

Chapter 3 – Probability

We have completed our study of Descriptive Statistics and are headed for Statistical Inference. Before we get there, we need to learn a little bit about probability.

3.1 – Basic Concepts of Probability

- Probability** – Probability is a numerical measure of the likelihood of some event occurring. It is measured on a scale from 0 to 1: $0 \leq \text{Probability} \leq 1$. The smaller (larger) the probability the less (more) likely it is that the event will occur. An event with probability 0 never occurs; an event with probability 1 always occurs. Mostly, probabilities are somewhere between these two extremes.
- Example: Consider flipping a fair coin. There are two possible outcomes and we generally assume that they are *equally likely*. Thus, $P(T) = P(H) = 1/2$

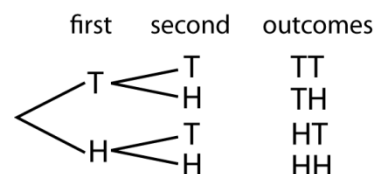


3. **Definitions:**

- Experiment** – a repeatable process whose output varies over a set of possible outcomes.
- Sample space** – The set of *all* possible experimental outcomes.

4. Examples:

- Experiment: flip a coin
Sample space: {H, T}
- Experiment: flip two coins
Sample space: {TT, TH, HT, HH}



Sometimes a *tree diagram* is useful to enumerate a sample space:

- Experiment: roll a die
Sample space: {1, 2, 3, 4, 5, 6}

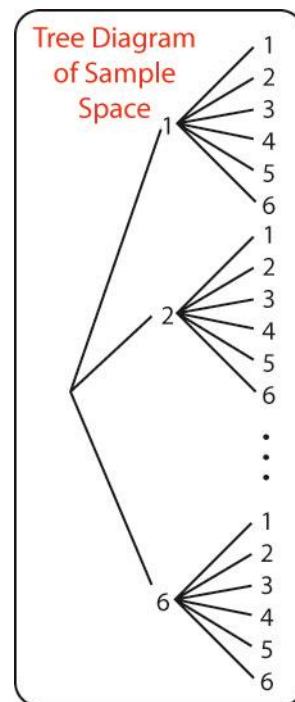
5. Simple Sample Space – Suppose there are n outcomes in a sample space and they are all equally likely then the probability of each outcome is the same, $1/n$.

6. Examples

- Experiment: Roll a six-sided die.

$$P(1) = P(2) = P(3) = P(4) = P(5) = P(6) = 1/6$$

- Experiment: Roll two six-sided dice.

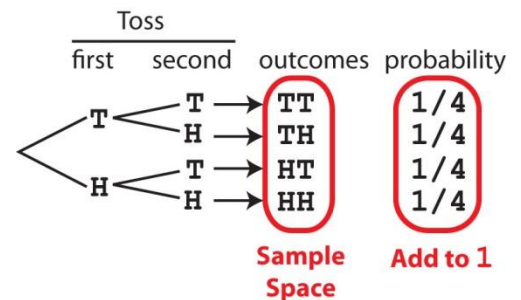


		2nd Roll					
		1	2	3	4	5	6
1st Roll	1	1,1	1,2	1,3	1,4	1,5	1,6
	2	2,1	2,2	2,3	2,4	2,5	2,6
	3	3,1	3,2	3,3	3,4	3,5	3,6
	4	4,1	4,2	4,3	4,4	4,5	4,6
	5	5,1	5,2	5,3	5,4	5,5	5,6
	6	6,1	6,2	6,3	6,4	6,5	6,6

Sample Space

- How many outcomes are there? $36 = 6 \times 6$
- Are they all equally likely? Yes.
- $P(1,1) = P(1,2) = P(2,1) = P(1,3) = \dots = P(6,6) = 1/36$

c. Experiment: Flip two coins



d. Experiment: Pick a person at random from a room with these people in it: A, B, C, D.

Sample space: {A, B, C, D}

What is the probability that C is picked? Answer: 1/4

e. Experiment: Pick two people from A, B, C, D to form a committee.

This experiment is a little different in two ways. First, once we pick the first person, they can't be picked again. Second, the order people are picked doesn't matter. For example picking A first and then B is the same as picking B first and then A. if we draw the tree diagram as shown on the right, then the sample space is:

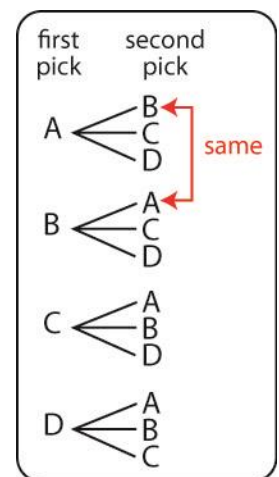
Sample space: {AB, AC, AD, BA, BC, BD, CA, CB, CD, DA, DB, DC }

where each outcome has probability 1/12.

However, noticing that the order doesn't matter, the sample space can be written:

Sample space: {AB, AC, AD, BC, BD, CD }

where each outcome has probability 1/6.



7. **Probability** is the mathematical theory that Statistical Inference is based upon. In this chapter, we learn just enough probability here to give us some insight into how inference works and that might be useful in other ways. Next, we refine earlier definitions and introduce some notation.

8. **Definitions and notation:**

a. **Sample Space** – We denote the sample space as the set of all possible outcomes (simple events) to an experiment, $S = \{E_1, E_2, \dots, E_n\}$, where n is the number of items in the sample space.

b. Probability Requirements for a given sample space:

- All probabilities are between 0 and 1: $0 \leq P(E_i) \leq 1, i = 1, 2, \dots, n$.
- The sum of all probabilities is 1: $\sum_{i=1}^n P(E_i) = 1$.

c. Example – Roll a six-sided die. Let: $E_i = \text{face } i \text{ is observed}$. Thus,

$$S = \{E_1, E_2, E_3, E_4, E_5, E_6\},$$

$$P(E_1) = P(E_2) = P(E_3) = P(E_4) = P(E_5) = P(E_6) = 1/6$$

and finally, we see that if you add up all the probabilities you get 1:

$$\sum_{i=1}^6 P(E_i) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = 1$$

3.2 – Events

1. **Definitions**

a. **Simple Event** – A *simple event* is an individual outcome to an experiment.

b. **Event** – An event is a subset of the sample space which is comprised of a set of simple events.

c. **Probability of an Event** – The probability of an event is the sum of the probabilities of the simple events that comprise it.

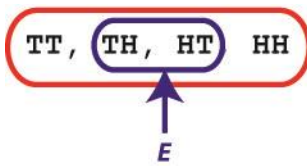
2. **Example** – Find the probability of an even outcome when a six-sided die is rolled.

- Let $A = \text{an even outcome is observed when a die is rolled}$, $A = \{E_2, E_4, E_6\} = \{2, 4, 6\}$.
- $P(A) = P(2) + P(4) + P(6) = \frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{3}{6} = \frac{1}{2} = 0.5$

3. **Example** – Flip two coins.

a. Find the probability of getting a head and a tail

- Let $E = \text{a head and a tail are observed} = \{TH, HT\}$



By Counting:

$$P(E) = P(\text{TH or HT}) = 2/4 = 1/2$$

By a Law of Probability:

$$P(E) = P(\text{TH or HT}) = P(\text{TH}) + P(\text{HT}) \\ = 1/4 + 1/4 = 2/4 = 1/2$$

b. Find the probability of getting at least one head.

- Let $B = \text{at least one head} = \{TH, HT, HH\}$
- $P(B) = P(TH) + P(HT) + P(HH) = \frac{1}{4} + \frac{1}{4} + \frac{1}{4} = \frac{3}{4} = 0.75$

4. **Example** – Choose a card from a standard deck.

- $S = \{2_H, 2_D, 2_C, 2_S, 3_H, 3_D, 3_C, 3_S, \dots\}$

a. Find the probability that an ace of hearts was picked. $P(A_H) = 1/52$.

b. Find the probability that an ace was picked.

- Let $A = \text{an ace was picked} = \{A_H, A_D, A_C, A_S\}$
- $P(A) = P(A_H) + P(A_C) + P(A_D) + P(A_S) = \frac{1}{52} + \frac{1}{52} + \frac{1}{52} + \frac{1}{52} = \frac{4}{52} = \frac{1}{13} = 0.077$

c. Find the probability that a spade was picked.

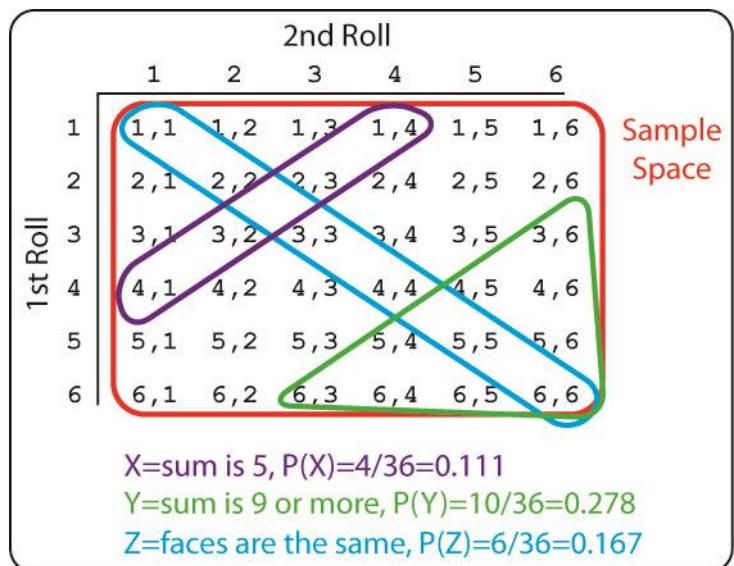
- Let $B = \text{an spade was picked} = \{2_S, 3_S, \dots, K_S, A_S\}$
- $P(B) = P(2_S) + P(3_S) + \dots + P(A_S) = \frac{1}{52} + \frac{1}{52} + \dots + \frac{1}{52} = \frac{13}{52} = \frac{1}{4} = 0.25$

d. Find the probability that an ace or a spade was picked.

- Let $C = \text{an ace or a spade was picked} = \{2_S, 3_S, \dots, K_S, A_S, A_H, A_D, A_C\}$
- $P(C) = \frac{16}{52} = \frac{4}{13} = 0.308$
- Note that $P(C) \neq P(A) + P(B)$

5. **Example** – Roll two six-sided die. Find the probability:

- The sum is 5
- The sum is 9 or more
- The faces are the same

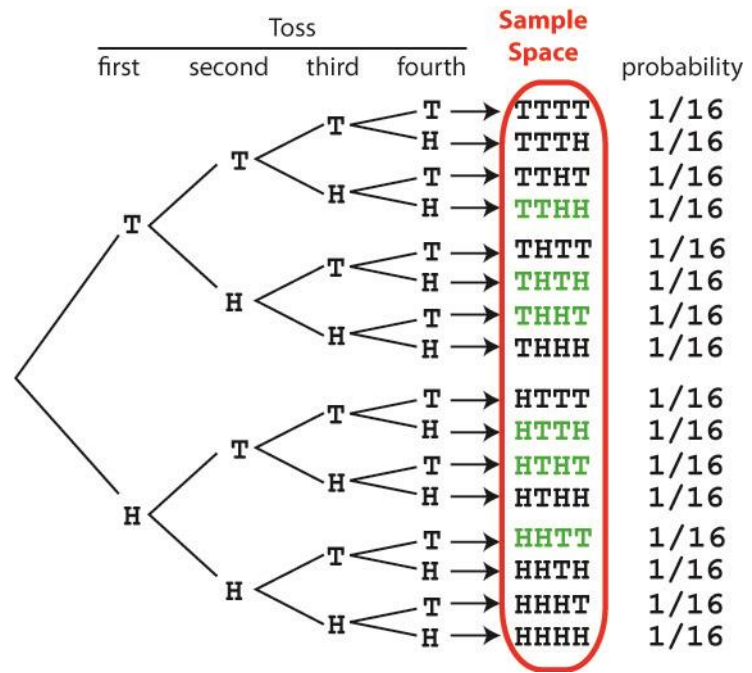


6. **Example** – Flip 4 coins. Find the probability of

a. Exactly 2 heads and 2 tails. (the outcomes shown in green) $P(A) = \frac{6}{16} = 0.375$

b. Exactly 3 heads or exactly 3 tails. $P(B) = \frac{8}{16} = 0.5$

c. No heads (or equivalently, all tails). $P(C) = \frac{1}{16} = 0.0625$



3.3 – Classical Methods of Assigning Probabilities

1. Classical methods of assigning probabilities are based on counting techniques such as the multiplication rule, combinations, and permutations. We have seen simple examples of counting in the previous section.

2. **Multiplication Rule** – Suppose we can think of an experiment as a sequence of S steps. If there are n_1 ways to do the first step, n_2 ways to do the second step, ..., and n_S ways to do the S^{th} step then the total number of ways the S steps can occur is: $n_1 * n_2 * \dots * n_S$.

3. Examples

a. Roll 2 dice. How many possibilities are there? $6 * 6 = 36$

b. Flip 4 coins. How many possibilities are there? $2 * 2 * 2 * 2 = 16$

c. A license plate is composed of 3 letters followed by 3 digits. How many possible license plates are there? $26 * 26 * 26 * 10 * 10 * 10 = 17,576,000$

d. A restaurant has 3 meats, 6 vegetables, and 2 deserts. A *lunch special* is composed of 1 of each. How many different *lunch specials* are there? $3 * 6 * 2 = 36$

4. **Combinations** – A combination is a way of selecting a number of items from a larger group *where the order of the items does not matter and repeats are not allowed*. Suppose there are n choices and we want to choose k of them ($k \leq n$) in such a way that the order doesn't matter. The total number of ways this can be done is:

$$\binom{n}{k} = n \text{ choose } k = \frac{n!}{k!(n-k)!}$$

5. Examples

- a. How many ways can we pick two people from A, B, C, D to form a committee?

