the principal

10.4.1 Amortization Schedule

End of Year	Payment	Interest	Principal Repayment	Ending Balance
0				\$15,000
1	\$5,280.50	$15,000(.08)$ $= 1,200$	$5280.50 - 1,200$ $= \$4,620.50$	$15,000 - 4,620.50$ $= \$10,379.50$
2	\$5820.50	$10,379.50(.08)$ $= 830.36$	$5,820.50 - 830.36$ $= \$4,990.14$	$10,379.50 - 4,990.14$ $= \$5,389.36$
3	\$5,820.40	$5,389.36(0.8)$ $= 431.15$	$5,820.50 - 431.15$ $= 5,389.36$	$5,389.36 - 5,389.36$ $= 0.00$