As we noted earlier, the order people are picked doesn't matter and each person can be picked only once. So, a *combination* can be used to determine the number of possibilities.

$$\binom{4}{2} = \frac{4!}{2!2!} = \frac{4 * 3 * 2 * 1}{(2 * 1) * (2 * 1)} = 6$$

And those possibilities are: {AB, AC, AD, BC, BD, CD }

- b. Suppose there are 6 vegetables on a buffet. How many ways can you choose 2 different ones.

$$\binom{6}{2} = \frac{6!}{2!(6-2)!} = \frac{6!}{2!4!} = \frac{6 * 5 * 4!}{2!4!} = \frac{6 * 5}{2} = 15$$

- e. A restaurant has 3 meats, 6 vegetables, and 2 deserts. A *lunch special* is composed of 1 meat, 2 vegetables and a desert. How many different *lunch specials* are there? Multiplication rule:

Number of ways to choose a meat		Number of ways to choose 2 vegetables		Number of ways to choose a desert
3	*	$\binom{6}{2}$	*	2

$$= 3 * 15 * 2 = 90$$

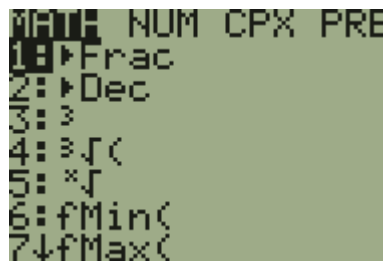
- f. How many ways are there to pick 5 cards from a standard deck? Answer: $\binom{52}{5}$.

We can work this by hand, or use the calculator to answer this question:

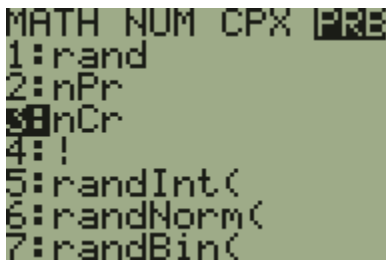
1. Type: 52



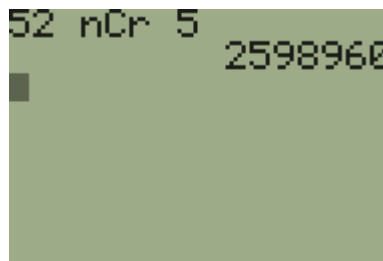
2. Press the *MATH* key



3. Use the right-arrow key to scroll over to *PRB*, then the down-arrow key to scroll down to *nCr*. Press *Enter*.



4. Type: 5 and press *Enter*.



g. Pick 2 cards from a standard deck. How many ways are there to pick 2 Aces? $\binom{4}{2} = 6$

e.g. $\{(A_H, A_D), (A_H, A_C), (A_H, A_S), (A_D, A_C), (A_D, A_S), (A_C, A_S)\}$

h. Pick 2 cards from a standard deck. How many different 2 card hands are there? $\binom{52}{2} = 1326$

i. Pick 2 cards from a standard deck. What is the probability of choosing exactly 2 Aces?

$$\frac{\# \text{ ways to pick 2 Aces}}{\# \text{ ways to pick 2 cards}} = \frac{\binom{4}{2}}{\binom{52}{2}} = \frac{6}{1326} = 0.00452$$

j. Pick 5 cards from a standard deck. What is the probability of choosing exactly 2 Aces?

$$\frac{\# \text{ ways to pick 2 Aces} * \# \text{ ways to pick 3 other cards}}{\# \text{ ways to pick 5 cards}} = \frac{\binom{4}{2} \binom{48}{3}}{\binom{52}{5}} = \frac{6 * 17,296}{2,598,960} = \frac{103,776}{2,598,960} = 0.0399$$

5. **Permutations** – A permutation is a way of selecting a number of items from a larger group *where the order of the items does matter and repeats are not allowed*. Suppose there are n choices and we want to choose k of them ($k \leq n$) in such a way that the order does matter. The total number of ways this can be done is:

$$P_{n,k} = \frac{n!}{(n-k)!}$$

6. Examples

a. How many ways are there of picking the first 4 batters from the starting 9 in a baseball game?

$$P_{9,4} = \frac{9!}{(9-4)!} = \frac{9!}{5!} = \frac{9 * 8 * 7 * 6 * 5!}{5!} = 9 * 8 * 7 * 6 = 3,024$$

Alternatively, you can think of this as the multiplication rule:

$$\frac{9}{\# \text{ ways to pick 1}^{\text{st}} \text{ person}} * \frac{8}{\# \text{ ways to pick 2}^{\text{nd}} \text{ person}} * \frac{7}{\# \text{ ways to pick 3}^{\text{rd}} \text{ person}} * \frac{6}{\# \text{ ways to pick 4}^{\text{th}} \text{ person}}$$

b. There are 20 people in a club. How many ways can you choose 1 chairperson, 1 secretary and 1 treasurer?

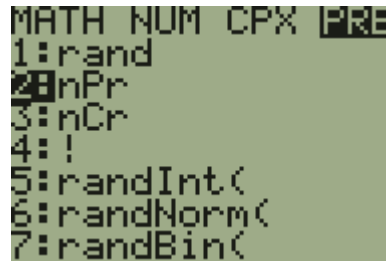
$$P_{20,3} = \frac{20!}{(20-3)!} = \frac{20!}{17!} = \frac{20 * 19 * 18 * 17!}{17!} = 20 * 19 * 18 = 6,840$$

Or, we can use the calculator to answer this question:

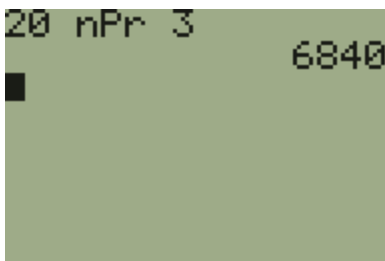
1. Type: 20



2. Press the *MATH* key. Then, use the right-arrow key to scroll over to *PRB*, then the down-arrow key to scroll down to *nPr*. Press *Enter*.



3. Type: 3 and press *Enter*.



- c. There are 10 children in a room. How many ways can they be lined up?

$$P_{10,10} = \frac{10!}{(10 - 10)!} = \frac{10!}{0!} = \frac{10!}{1} = 10! = 3,628,800$$

3.4 – Frequency Method of Assigning Probabilities

1. The probability examples and techniques we have seen so far are *exact* probabilities. Frequently, there is not such a convenient method to obtain exact probabilities. Another technique for assigning probabilities is to use the *frequency (or empirical)* approach. The idea is to look at past (historical) data and use it to *estimate* probabilities.
2. Example – Suppose that a Company X places orders with Company Y frequently and that you would like to determine the probability of when an order will be delivered. To do this you go back into your records and find the last 75 orders that have been placed. What you find is shown in the table below.

Days to Receive Order	Number of Orders	Probability
1	6	6/75 = 0.08
2	24	24/75 = 0.32
3	30	30/75 = 0.40
4	15	15/75 = 0.20
	n=75	sum=1

3. Example – Suppose an order is late if it arrives on the 3rd or 4th day. Thus,

$Late\ Order = \{ 3\ days, 4\ days \}$ is an *Event*.

And the probability is:

$$P(Late\ Order) = P(3\ days) + P(4\ days) = 0.4 + 0.2 = 0.6$$

We can think of this in terms of an experiment and a sample space:

Experiment: Place an order

Sample space: { 1 day, 2 days, 3 days, 4 days }

We will mostly use this approach.

3.5 – Other Methods of Assigning Probabilities

- Subjective** – Sometimes we estimate probabilities using intuition, knowledge of a situation. For instance, an investment banker may state that there is a 75% chance that the price of a stock will go up in the next month.
- Probability models** – Sometimes experiments follow a *pattern*. When they do, we can use a formula to calculate probabilities. This will be covered in chapters 4 and 5 where we consider the Binomial distribution and the Normal distribution.

Homework 3.1

- Over the course of 20 games, a basketball player has taken 186 2-point shots and made 103 of them. What is the approximate probability that this player will make a 2-point shot?
- A driver for a trucking company gets a per diem for any trips 500 miles or more. Over the last 3 years, the driver has taken trips whose distance is distributed as shown in the table below. What is an estimate of the probability of a trip earning a per diem?

Distance	100-249	250-499	500-749	750-999	1000-1249	1250-1500
Num Trips	48	122	23	55	33	17

- There are 6 ways to travel from Valdosta to Augusta, 3 ways to travel from Augusta to Columbia, SC, and 5 ways to travel from Columbia to Charlotte, NC. How many possible ways are there to travel from Valdosta to Charlotte?
- A team of 9 people is following a path in the woods as they search for someone lost. At noon, they decide that 4 will continue on the path and 5 will remain behind and set up a base camp. How many ways can the 4 people be chosen to continue searching the paths?
- A team of 9 people is searching for someone lost in the woods. They come to a junction where there are 4 paths leading in different directions. It is decided that 5 people will wait at the junction and the other 4 will continue the search, each on one of the 4 different paths. How many ways can the 4 people be chosen to continue searching the paths? Since the paths are different, sending a person, say John, down the first path

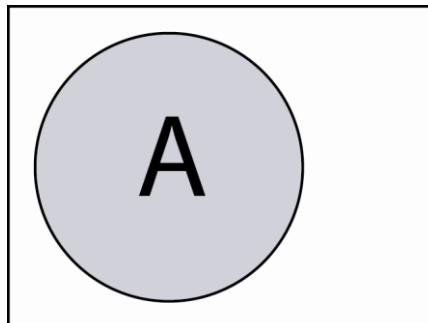
is a different possibility than sending him down the 2nd (or 3rd or 4th path).

6. How many 7-digit phone numbers are possible if the first digit cannot be a 1 or 0?
7. How many distinct bridge hands are possible? A bridge hand is 13 cards chosen from a standard deck.
8. There are 10 sodas and 5 beers in a cooler. Pick 2 cans at random. What is the probability that you pick 2 beers?

3.6 – Venn Diagrams

1. **Venn diagrams** – A graphical technique that is useful for calculating probabilities. We use a box to represent the sample space and circles to represent events. The box contains *all* the probability, so we say the box has a probability of 1.

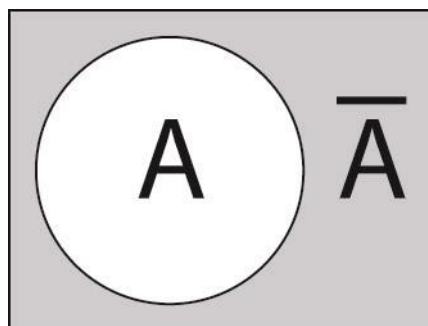
Example: Let $A = \text{person votes for Sanders}$. Suppose that $P(A) = 0.65$.



2. **Complement** – If A is an event, then \bar{A} (some books use A' or A^c) is defined to be *everything that is not in A* . \bar{A} is referred to as “*A complement*” or “*not A*”. It is easy to see that if we add A and \bar{A} we get 1: $P(A) + P(\bar{A}) = 1$

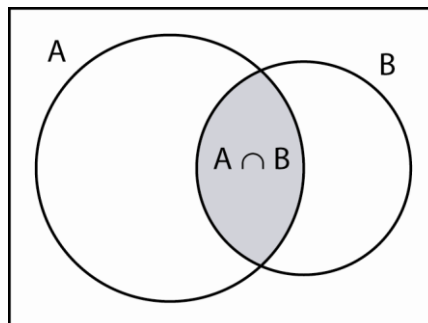
Thus, the complement rule states that: $P(\bar{A}) = 1 - P(A)$

Example: Let $\bar{A} = \text{person does not vote for Sanders}$. Thus, $P(\bar{A}) = 1 - P(A) = 1 - 0.65 = 0.35$.



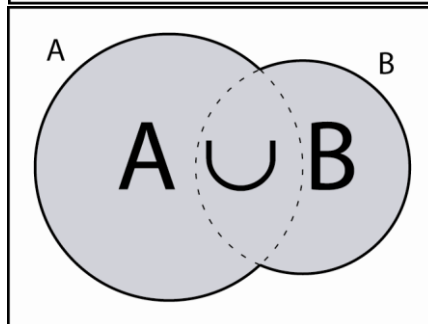
3. **Intersection** – When we have two events, A and B , we denote their intersection by $A \cap B$. The intersection of two events is *everything that is in A and in B* .

Example: Let $B = \text{person has a college degree}$. Then $A \cap B$ is the event where someone will vote for Sanders *and* has a college degree. Suppose that



4. **Union** – The union of two events is denoted by $A \cup B$ and is defined to be *everything in A or in B (or in both)*.

Example: $A \cup B$ is the event where someone will vote for Obama *or* has a college degree (or both).



5. **Addition Rule** – tells us how to calculate the probability of the *union* of two events:

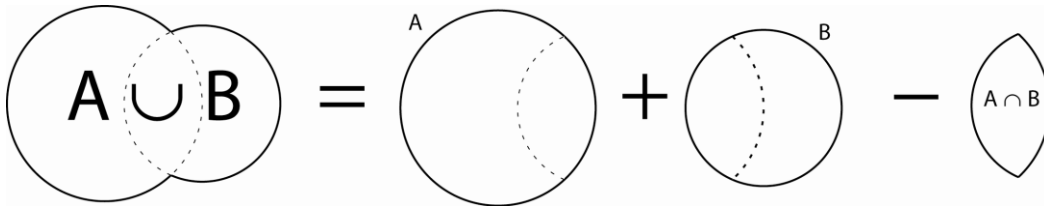
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

a. Example: Suppose that: $P(A) = 0.65$, $P(B) = 0.45$, $P(A \cap B) = 0.20$. Then,

$$P(A \cup B) = 0.65 + 0.45 - 0.20 = 0.90$$

Thus, 90% of people will vote for Sanders or have a college degree.

b. Explanation of formula: We subtract the intersection because it has already been added twice:



c. Similarly, you can rearrange this formula using algebra to get the formula shown below. We will see how to use this in later examples.

$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$

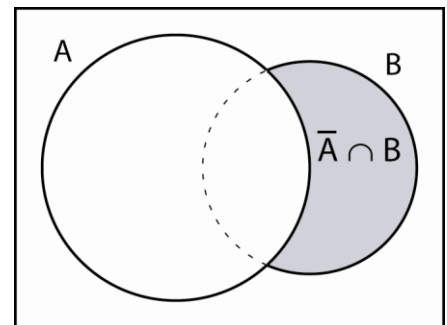
6. **Other events of interest:**

a. $\bar{A} \cap B$ is referred to as “only B” or “B but not A”. It is defined to be *everything that is in B but not in A*.

Example: $\bar{A} \cap B$ is the event someone has a college degree but did not vote for Sanders. The probability is:

$$P(\bar{A} \cap B) = P(B) - P(A \cap B) = 0.45 - 0.20 = 0.25$$

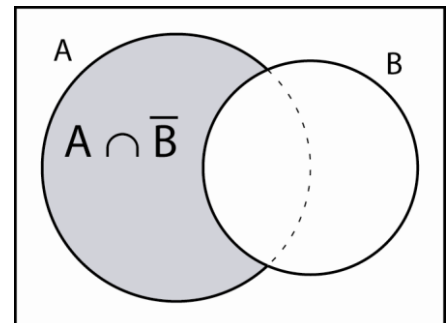
Thus, 25% of people have a college degree but did not vote for Sanders.



b. $A \cap \bar{B}$ is referred to as “only A” or “A but not B”. It is defined to be *everything that is in A but not in B*.

Example: $A \cap \bar{B}$ is the event someone votes for Sanders but does not have a college degree. The probability is:

$$P(A \cap \bar{B}) = P(A) - P(A \cap B) = 0.65 - 0.20 = 0.45$$



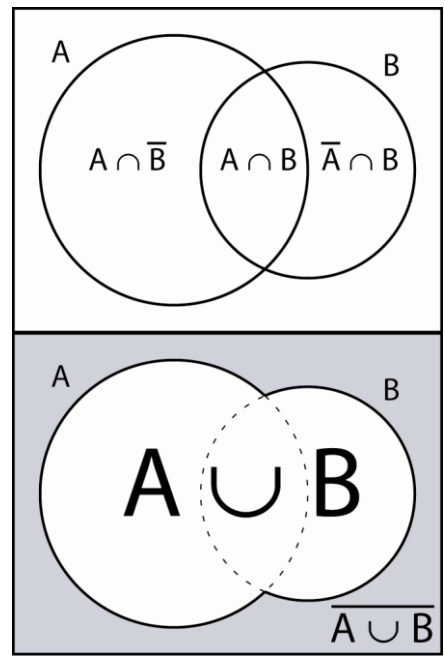
c. Thus, we can see that the Union is composed of three pieces.

- $A \cap \bar{B}$ = vote for Sanders but do not have a college degree.
- $A \cap B$ = vote for Sanders and have a college degree.
- $\bar{A} \cap B$ = didn't vote for Sanders but do have a college degree.

d. $\overline{A \cup B}$ is referred to as *neither A nor B*. It is defined to be *everything that is not in the union* (or as having neither characteristic A nor B).

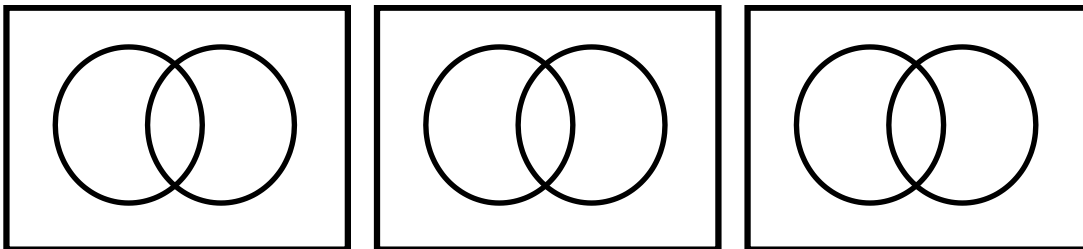
Example: from above. $\overline{A \cup B}$ is the event that someone neither votes for Sanders nor has a college degree. We use the complement rule to calculate the probability of this event:

$$P(\overline{A \cup B}) = 1 - P(A \cup B) = 1 - 0.90 = 0.10$$

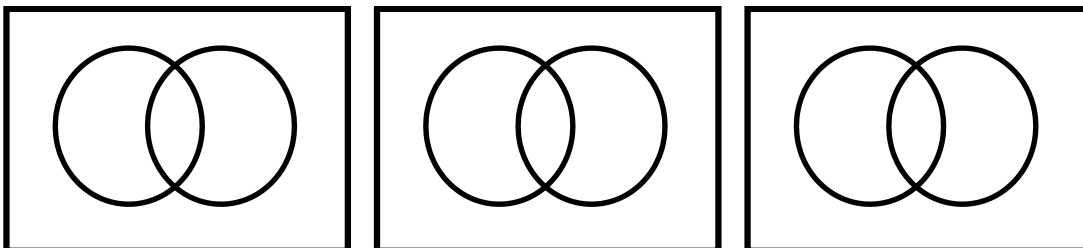


7. Examples:

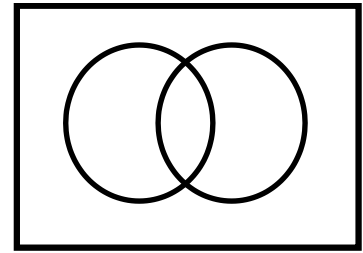
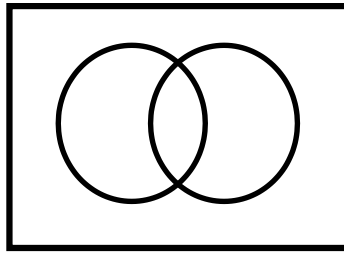
- There are two machines in a shop. Machine A is down (not working) 15% of the time, Machine B is down 20% of the time, and both are down 5% of the time. What is the probability that...
 - either machine is down?
 - neither machine is down?
 - A is not down?
 - only A is down?
 - machine B is down, but not machine A?
 - the only machine down is A?



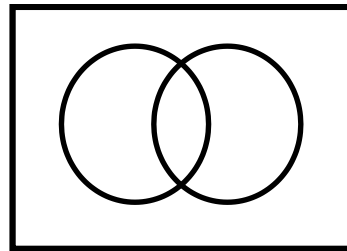
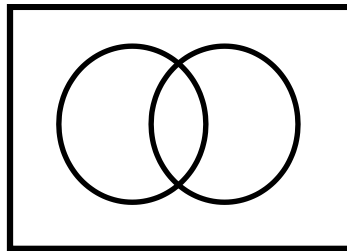
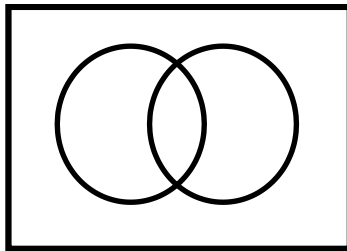
- Two types of defects occur on a product. Type A defects occur with probability 0.05, Type B with probability 0.07, and at least one type of defect occurs with probability 0.09. What fraction of products have ...
 - both types of defects occur?
 - only Type A defects occur?
 - neither type of defect occurs?
 - Type B defects but not Type A defects?



3. 500 people were sampled to see investigate the fraction of people who had had either of two diseases. The data showed that 50 had disease A, 100 had disease B and 30 had both. What percentage of people had...
- either disease?
 - exactly one disease?
 - neither disease
 - did not have disease B?

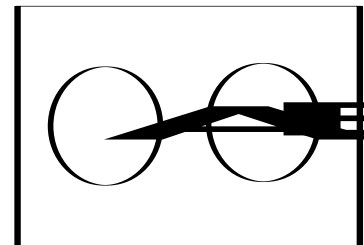


4. In a recent survey of homeowners, testing to see how what percentage own dogs and cats, it was found that 10% had only a dog as compared to 5% that had only a cat, while 60% had neither a a dog nor a cat. What percentage of homeowners
- have either a dog or a cat?
 - both a dog and a cat?
 - a dog?
 - a cat?



8. **Mutually Exclusive** – Two events are mutually exclusive if they have no intersection.

- a. Example - Pick a card at random from a standard deck. Let A be the event that you draw a *Heart* and B be the event that you draw a *Spade*. The two events, A and B are mutually exclusive because they have no intersection. In other words, if you draw a heart, there is no way you can draw a spade and *visa-versa*.



The probability of the union of two mutually exclusive events is simply the sum of the two probabilities:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = P(A) + P(B) - 0 = P(A) + P(B)$$

Thus,

$$P(\text{Heart} \cup \text{Spade}) = P(\text{Heart}) + P(\text{Spade}) = \frac{13}{52} + \frac{13}{52} = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

Homework 3.2

1. A company sells two products: gasoline and soft drinks. 90% of people who enter the store purchase gasoline and 20% purchase a soft drink. 15% of people purchase both gasoline and a soft drink. (The people who don't purchase anything are simply asking directions.) What percentage of people purchase... (a) either gasoline or a soft drink? (b) gasoline but not a soft drink? (c) a soft drink but not gasoline? (d) What percentage of people are simply asking directions (*i.e.* not purchasing anything)? (e) only gas?
2. A survey of 200 people revealed that 140 people have neither a dog nor a cat, 20 have a cat but not a dog, and 30 have a dog but not a cat. What percentage of people have... (a) only a dog? (b) only a cat? (c) ...either a dog or a cat? (d) ...a dog and a cat? (e) ...a dog? (f) ...a cat?
3. The probability that a person owns a car is 0.7 and the probability that a person owns a home is 0.4. The probability that a person owns a car or a home is 0.9. What is the probability that a person... (a) ...owns a home and a car? (b) ...owns neither a home nor a car? (c) ...owns only a home? (d) ...owns only a car? (e) ...owns either a car or a home, but not both?
4. 70% of people play a musical instrument, 40% play a sport, and 20% play neither an instrument nor a sport. What percentage of people (a) play an instrument or a sport (b) play an instrument and a sport? (c) only play a sport? (d) don't play a sport?
5. Approximately 10% of children have had measles, but not small pox. 5% of children have had both measles and small pox. Finally, 7% of children have had small pox. What percentage of children have had (a) measles? (b) measles or small pox? (c) small pox but not measles (d) neither measles nor small pox? (e) exactly one of the diseases.
6. Roll two 6-sided dice. Find the probability that the faces are the same or the sum of the faces is 9 or larger.
7. Pick a card from a standard deck. Find the probability of obtaining a heart or a king.

3.7 – Conditional Probability

1. **Idea:** Conditional probability is concerned with how to calculate probabilities *when we know certain information about the outcome of the experiment.*

a. Example:

Regular	Pick 2 cards from a deck. What is the probability of 2 Aces?
Conditional	Deal two cards from a deck, the first face-up and the second face-down. If the face-up card is an Ace what is the probability the face-down card is an Ace?

b. Example:

Regular	Pick someone at random. What is the probability that they support gun control?
Conditional	Pick someone at random. They tell you they are a Democrat. What is the probability they support gun control?

c. Example:

Regular	What is the probability a machine will last 10 years?
Conditional	What is the probability a machine will last 10 years given that it has already lasted 4 years?

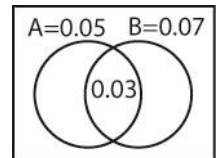
2. **Definition** – The conditional probability of event A given event B is:

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

The phrase *given event B* means that B is a fact, it has definitely occurred. Using this definition, we can also write:

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

3. **Example:** Two types of defects occur on a product. Type A defects occur with probability 0.05, Type B with probability 0.07, and both types of defects occur with probability 0.03.



a. Given that a Type B defect occurred, what is the probability that a Type A defect occurred?

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{0.03}{0.07} = 0.43 = 43\%$$

b. Given that a Type A defect occurred, what is the probability that a Type B defect occurred?

$$P(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{0.03}{0.05} = 0.6 = 60\%$$

4. **Example:** Roll two dice. Given that the sum is 9 or higher, what is the probability that the sum is 10 or higher? See diagram on page 77 if necessary.

$$\begin{aligned} P(\text{sum} \geq 10 | \text{sum} \geq 9) &= \frac{P(\text{sum} \geq 10 \cap \text{sum} \geq 9)}{P(\text{sum} \geq 9)} = \frac{P(\text{sum} \geq 10)}{P(\text{sum} \geq 9)} \\ &= \frac{6/36}{10/36} = 0.6 = 60\% \end{aligned}$$

3.7 – Independence

1. **Independence** – Two events are independent if they are unrelated. In other words, if knowledge of the outcome of one event in no way influences the probability of the outcome of the other. Mathematically, two events, A and B are independent if:

$$P(A|B) = P(A).$$

2. **Example:** Using the data from Example 3 above, are Type A and Type B defects independent?

No, because $P(A|B) = 0.43$ and $P(A) = 0.05$, or because $P(B|A) = 0.6$ and $P(B) = 0.07$

3. **Example:** Roll two dice, one-at-a-time. Dice rolls (and coin tosses) are independent. It is obvious! If you roll a 6 on the first roll there is no way that can influence the second roll.

Suppose you have rolled four 6's in a row. What is the probability that the 5th roll is a 6? The rolls are independent so the probability is 1/6. The previous 4 sixes don't influence the 5th roll.

4. **Example:** Cards are generally not independent. For example, let:

$$A = \text{pick an ace from a deck of cards}$$

$$B = \text{pick an jack from a deck of cards}$$

Then:

$$P(A|B) = 4/51 \neq P(A) = 4/52$$

Note, this is obvious too! If you pick a jack from the deck, of course that will change the probability of obtaining an ace because there is one less card in the deck.

5. If A and B are independent then $P(A \cap B) = P(A)P(B)$

- a. Example: Flip two coins. What is the probability of 2 heads?

$$A = \text{head on first flip}$$

$$B = \text{head on second flip}$$

$$\text{A and B are independent, so: } P(\text{two heads}) = P(A \cap B) = P(A)P(B) = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

- b. Example: A batter has a 0.350 batting average. Suppose each time the batter goes to the plate, the outcome is independent of the others. What is the probability that in the batters next 4 at-bats, he gets first a hit, then another hit, followed by an out, and a hit?

$$P(H_1 \cap H_2 \cap \bar{H}_3 \cap H_4) = 0.35 \cdot 0.35 \cdot (1 - 0.35) \cdot 0.35 = 0.028$$

Note: this question is different from: What is the probability that the player gets 3 hits in his next 4 at bats.

Homework 3.3

1. A company sells two products: gasoline and soft drinks. 90% of people who enter the store purchase gasoline and 20% purchase a soft drink. 15% of people purchase both gasoline and a soft drink. (The people who don't purchase anything are simply asking directions.) (a) Suppose that someone purchases gas in the store, what is the probability that they will also purchase a soft drink? (b) Suppose that someone purchases a soft drink in the store, what is the probability that they will also purchase gasoline? (c) Are *purchasing gasoline* and *purchasing a soft drink* independent? Why or why not?
2. A survey reveals that 50% of Americans own a home, 20% own a handgun, and 10% own a home and a handgun. (a) Suppose that someone owns a gun, what is the probability that they also own a home? (b) Are handgun ownership and home ownership independent? Why or why not?
3. 70% of people play a musical instrument, 40% play a sport, and 20% play neither an instrument nor a sport. (a) Suppose that you see someone playing basketball, what is the probability that they also play a musical instrument? (b) Suppose that you see someone playing a guitar, what is the probability that they also play a

- sport? (c) Suppose that you see someone playing a guitar, what is the probability that they do not play a sport?
- Approximately 15% of children have had measles, 7% small pox, and 5% of children have had both measles and small pox. Are measles and small pox independent? Justify your answer.
 - Pick a card from a standard deck. Suppose someone looks at the card and tells you it is a King, what is the probability that the card is a Heart?
 - Pick a card from a standard deck. Suppose someone looks at the card and tells you it is a Heart, what is the probability that the card is a King?
 - There are 2 kinds of soda in a cooler: 5 Cokes and 10 Diet-Cokes. Also, there are 10 beers in the cooler. A friend picks one can at random and tells you it is a soda. What is the probability that it is a Diet-Coke?
 - Roll a die 3 times. What is the probability that you get
 - you get a 1 on first roll, 4 on second, and a 3 on third?
 - three 6's?
 - first roll is 3 or more, second roll is 2 or more, and third roll is 3 or less?

3.8 – Two-way Tables

- A two-way table is essentially the same as Venn diagrams, only we arrange the data in a table as opposed to the Venn diagram. The “two” refers to the fact that we are dealing with two variables. In the example below, the two variables are:

$$Sex \in \{Male, Female\}$$

$$Promotion Status \in \{Promoted, Not Promoted\}$$

- Example:** A major metropolitan police force in the eastern U.S. consists of 1200 officers - 960 men and 240 women. Over the past 2 years, 324 officers on the police force have been awarded promotions. The two-way table below summarizes the breakdown for male and female officers. After reviewing the data, a committee of female officers charged discrimination on the basis that 288 male officers had received promotions, whereas only 36 female officers had received promotions. The police administration countered with the argument that the relatively low number of promotions for female officers was not due to discrimination but due to the fact that there are fewer female officers on the police force.

	Male	Female	Totals
Promoted	288	36	324
Not Promoted	672	204	876
Totals	960	240	1200

In analyzing a two-way table, the first thing we do is find the *joint* and *marginal probabilities*.

	Male	Female	Totals
Promoted	$\frac{288}{1200} = 0.24$	$\frac{36}{1200} = 0.03$	$\frac{324}{1200} = 0.27$
Not Promoted	$\frac{672}{1200} = 0.56$	$\frac{204}{1200} = 0.17$	$\frac{876}{1200} = 0.73$
Totals	$\frac{960}{1200} = 0.8$	$\frac{240}{1200} = 0.2$	

Joint Distribution of promotion status & sex

	Male	Female	Totals
Promoted	0.24	0.03	0.27
Not Promoted	0.56	0.17	0.73
Totals	0.80	0.20	

Marginal Distribution of sex

Marginal Distribution of promotion status

Using some notation, we interpret these probabilities as shown below:

	Male (M)	Female (\bar{M})	Totals
Promoted (P)	$P(M \cap P)$	$P(\bar{M} \cap P)$	$P(P)$
Not Promoted (\bar{P})	$P(M \cap \bar{P})$	$P(\bar{M} \cap \bar{P})$	$P(\bar{P})$
Totals	$P(M)$	$P(\bar{M})$	

3. **Example** – If you pick someone at random from this police force, what is the probability that you pick...

a. ...a woman?	h. ...a woman who was not promoted?
b. ...a man?	i. ...someone who was promoted or a woman?
c. ...someone who was promoted?	j. ...someone who was not promoted or a woman?
d. ...a person who was not promoted?	k. ...someone who was promoted or a man?
e. ...a man who was promoted?	l. ...someone who was not promoted or a man?
f. ...a man who was not promoted?	
g. ...a woman who was promoted?	

Suppose you pick someone at random from the police force and discover that they...

- m. ...are a woman, what is the probability that she was promoted?
- n. ...are a woman, what is the probability that she was not promoted?
- o. ...are a man, what is the probability that he was promoted?
- p. ...are a man, what is the probability that he was not promoted?
- q. Are promotion status and sex independent? Why?

Homework 3.4

1. The data shown on the right was collected about voter preference for males and females. Each person in the survey was asked whether they would vote for McCain or Obama, and their sex was noted.

	Male	Female
McCain	420	140
Obama	490	350

Suppose that you pick someone at random, what is the probability that the person...

a. ...votes for McCain?	f. Suppose that you pick a person at random and observe that they are a female, what is the probability that the person voted for Obama?
b. ...is a female that votes for Obama?	g. Suppose that you pick a person at random and discover that they voted for McCain, then what is the probability that the person is a Male?
c. ...is a male who didn't vote for McCain?	h. Are voter's preference and sex independent? Why or why not?
d. ...is a female?	
e. ...is a female or voted for Obama?	

2. A survey of 207 college students is conducted to find the numbers of students that have cell phones and cars. The survey reveals that 42 have a car and a cell phone; 126 have a car but not a cell phone; 28 do not have a car but do have a cell phone; and 11 have neither a car nor a cell phone.

a. Make a two-way table to represent this data.	Suppose that you pick someone at random and the person tells you they...
Suppose you pick someone at random, what is the probability that the person...	h. have a car. What is the probability that they have a cell phone?
b. ...has a cell phone?	i. have a cell phone. What is the probability that they have a car?
c. ...has a car?	j. do not have a car. What is the probability that they have a cell phone?
d. ...does not have a cell phone?	k. do not have a car. What is the probability that they do not have a cell phone?
e. ...has a cell phone and a car?	l. Are owning a car and having a cell phone independent? Why?
f. ...has a cell phone but not a car?	
g. ...does not have a car or a cell phone?	

Chapter 4 – Discrete Probability Distributions

In this chapter we get slightly more formal about probability.

4.1 – Random Variables

- Random Variable (RV)** – A *random variable* is a numerical description of the outcome of an experiment.
- Example – When we flip two coins, the sample space is: $\{TT, TH, HT, HH\}$.

Suppose we let $X =$ the number of Heads in two flips of a coin.

X is called a *discrete random variable* and takes on the values $\{0, 1, 2\}$. Note that X is a *numerical description* of the outcome of this experiment, it *counts* the number of Heads in two coin tosses. Thus:

$$P(X = 0) = P(TT) = \frac{1}{4}$$

$$P(X = 1) = P(TH \text{ or } HT) = \frac{1}{4} + \frac{1}{4} = \frac{1}{2}$$

$$P(X = 2) = P(HH) = \frac{1}{4}$$

Notice that these three probabilities add up to 1.

- Example – Let $X =$ number of heads in 4 coin tosses

$$X \in \{0, 1, 2, 3, 4\}$$

What are the probabilities of each of the values the random variable can take on?

Sample Space	# of Heads		
TTTT	0	$P(X=0) = P(TTTT)$	$=1/16$
TTTH	1	$P(X=1) = P(TTTH, TTHT, THTT, HTTT)$	$=4/16$
TTHT	1		
TTHH	2		
THTT	1		
THTH	2	$P(X=2) = P(TTHH, THTH, THHT, HTTH, HTHT, HHTT)$	$=6/16$
THHT	2		
THHH	3		
HTTT	1	$P(X=3) = P(THHH, HTHH, HTHH, HHHT)$	$=4/16$
HTTH	2		
HTHT	2	$P(X=4) = P(HHHH)$	$=1/16$
HTHH	3		
HHTT	2	Sum=16/16=1	
HHTH	3		
HHHT	3		
HHHH	4		

3. **Discrete Random Variable** – A *discrete RV* takes on a finite number of values. Thus, we can count all the values, there is separation between the values. Many times, a discrete random variable is a *count* of something. Examples:

- Let X = the number of defects on a product, X can be: $\{0, 1, 2, \dots\}$
- Let X = the number of boxes of cereal out of 50 that have more than 5 fat grams per serving. Thus, X can be $\{0, 1, 2, \dots, 49, 50\}$
- Let X = the sum of the faces of two dice that are rolled. Thus, X can be $\{2, 3, 4, \dots, 10, 11, 12\}$
- Let X = the number of accidents at an intersection in an hour. Thus, X can be $\{0, 1, 2, \dots\}$

4. **Continuous Random Variable** – A *continuous RV* can take on *ALL* values in a given interval. Many times, a continuous random variable is a scientific measurement. We will study these in Chapter 6. Examples:

- Let X = lifetime of a light bulb. Thus, X might take on any value in [1 hour, 5000 hours]
- Let X = the height of a plant. Thus, X might take on any value in [6in, 18in]
- Let X = the weight of a person. Thus, X might take on any value in [90 lbs, 400 lbs]
- Let X = the time between successive accidents. Thus, X might take on any value in [0min, 3000 min]
- Let X = the fat grams in a serving of cereal. Thus, X might take on any value in [2.5 gms, 4.5 gms.]
- Let X = the gas mileage of a car. Thus, X might take on any value in [24 mpg, 33 mpg]

4.2 – Calculating Discrete Probabilities

1. **Discrete Probability Function, $f(x)$** – provides the probabilities of all the values a random variable can take on. For a discrete RV, the probability that X takes on a particular value is simply the function evaluated at that value:

$$P(X = c) = f(c)$$

2. **Example** – Suppose we have a probability function, $f(x) = \frac{x}{6}$, $x = 1,2,3$. This tells us that the discrete random variable takes on exactly the value 1,2, or 3 and that:

$$\begin{aligned}P(X = 1) &= f(1) = 1/6 \\P(X = 2) &= f(2) = 2/6 \\P(X = 3) &= f(3) = 3/6\end{aligned}$$

3. **Properties of a discrete probability function:**

1. $0 \leq f(x) \leq 1$, as we said earlier, probabilities are always between 0 and 1.
2. $\sum f(x) = 1$, as we said earlier, all the probabilities in the sample space must add to 1.
4. Probability functions are given in a table, graph, mathematical formula, or in words.

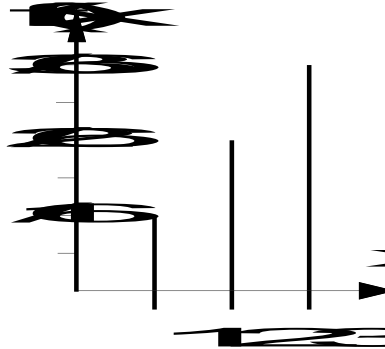
Example – Note that the four examples below all convey the same probability function.

1. Formula $f(x) = \frac{x}{6}$, $x = 1,2,3$

2. Table

x	$f(x)$
1	1/6
2	1/3
3	1/2

3. Graph



4. Words

A random variable takes on values 1, 2, and 3 with probabilities 1/6, 1/3, and 1/2, respectively.

5. Important Rules for Calculating Probabilities for Discrete Random Variables.

Example: $f(x) = \frac{x}{42}$, $x = 2, 4, 6, 8, 10, 12$

a. $P(X = c) = f(c)$

- $P(X = 8) = f(8) = 8/42$

b. $P(X \leq c) = \sum_{x \leq c} f(x)$

- $P(X \leq 8) = f(2) + f(4) + f(6) + f(8) = \frac{2}{42} + \frac{4}{42} + \frac{6}{42} + \frac{8}{42} = 20/42$

c. $P(X < c) = \sum_{x < c} f(x)$

- $P(X < 8) = f(2) + f(4) + f(6) = \frac{2}{42} + \frac{4}{42} + \frac{6}{42} = 12/42$

- Note: $P(X < 8) = P(X \leq 6)$ for this particular problem.

d. $P(X \geq c) = \sum_{x \geq c} f(x) = 1 - P(X < c)$

- $P(X \geq 8) = f(8) + f(10) + f(12) = \frac{8}{42} + \frac{10}{42} + \frac{12}{42} = 30/42$

or

- $P(X \geq 8) = 1 - P(X < 8) = 1 - [f(2) + f(4) + f(6)] = 1 - \frac{12}{42} = 30/42$

e. $P(X > c) = \sum_{x > c} f(x) = 1 - P(X \leq c)$

- $P(X > 8) = f(10) + f(12) = \frac{10}{42} + \frac{12}{42} = 22/42$

Or

- $P(X > 8) = 1 - P(X \leq 8) = 1 - [f(2) + f(4) + f(6) + f(8)] = 1 - \frac{20}{42} = 22/42$

f. $P(a \leq X \leq b) = P(X \leq b) - P(X < a)$

- $P(6 \leq X \leq 10) = P(X \leq 10) - P(X < 6) = P(X \leq 10) - P(X \leq 4)$

- $P(X \leq 10) = f(2) + f(4) + f(6) + f(8) + f(10) = 30/42$

- $P(X \leq 4) = f(2) + f(4) = 6/42$

$$= \frac{30}{42} - \frac{6}{42} = 24/42$$

or

- $P(6 \leq X \leq 10) = f(6) + f(8) + f(10) = 24/42$

6. Example 1 – Given the probability function shown in the table below, find the probability that X is

- 6 or less.
- less than 6.
- 5 or 7.
- not 6.

x	$f(x)$
4	0.2
5	0.2
6	0.3
7	0.3

7. Example 2 – A probability function is: $(x) = \frac{|x-3|}{15}$, $x = -2, -1, 0, 1, 2$. Find the probability that X is ...

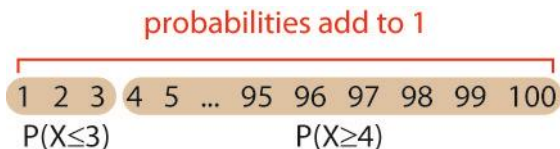
- 2
- 2
- 0
- negative
- between -1 and 1, inclusive

8. Example 2 – A probability function is: $(x) = \frac{x}{5050}$, $x = 1, 2, \dots, 100$. Find the probability that X is ...

- at most 3. Thus, $P(X \leq 3) = ?$. Note:

$$P(X \leq 3) = f(1) + f(2) + f(3) = \frac{1}{5050} + \frac{2}{5050} + \frac{3}{5050} = \frac{6}{5050} = 0.0012$$

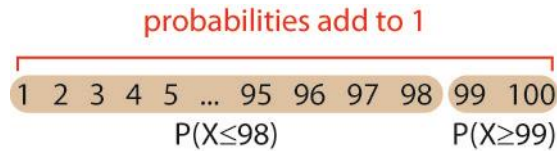
- at least 4. Thus, $P(X \geq 4) = ?$ Note:



Thus, $P(X \leq 3) + P(X \geq 4) = 1 \implies P(X \geq 4) = 1 - P(X \leq 3)$

$$\text{So, } P(X \geq 4) = 1 - \frac{6}{5050} = \frac{5044}{5050} = 0.9988$$

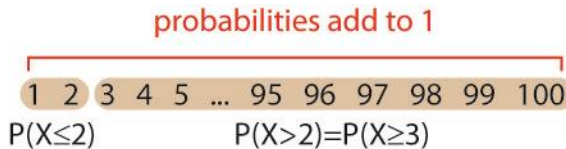
c. 98 or less. Note:



$$\text{So, } P(X \leq 98) = 1 - P(X \geq 99) = 1 - [f(99) + f(100)]$$

$$= 1 - \left(\frac{99}{5050} + \frac{100}{5050} \right) = 1 - \frac{199}{5050} = \frac{4851}{5050} = 0.9606$$

d. greater than 2. Note:



$$\text{So, } P(X > 2) = 1 - P(X \leq 2) = 1 - [f(1) + f(2)]$$

$$= 1 - \left(\frac{1}{5050} + \frac{2}{5050} \right) = 1 - \frac{3}{5050} = \frac{5047}{5050} = 0.9994$$

9. Example 4 – Suppose that X has probability function, $f(x) = cx$, $x = 2, 4, 6$, where c is a constant. Find the value of c.

$$\text{Remember that: } \sum f(x) = 1$$

$$\sum f(x) = 2c + 4c + 6c = 12c = 1 \Rightarrow c = 1/12$$

Thus, $f(x) = \frac{x}{12}$, $x = 2, 4, 6$ is a valid probability function.

10. Example 4 – Suppose that X has probability function, $f(x) = \frac{x-10}{c}$, $x = 20, 30, 40$ where c is a constant. Find the value of c.

$$\sum f(x) = \frac{20-10}{c} + \frac{30-10}{c} + \frac{40-10}{c} = \frac{60}{c} = 1 \Rightarrow c = 60$$

Thus, $f(x) = \frac{x-10}{60}$, $x = 20, 30, 40$, is a valid probability function.

Homework 4.1

Consider the following probability function for the next six questions.

$$f(x) = \frac{11-x}{55}, \quad x = 1, 2, \dots, 9, 10$$

Find the probability that X is ...

1. ...exactly 8. (a) 0.8 (b) 0.1 (c) 0.054 (d) 0.11 (e) 0.89

2. ...less than 4. (a) 0.49 (b) 0.75 (c) 0.6 (d) 0.7 (e) 0.3
3. ...3 or greater. (a) 0.8 (b) 0.65 (c) 0.3 (d) 0.33 (e) 0.75 (f) 0.67
4. ...not 5. (a) 0.65 (b) 0.5 (c) 0.33 (d) 0.67 (e) 0.89 (f) 0.94
5. ...more than 8.
6. ...at least 6 but no more than 8.

Consider the following probability function for the next four questions.

$$f(x) = \frac{|x|}{12}, \quad x = -4, -2, 2, 4$$

Find the probability that X is ...

7. ...-2.
8. ...negative.
9. ...positive
10. ...either 2 or -2
11. Consider this probability function, $f(x) = \frac{x^2}{c}$, $x = 1, 2, 3$, where c is a constant. Find the value of c that makes this a probability function.
12. Consider this probability function, $f(x) = cx^3$, $x = 1, 2, 3$, where c is a constant. Find the value of c that makes this a probability function.

4.3 – Expected Value

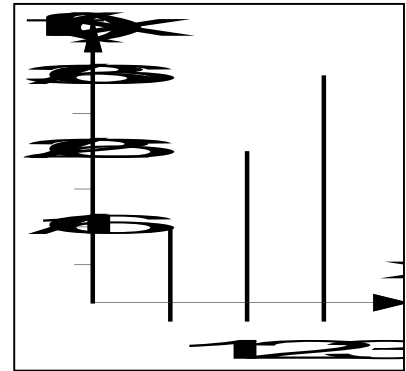
1. **Expected Value of X** (of a discrete RV) – The *expected value of a random variable* is the *long-run average value of a random variable*. It is what we *expect* to happen, *on average* in the long-run. It is a measure of the *center* of X . It is also called the **mean of the RV X**.
2. We use the symbol $E[X]$ or μ to denote the expected value of a random variable. It is computed by taking each value and multiplying by the respective probability, then adding these products:

$$\mu = E[X] = \sum_{\text{all } x} xf(x)$$

3. **Example** – Consider the probability function: $f(x) = \frac{x}{6}$, $x = 1, 2, 3$. Find the expected value of the random variable, X .

$$\mu = E[X] = \sum_{\text{all } x} xf(x) = 1\left(\frac{1}{6}\right) + 2\left(\frac{2}{6}\right) + 3\left(\frac{3}{6}\right) = \frac{7}{3} = 2.33$$

Does this answer make sense? The idea is that there is an experiment that results in an outcome of 1, 2, or 3. Do this experiment *many* times and then compute the *sample average* of all the outcomes. What do we expect the answer to be? Look at the graph below. Most of the time you do the experiment, you get a "3", sometimes you get a "2", and occasionally you get a "1". So, it should make sense that the *long-run average* value from doing this experiment is between 2 and 3, *i.e.* 2.33.



4. **Example** – Suppose you have the opportunity to play the following game as frequently as you like. You win \$10 with probability 0.4 and lose \$5 with probability 0.6. Ignore the cost of playing the game (assume that is factored into the *payoff*). Would you want to play this game once? A few times? Never? As frequently as possible?

Solution:

x	$f(x)$
\$10	0.4
-\$5	0.6

$$\mu = E[X] = \sum xf(x) = \$10(0.4) + (-\$5)(0.6) = \$4 + (-\$3) = \$1$$

The expected value of this game is \$1. What does this number mean? It means that if I am going to play this game *many* times, on average I will win \$1. So, I want to play the game as frequently as possible! If I play the game 1000 times, I would win approximately \$1000. If I play the game a *few* times, I'm not sure what will happen. Another way to think about this is as follows. If I agree to play the game *many* times, we don't have to play the game at all! Every time I want to play, I'll tell you, and you can give me \$1. The result will be the same in the *long-run*, the same as if we had played the game!

5. **Example** – Consider the following probability function, $f(x) = \frac{|x-3|}{15}$, $x = -2, -1, 0, 1, 2$. Find the mean of X .

$$\begin{aligned} \mu = E[X] &= \sum xf(x) = -2\left(\frac{5}{15}\right) + (-1)\left(\frac{4}{15}\right) + 0\left(\frac{3}{15}\right) + 1\left(\frac{2}{15}\right) + 2\left(\frac{1}{15}\right) = \\ &= -\frac{10}{15} = -\frac{2}{3} = -0.667 \end{aligned}$$

6. **Example** – What is the expected value of a single roll of a die.

$$\mu = E[X] = \sum xf(x) = 1\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right) = \frac{21}{6} = 3.5$$

4.4 – Variance

- Variance of X** – The *variance* of a random variable is a measure of how *spread out* the values of X are.
- We use the symbol $V[X]$ or σ^2 to denote the variance of a random variable. The definition is:

$$\sigma^2 = V[X] = \sum (x - \mu)^2 f(x)$$

However, a simpler formula for computing is:

$$\sigma^2 = V[X] = E[X^2] - \mu^2$$

Where:

$$E[X^2] = \sum x^2 f(x)$$

3. **Standard Deviation of X** – The standard deviation of a random variable is the square-root of the variance: $\sigma = \sqrt{\sigma^2}$. Usually, we are interested in the standard deviation as opposed to the variance because it has units which are the same as the units of the variable.

4. To calculate the standard deviation we use these steps:

1. Find the expected value of X : $\mu = E[X] = \sum xf(x)$
2. Find the expected value of X^2 : $E[X^2] = \sum x^2 f(x)$
3. Find the variance of X : $\sigma^2 = V[X] = E[X^2] - \mu^2$
4. Find the standard deviation of X : $\sigma = \sqrt{\sigma^2}$

5. **Example:** Consider the probability function: $f(x) = \frac{x}{6}$, $x = 1, 2, 3$. Find the variance and standard deviation of the random variable, X .

$$a. \mu = E[X] = \sum xf(x) = 1\left(\frac{1}{6}\right) + 2\left(\frac{2}{6}\right) + 3\left(\frac{3}{6}\right) = \frac{7}{3} = 2.33$$

$$b. E[X^2] = \sum x^2 f(x) = 1^2\left(\frac{1}{6}\right) + 2^2\left(\frac{2}{6}\right) + 3^2\left(\frac{3}{6}\right) = 6$$

$$c. \sigma^2 = V[X] = E[X^2] - \mu^2 = 6 - \left(\frac{7}{3}\right)^2 = 5/9 = 0.556$$

$$d. \sigma = \sqrt{\sigma^2} = \sqrt{\frac{5}{9}} = 0.745$$

6. **Example:** Same as Example 4 in the previous section (You win \$10 with probability 0.4 and lose \$5 with probability 0.6). Find the variance and standard deviation of the payoff from this game.

$$a. \mu = E[X] = \sum xf(x) = \$10(0.4) + (-\$5)(0.6) = \$4 + (-\$3) = \$1$$

$$b. E[X^2] = \sum x^2 f(x) = (\$10)^2(0.4) + (-\$5)^2(0.6) = 55$$

$$c. \sigma^2 = V[X] = E[X^2] - \mu^2 = 55 - (1)^2 = 54$$

$$d. \sigma = \sqrt{\sigma^2} = \sqrt{54} = \$7.348$$

7. **Example:** Consider the following probability function, $f(x) = \frac{|x-3|}{15}$, $x = -2, -1, 0, 1, 2$. Find the variance and standard deviation of X .

a. $\mu = -\frac{2}{3} = -0.667$ (from a previous example)

b. $E[X^2] = \sum x^2 f(x) = (-2)^2 \left(\frac{5}{15}\right) + (-1)^2 \left(\frac{4}{15}\right) + 0^2 \left(\frac{3}{15}\right) + 1^2 \left(\frac{2}{15}\right) + 2^2 \left(\frac{1}{15}\right) = \frac{30}{15} = 2$

c. $\sigma^2 = V[X] = E[X^2] - \mu^2 = 2 - \left(-\frac{2}{3}\right)^2 = \frac{14}{9} = 1.555$

d. $\sigma = \sqrt{V[X]} = \sqrt{\frac{14}{9}} = \frac{\sqrt{14}}{3} = 1.247$

8. Comparing Investments – Investment opportunities are frequently compared by answering these questions:

1. What is *expected return*? We answer this by interpreting the expected value, μ
2. How much *volatility* is there? We answer this by interpreting the standard deviation, σ
3. What is the probability that we will return a profit?

9. Example: Investment A pays \$100 with probability 1/10, \$500 with probability 3/5, and -\$100 with probability 3/10. Investment B pays \$300 with probability 2/5, \$400 with probability 2/5, and -\$50 with probability 1/5. Which investment do you prefer and why?

Investment A		Investment B	
x	$f(x)$	x	$f(x)$
\$100	1/10	\$300	2/5
\$500	3/5	\$400	2/5
-\$100	3/10	-\$50	1/5

Criteria	
Investment A	Investment B
$\mu_A = \$280$	$\mu_B = \$270$
$\sigma_A = \$275$	$\sigma_B = \$166$
$P(A \text{ returns a profit}) = 0.7$	$P(B \text{ returns a profit}) = 0.8$

Summary: The answer to the problem is, *it depends*:

1. If you are going to invest many times, you probably only need to look at the expected return, in which case you choose Investment A.
2. If you are going to invest a few times, you juggle the expected return and the volatility. How you do this depends on your tolerance for risk. In other words, if your tolerance for risk is low, you may choose an investment with a lower expected return because it has less volatility. For this problem, Investment B has less risk, $\sigma_B = \$166 < \$275 = \sigma_A$, but a smaller expected return, $\mu_B = \$270 < \$280 = \mu_A$.
3. If you are going to invest one time, you might also want to take into account the probability that the investment will return a profit. For this problem, Investment B has a 10% higher chance of returning a profit.

Homework 4.2

Consider the following probability function for the next two questions:

x	3	4	5
$f(x)$	0.7	0.2	0.1

1. The expected value of the random variable X is: (a) 4 (b) 3.4 (c) 1.6 (d) 12 (e) 0.33
2. The **standard deviation** of the random variable X is: (a) 16.67 (b) 12 (c) 19.22 (d) 11.56 (e) 0.66 (f) 3.47

Consider the following probability function, $f(x) = \frac{x}{12}$, $x = 2, 4, 6$.

3. Find the mean of the random variable X .
4. Find the variance of the random variable X .
5. Find the standard deviation of the random variable X .

Consider an investment that returns \$5 with probability 0.4, \$2 with probability 0.2, and -\$9 with probability 0.4.

6. What is the expected return of this investment?
7. Is this a good investment for the long-run? Why or why not?
8. What is the volatility of this investment?

4.5 – Binomial Distribution

1. Binomial Distribution –

- The binomial distribution is a *special* probability function. This is because it explains a *pattern* that occurs in some real problems.
- We use the notation, $X \sim \text{Bin}(n, p)$ to say that the random variable X follows (or *has*) the binomial distribution.
- n and p are called the *parameters* of the distribution. This means that we must *know* the values of n and p before we can do any calculations.
- The probability function for a binomial random variable is:

$$f(x) = \binom{n}{x} p^x (1-p)^{n-x}, \quad x = 0, 1, \dots, n$$
$$= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}$$

We will not need this formula because our calculator will calculate it for us.

- **Mean and Variance of Binomial RV** – Since the binomial distribution follows a pattern, it can be shown that the expected value of a Binomial random variable is: $\mu = np$ and the variance is: $\sigma^2 = np(1-p)$.

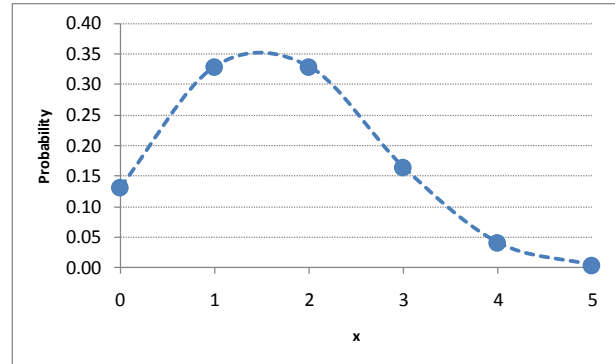
2. **Example** – Suppose a baseball player goes to bat 5 times in a game and that his batting average is 0.333. Find the probability distribution for the number of hits in five at-bats assuming that each time at the plate is independent of the others.

Let X = number of hits in 5 at-bats, $X \in \{0,1,2,3,4,5\}$

This problem can be answered with the binomial distribution (we will show why later) and we can write:

$$X \sim \text{Bin}(n = 5, p = 0.333)$$

This is a graph of the probabilities is shown on the right.



3. **Example** – A drug works 90% of the time in curing an infection. A sample of 100 people with the infection is chosen and each is given the drug. Find the probability distribution for the number of people who are cured.

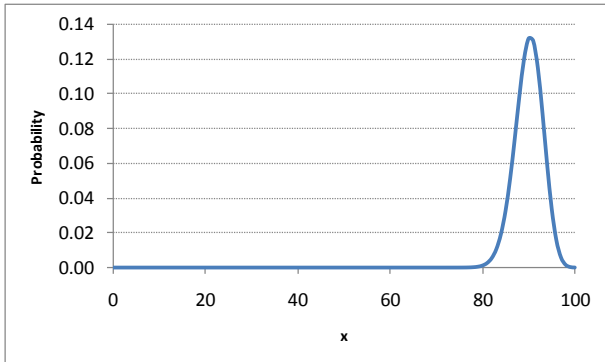
Let X = number of people cured out of 100, $X \in \{0,1, \dots, 100\}$

This problem can also be answered with the binomial distribution and we can write:

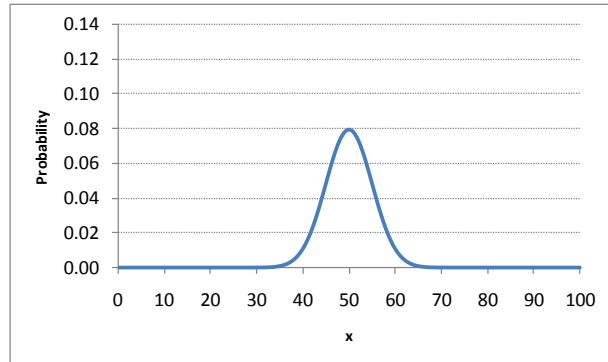
$$X \sim \text{Bin}(n = 100, p = 0.9)$$

These are some graphs of the probabilities:

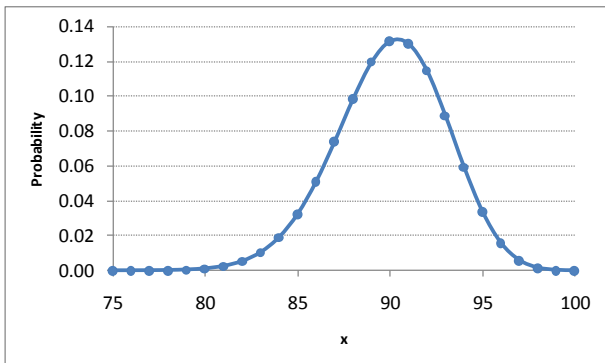
90% Success Rate



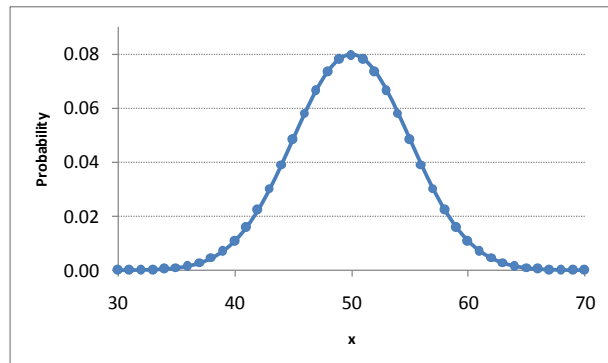
50% Success Rate



90% Success Rate



50% Success Rate



What is the mean number of people who are cured? Another way to ask this is, *What is the expected number of people who will be cured?*

$$\mu = np = 100 * 0.9 = 90 \text{ people} \quad \left| \quad \mu = np = 100 * 0.5 = 50$$

What is the variance and standard deviation of the number of people who are cured?

$$\sigma^2 = np(1 - p) = 100 * 0.9 * (1 - 0.9) = 9 \quad \left| \quad \sigma^2 = np(1 - p) = 100 * 0.5 * (1 - 0.5) = 25$$

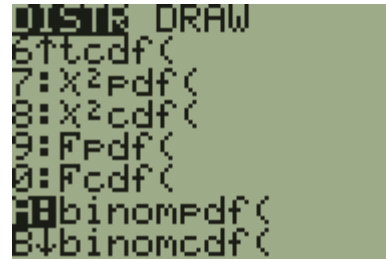
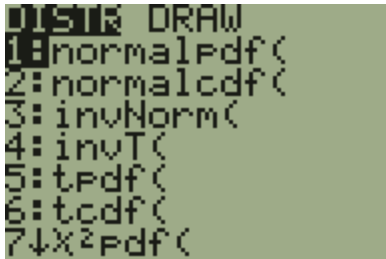
$$\sigma = \sqrt{\sigma^2} = \sqrt{9} = 3 \text{ people.} \quad \left| \quad \sigma = \sqrt{\sigma^2} = \sqrt{25} = 5 \text{ people.}$$

4. TI 83/84 Binomial Distribution Directions

For the following directions, we assume we know that $X \sim \text{Bin}(n, p)$ and that we know the values of n and p .

a. We can answer this question: $P(X = x)$ by following these directions:

Choose *DISTR* by pressing 2nd and the VARS key. Scroll down to *A:binompdf()* and press *Enter*.



Type in: n, p, x and press *Enter*.

Thus, $P(X = x) = \text{binompdf}(n, p, x)$

b. We can answer this question: $P(X \leq x)$ by using : “*B:binomcdf*” (in step 2 above). Thus,

$P(X \leq x) = \text{binomcdf}(n, p, x)$

Any other probability questions will require using simple probability rules to rearrange the question to conform to b immediately above.

- c. $P(X < x) = \text{binomcdf}(n, p, x - 1)$
- d. $P(X \geq x) = 1 - \text{binomcdf}(n, p, x - 1)$
- e. $P(X > x) = 1 - \text{binomcdf}(n, p, x)$
- f. $P(a \leq X \leq b) = \text{binomcdf}(n, p, b) - \text{binomcdf}(n, p, a - 1)$

5. **Example** – Suppose a baseball player goes to bat 5 times in a game and that his batting average is 0.333. What is the probability of getting 3 hits? 3 or fewer hits? 3 or more hits? more than 3 hits? between 2 and 4 hits, inclusive?

As we saw previously:

Let X = number of hits in 5 at-bats, $X \in \{0,1,2,3,4,5\}$

$$X \sim \text{Bin}(n = 5, p = 0.333)$$

Thus,

$$P(X = 3) = \text{binompdf}(5, 0.333, 3) = 0.1643 = 0.16$$

$$P(X \leq 3) = \text{binomcdf}(5, 0.333, 3) = 0.9549 = 0.95$$

$$P(X \geq 3) = 1 - \text{binomcdf}(5, 0.333, 2) = 0.2094 = 0.21$$

$$P(X > 3) = 1 - \text{binomcdf}(5, 0.333, 3) = 0.0451 = 0.04$$

$$P(2 \leq X \leq 4) = \text{binomcdf}(5, 0.333, 4) - \text{binomcdf}(5, 0.333, 1) = 0.5343 = 0.53$$

6. When can we use the Binomial distribution? We can use the binomial distribution when three conditions have been met:
1. Experiment is a sequence of n independent trials.
 2. Each trial results in one of two possible outcomes, *Success* or *Failure*. Success is defined as *what you are interested in*.
 3. The probability of *Success* on each trial is p and is constant from trial to trial.

If the three conditions are met, then the random variable,

X = the number of Successes in n trials.

Follows the binomial distribution with parameters n and p :

$$X \sim \text{Bin}(n, p)$$

7. **Example** – Roll six dice. What is the probability of obtaining exactly two "4's"? Have the conditions for the Binomial distribution been met?
1. **Experiment is a sequence of n independent trials.** We can think of rolling the dice one-at-a-time, thus we have six *trials* (where a *trial* is a roll of a single die) and we say that $n = 6$. Are the trials *independent*? Yes, the outcome of the first roll in no way influences the outcome of the second roll, *etc.* So, the first condition has been met.
 2. **Each trial results in one of two possible outcomes, *Success* or *Failure*.** *Success* for this problem is when we roll a die and obtain a "4" and a *Failure* is anything else. So, we can classify the outcome of each of the six *trials* (roll of the die) as a *Success* or a *Failure*. For instance if I roll: 3, 6, 4, 2, 5, 3, I can classify these as: *Failure, Failure, Success, Failure, Failure, Failure*. Thus, the second condition has been met.

3. The probability of Success on each trial is p and is constant from trial to trial. The probability of *Success* for this problem is the probability of rolling a single die and obtaining a "4." This probability is known to be $1/6$, thus $p = P(\text{Success}) = P(\text{roll a 4}) = 1/6$. Is p constant from trial to trial (roll to roll). Yes, the probability of rolling a "4" never changes. Thus, the third condition has been met.

Thus, we see that the three conditions have been met. So,

Let X = the *number* of "4's" obtained in six rolls of the die.

And we say that X has a Binomial distribution with parameters $n=6$ and $p=1/6$.

$$X \sim \text{Bin}(n = 6, p = 1/6)$$

We can also see that the *values* that X can take on are:

$$\{0,1,2,3,4,5,6\}$$

and the question being asked is

$$P(X = 2) = ? \quad (\text{Answer: } 0.2009).$$

8. Example – A drug is successful 90% of the time in curing a particular disease. The drug is given to a group of 10 people with the disease. Find the probability that eight people are cured.

1. Experiment is a sequence of n independent trials. We can think of giving the drug to the 10 people one-at-a-time, thus we have ten *trials* (where a *trial* is when we give the drug to a single person) and we say that $n = 10$. Are the trials *independent*? Yes, whether the drug works on the first person in no way influences whether the drug works on the second person, *etc.* So, the first condition has been met.

2. Each trial results in one of two possible outcomes, Success or Failure. *Success* for this problem is when someone is *cured* and a *Failure* is when the person is not cured. So, we can classify the outcome of each of the ten *trials* as a *Success* or a *Failure*. Thus, the second condition has been met.

3. The probability of Success on each trial is p and is constant from trial to trial. The probability of *Success* for this problem is the probability that a person is cured after they take the drug. We assume that this probability is 0.9 since we are told that 90% of people are cured with the drug., thus $p = P(\text{Success}) = P(\text{a person is cured}) = 0.9$. Is p constant from trial to trial (person to person). Technically, no. Everybody is different. However, in a situation like this we frequently *assume* that it is constant for all people.

Thus, we see that the three conditions have been met. So,

Let X = the *number* of people who are cured out of ten.

And we say that X has a Binomial distribution with parameters $n=10$ and $p=0.9$.

$$X \sim \text{Bin}(n = 10, p = 0.9)$$

We can also see that the *values* that X can take on are:

$$\{0,1,2,3,4,5,6,7,8,9,10\}$$

and the question being asked is

$$P(X = 8) = ? \quad (\text{Answer: } 0.1937).$$

9. **Example** – Choose 5 cards from a deck. Find the probability of choosing exactly two aces.
1. **Experiment is a sequence of n independent trials.** We can think of drawing the cards one-at-a-time, thus we have five *trials* (where a *trial* is a single card being drawn) and we say that $n = 5$. Are the trials *independent*? No, see condition 3 below.
 2. **Each trial results in one of two possible outcomes, *Success* or *Failure*.** *Success* for this problem is when an Ace is drawn and a *Failure* is when any other card is drawn. So, we can classify the outcome of each of the five *trials* as a *Success* or a *Failure*. Thus, the second condition has been met.
 3. **The probability of *Success* on each trial is p and is constant from trial to trial.** The probability of *Success* for this problem is the probability that an Ace is drawn. Note this probability is *not* constant from draw-to-draw. Why? What is the probability of an Ace on the first draw? The second draw?

This problem *does not* fit the binomial pattern.

10. **Example** – A drug is successful 90% of the time in curing a particular disease. In a study of the drug's effectiveness, a group of 10 people with the disease are given the drug. Assume that the people are independent.

As we saw before, $X =$ the number of people who are cured and $X \sim \text{Bin}(n = 10, p = 0.9)$.

What is the probability that...

- a. ...8 people are cured? $P(X = 8) =$
- b. ...8 or fewer people are cured? $P(X \leq 8) =$
- c. ...less than 8 people are cured? $P(X < 8) = P(X \leq 7) =$
- d. ...at least 7 are cured? $P(X \geq 7) = 1 - P(X \leq 6) = 1 -$
- e. ...more than 7 are cured? $P(X > 7) = P(X \geq 8) = 1 - P(X \leq 7) = 1 -$
- f. What is the mean number of people who are cured?
- g. What is the variance and standard deviation of the number of people who get well?
- h. ...3 people are not cured?

We can answer this question rather simply by noting that if 3 people are *not* cured then 7 people are cured: $P(X = 7) =$

However, a better approach, especially for the following problems that are a little more tricky is to

Let $Y =$ the number of people who are *not* cured and $Y \sim \text{Bin}(n = 10, p = 0.1)$. Then, we can answer the question with: $P(Y = 3) =$

- i. ...that more than 2 people are not cured? $P(Y > 2) = P(Y \geq 3) = 1 - P(Y \leq 2) = 1 -$
- j. What is the mean number of people who are not cured?
- k. What is the variance of the number of people who are not cured?

11. Example – Suppose that a company places orders with a manufacturer frequently. Consider the table below that shows the number of days it takes to receive an order and the respective probability. An order is considered *Late* if it is received on the third or fourth day. Assume that orders are independent of one another.

Day	1	2	3	4
Probability	0.20	0.65	0.10	0.05

- a. Out of the next ten orders, what is the probability of three or fewer late orders?

Let X = the number of late orders out of the next 10, and $X \sim \text{Bin}(n = 10, p = 0.1 + 0.05 = 0.15)$.

$$P(X \leq 3) =$$

- b. Out of the next 200 orders, how many do we expect to be late?

Let X = the number of late orders out of the next 200, and $X \sim \text{Bin}(n = 200, p = 0.15)$.

Homework 4.3

Consider the following information for the next three questions. A company has purchased many computers from the same supplier over several years and has found that 95% of the computers operate properly upon initial configuration and 5% do not operate properly. Assume the computers are independent of one another.

- Consider the next shipment of 6 computers to the company. Suppose you want to find the probability that there are exactly two computers that do not operate properly. (a) Identify the values of n and p and write the probability question. (b) Find the answer using your calculator.
- Consider the next shipment of 20 computers to the company. Suppose you want to find the probability that there are at least 17 computers that operate properly. (a) Identify the values of n and p and write the probability question. (b) Find the answer using your calculator.
- Suppose that over the course of the next year the company plans to purchase 250 computers. How many computers should the company expect to have problems with?
(a) 5 (b) 10 (c) 11.875 (d) 12.5 (e) 13.825 (f) 21
- A simple random sample survey in New York City revealed that 15% of city residents use a car as their primary means of transportation, 5% use a motorcycle, 60% use public transportation, 15% use taxis, and the rest walk. Suppose you take a sample of 60 NYC residents and want to find the probability that more than twenty people say they use either a car or a taxi. (a) Identify the values of n and p and write the probability question. (b) Find the answer using your calculator.
- A baseball player has a batting average of 0.375 and each time he goes to bat is independent of other at bats. (a) Suppose you want to calculate the probability that the player gets 4 hits in his next 4 at-bats. (a) Identify the values of n and p and write the probability question. (b) Find the answer using your calculator. (c) What about no hits in his next 4 at-bats? (d) Suppose this player goes to bat exactly 4 times in each of 162 games. In how many games do we expect the player to go hitless?

Chapter 5 – The Normal Distribution

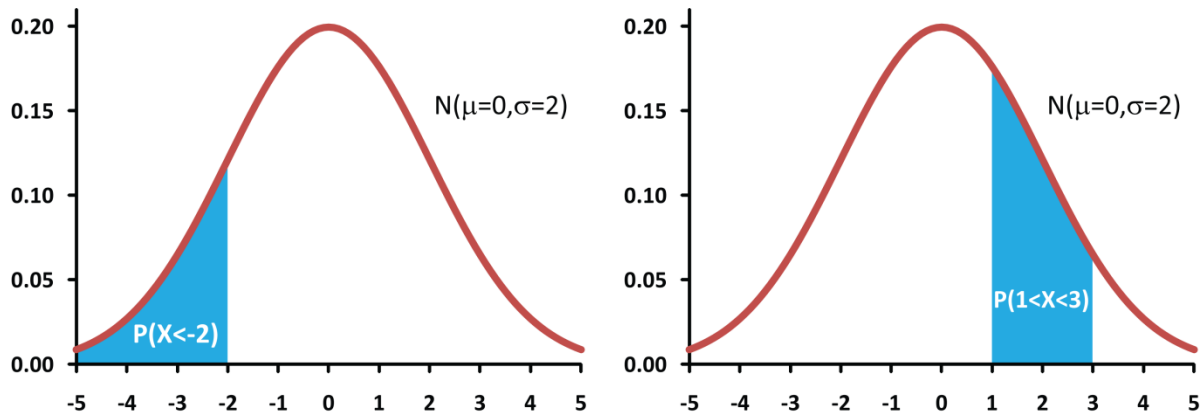
In this chapter we consider the normal probability distribution

5.1 – Introduction and Graphing

1. The normal distribution is a *continuous* distribution and is considered a special distribution because of its widespread use. Every area of science and business makes use of the normal distribution. The normal distribution is a *model* for the *bell-shaped* or *mound-shaped* histogram that we discussed earlier. When we see a bell-shaped histogram, many times that will lead us to assume (hypothesize) that the underlying nature of the distribution is governed by the normal distribution. Also, the normal distribution is the basis for many statistical procedures.
2. We use the notation, $X \sim N(\mu, \sigma)$ to say that X follows a normal distribution with parameters μ and σ . The *mean* (or *expected value*) of the normal distribution is μ and the *standard deviation* is σ . The normal distribution has probability density function:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2}, \quad -\infty < x < \infty$$

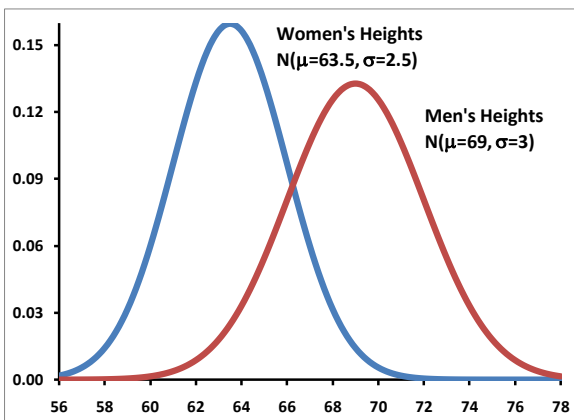
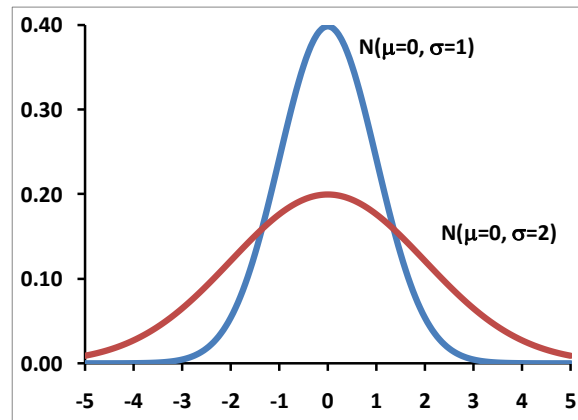
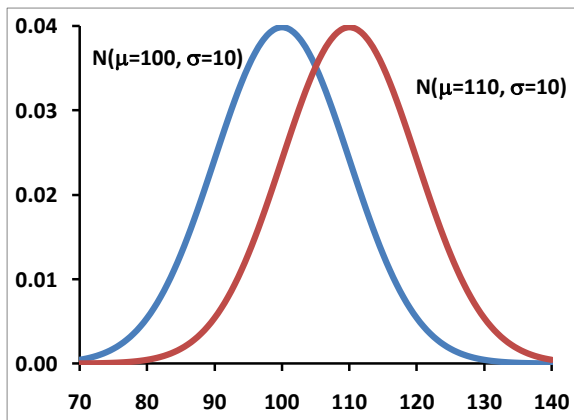
3. With discrete random variables, we use a *probability function* which provides the probabilities for each value the random variable can take on. Remember that continuous random variables can take on *any* value in a given range. Thus, it no longer makes sense to talk about the probability of an individual outcome because the outcome would be just one of an infinite number. With continuous random variables, we use a *probability density function*, $f(x)$, which describes the density or concentration of probabilities. With continuous random variables, probability is defined as *area under $f(x)$* . Thus, the total area under $f(x)$ is 1.



4. Notice that the Normal distribution is defined for $-\infty < x < \infty$. Thus, it goes on forever in either direction, getting closer and closer to the x -axis. Many times this domain doesn't make sense when we are using the normal distribution for a real problem. For example, the normal distribution is used to model men's and women's heights (and weights), but obviously heights cannot be negative, nor can they be arbitrarily large. However, as we saw with the Empirical Rule, 99.7% of the area (probability) is within $\mu \pm 3\sigma$. In such situations, we effectively ignore the *tails* of the distribution. This seems intuitively valid because the tails hold so little probability.

5. When we look at a normal distribution, we see:
 1. Graph is bell-shaped and centered at μ .
 2. The *smaller (larger)* the standard deviation the *taller (shorter)* the graph.
 3. Approximately 95% of area of the graph should fall within $\mu \pm 2\sigma$.

6. Example – Remembering that the total area under the graph is 1, it is easy to see in the left figure below that the two curves are exactly the same, just that the red one has been shifted 10 units to the right. Similarly, the graphs on the right are both centered at 0, but their spreads are different. If we pick up the two tails of the blue graph and spread them out (larger σ), then the peak must come down (the blue graph) to ensure that the total area remains 1.

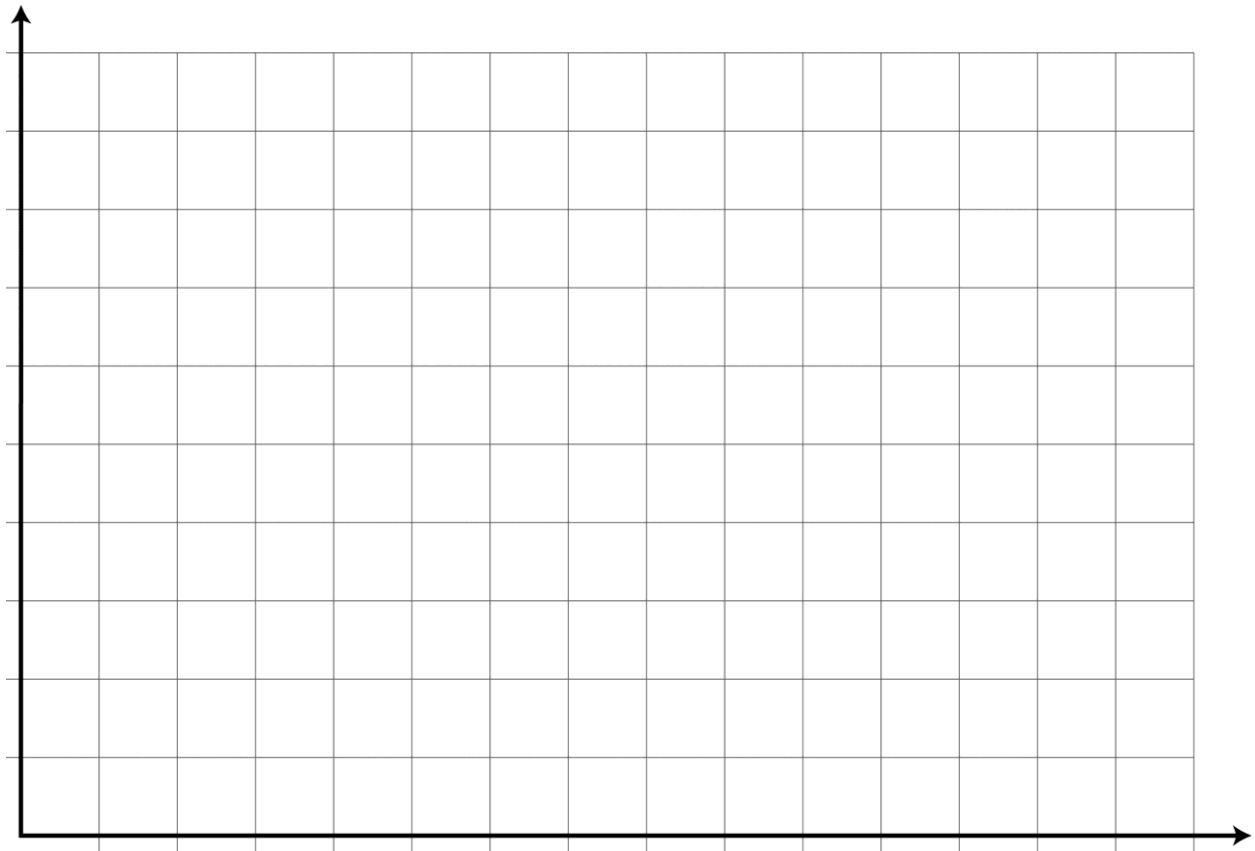


7. Several notes about continuous probabilities:

1. $P(X = a \text{ number}) = 0$. Probability is defined as *area*. There is no area under a number. Thus, we can only talk about probabilities for *ranges* of numbers. For instance, $P(X < 200) = 0$.
2. $P(X \leq c) = P(X < c)$. This is so by observing the following:

$$P(X \leq c) = P(X < c) + P(X = c) = P(X < c) + 0 = P(X < c)$$

8. **Example** – Graph $X \sim N(12,4)$ and $Y \sim N(10,2)$ on the same set of axes. The *x-axis* should be correct, but the vertical axis can be relative.



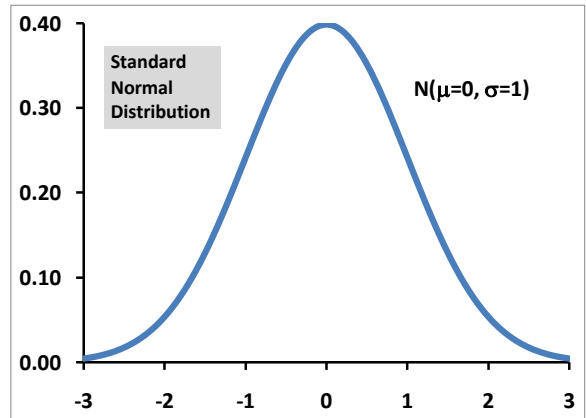
Homework 5.1

1. The total sales per day at Restaurant A are normally distributed with a mean of \$800 and standard deviation of \$200. The total sales per day at restaurant B are normally distributed with a mean of \$700 and standard deviation of \$100.
 - a. Graph these two normal distributions on the same axis. Be sure to label the *x-axis* with appropriate values.
 - b. Which restaurant has the more likely to make less than \$500 in a day? Demonstrate that your answer is correct on the graphs.

2. Test scores for Men are normally distributed with a mean of 70 and standard deviation of 15. Test scores for Women are normally distributed with a mean of 80 and standard deviation of 5.
 - a. Graph these two normal distributions on the same axis. Be sure to label the x-axis with appropriate values.
 - b. Are Men or Women more likely to score above a 90 on the test? Demonstrate that your answer is correct on the graphs.

5.2 – Standard Normal Distribution

1. There are an infinite number of normal distributions, $N(\mu, \sigma)$, as μ can be any real value and σ can be any positive real value. One normal distribution, $N(0,1)$ is particularly important in statistics. A $N(0,1)$ random variable is given the special symbol, Z , so that we say, $Z \sim N(0,1)$ and is referred to as the *standard normal distribution*.
2. An interesting and useful fact is that any normal distribution can be converted into the standard normal distribution. If $X \sim N(\mu, \sigma)$, then:

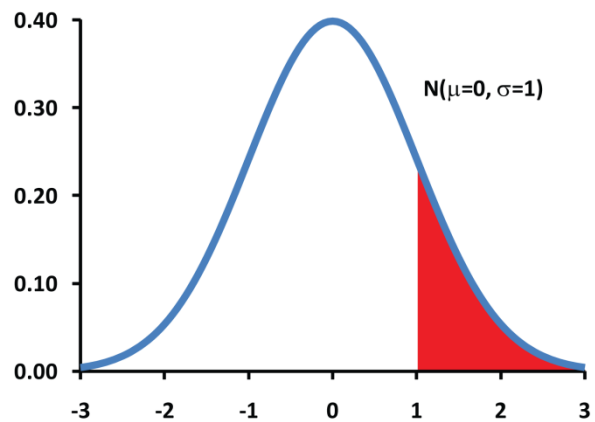
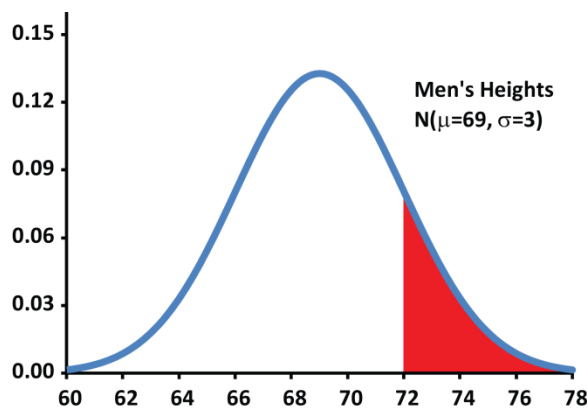


$$Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

This is the basis for most of the statistical procedures we will consider. Note the formula above is similar to the one we used for *z-scores* earlier.

3. Example –As an example of the formula above, consider this question: Suppose that men’s heights are normally distributed, $X \sim N(69,3)$ and we want to know $P(X > 72)$. This question is the same as:

$$P(X > 72) = P\left(Z > \frac{72 - 69}{3}\right) = P(Z > 1)$$



5.3 – Normal Probability Calculations

1. Assume we know that $X \sim N(\mu, \sigma)$. The TI-83 will calculate probabilities for a normal distribution by pressing 2^{nd} +DISTR and scrolling down to *normalcdf*. The calculator will answer two types of Normal probability questions:

a. $P(a \leq X \leq b) = \text{normalcdf}(a, b, \mu, \sigma)$

b. $P(X \leq c) = P(X < c) = \text{normalcdf}(-1E99, c, \mu, \sigma)$

Note that “-1E99” is the way you denote a very small number on the calculator. You can think of it as “negative infinity”. To enter it on your calculator: type in “-1” followed by: 2^{nd} key, then the *EE* key (above the comma).

If we have a “greater than” question, then we must use the complement rule:

c. $P(X \geq c) = P(X > c) = 1 - \text{normalcdf}(-1E99, c, \mu, \sigma) = \text{normalcdf}(c, 1E99, \mu, \sigma)$

2. Example – Suppose that $X \sim N(10, 2)$. Find $P(X < 12)$.

$$P(X < 12) = \text{normalcdf}(-1E99, 12, 10, 2) = 0.8413$$

Note, this can be converted to the standard normal distribution:

$$P(X < 12) = P\left(Z < \frac{12-10}{2}\right) = P(Z < 1) = \text{normalcdf}(-1E99, 1, 0, 1) = 0.8413$$

3. Example – A company makes wine glasses. The number of wine glasses made per day is normally distributed with mean 300 and standard deviation 20. Find the probability that the company makes, (a) fewer than 335 wine glasses in a day, (b) more than 285 wine glasses in a day, (d) between 283 and 291 wine glasses in a day

a. $P(X < 335) = \text{normalcdf}(-1E99, 335, 300, 20) = 0.9599 = 0.96$

Note, this can be converted to the standard normal distribution:

$$P(X < 335) = P\left(Z < \frac{335-300}{20}\right) = P(Z < 1.75) = \text{normalcdf}(-1E99, 1.75, 0, 1) = 0.9599$$

b. $P(X > 285) = 1 - P(X \leq 285) = 1 - \text{normalcdf}(-1E99, 285, 300, 20) =$

$$1 - 0.2266 = 0.7734 = 0.77$$

c. $P(283 < X < 291) = \text{normalcdf}(283, 291, 300, 20) = 0.1287 = 0.13$

4. Example – The personnel manager of a large company requires job applicants to take a certain test and achieve a score of at least 500. If the test scores are normally distributed with mean 485 and standard deviation 30, what percentage of the applicants pass the test?

$$P(X \geq 500) = 1 - P(X < 500) = 1 - \text{normalcdf}(-1E99, 500, 485, 30) = 1 - 0.6915 = 0.3085 = 0.31$$

5. Example – Experience indicates that the development time for a photographic printing paper is distributed as $X \sim N(30 \text{ seconds}, 1.1 \text{ seconds})$. Find the probability that it will take between 28.5 and 31.2 seconds for a randomly selected piece of photographic printing paper to develop.

$$P(28.5 < X < 31.2) = \text{normalcdf}(28.5, 31.2, 30, 1.1) = 0.7760 = 0.78$$

6. Example – Ball bearings are manufactured with a diameter that is normally distributed with a mean of 3 mm and a standard deviation of 0.1 mm. A customer has specifications that require that the ball bearings have diameter between 2.85 and 3.1 mm. (a) What fraction of ball bearings manufactured meet specifications? (b) What fraction of ball bearings manufactured do not meet specifications?

a. $P(\text{meet spec}) = P(2.85 < X < 3.1) = \text{normalcdf}(2.85, 3.1, 3, 0.1) = 0.7745 = 0.77$

b. $P(\text{don't meet spec}) = 1 - P(\text{meet spec}) = 1 - 0.7745 = 0.2255 = 0.22$

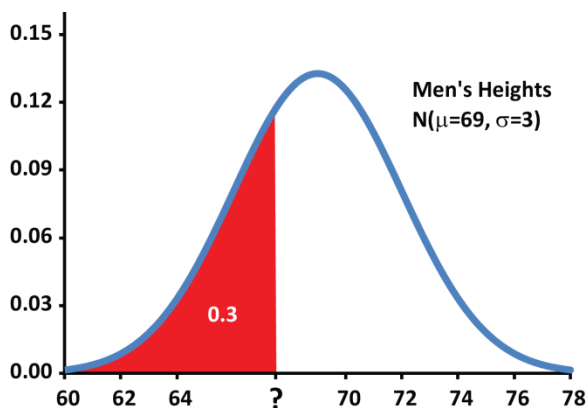
Homework 5.2

1. The total sales per day at Restaurant A are normally distributed with a mean of \$800 and standard deviation of \$200. The total sales per day at restaurant B are normally distributed with a mean of \$700 and standard deviation of \$100.
- Which restaurant has the more likely to make less than \$500 in a day? Demonstrate that your answer is correct by computing the appropriate probabilities.
 - What percentage of time will Restaurant A have sales between \$600 and \$1000?
 - What percentage of time will Restaurant B have sales more than \$865?
2. The speed of cars on a local interstate is normally distributed with mean of 76 mph and standard deviation of 5 mph.
- The DOT has determined that safe drive speeds are between 65 and 73 mph. What is the probability that a randomly selected driver is driving a safe speed?
 - What percentage of drivers are not driving a safe speed.
 - What is the probability that a randomly selected driver is driving 90 mph?
 - What is the probability that a randomly selected driver is driving the speed limit (70mph) or less?

3. The number of customers per day at a large store is normally distributed with mean of 1500 and standard deviation of 200.
 - a. Find the probability that there are fewer than 1200 people on a random day.
 - b. Find the probability that there are at least 1200 but less than 1700?
 - c. Find the probability that there are more than 1700 people on a random day.
 - d. Find the probability that there are fewer than 1200 people or more than 1700.

5.4 – Inverse Normal Probability Calculations

1. In the previous section, we were given a normal distribution, and asked what the probability is for a particular value, e.g. $P(X < 10) = ?$. In this section, we consider a slightly different problem. We are given the probability, but want to know what the value is, e.g. $P(X < ?) = 0.2$. Sometimes we call these *cutoff* problems.
2. For instance, if men's heights are normally distributed with mean 69" and standard deviation 3", then 30% of men have height below what value? Thus, $P(X < \text{cutoff}) = 0.3$ and our goal is to find the *cutoff* value.



3. To solve this problem, we use the *invNorm* function on our calculator by choosing: 2nd + DISTR and scroll down to 3:*invNorm*. Suppose you have $X \sim N(\mu, \sigma)$, and are given a probability, p . Two types of cutoff problems are:
 - a. $P(X < c) = p \Rightarrow c = \text{invNorm}(p, \mu, \sigma)$
 - b. $P(X > c) = p \Rightarrow c = \text{invNorm}(1 - p, \mu, \sigma)$
4. **Example** – A reading test has scores that are normally distributed with mean 70 and standard deviation 5.4. Students exiting 3rd grade take this test. If a student scores below the 20th percentile, then they must take summer reading classes. What score should be used as the cutoff?

$$P(X < \text{cutoff}) = 0.2 \Rightarrow \text{cutoff} = \text{invNorm}(0.2, 70, 5.4) = 65.455$$

5. **Example** – The scores on an aptitude test are normally distributed with mean 485 and standard deviation 30. Job applicants take this test and if they score in the top 15% then they are called back for an interview. What value should be used for the cutoff on the test?

$$P(X > \text{cutoff}) = 0.15 \Rightarrow \text{cutoff} = \text{invNorm}(0.85, 485, 30) = 516.09$$

Homework 5.3

- Test scores for Men are normally distributed with a mean of 70 and standard deviation of 15. Test scores for Women are normally distributed with a mean of 80 and standard deviation of 5.
 - 15% of men will score below what value?
 - 20% of women will score above what value?
- The number of customers per day at a large store is normally distributed with mean of 1500 and standard deviation of 200.
 - If the store has a busy day, then the manager gets a bonus. The owner would like to give these bonuses approximately 7% of the time. What value should she use as a cut-off so that whenever the number of customer exceeds this value, the manager will get a bonus?
 - Find the first quartile. (Hint: The first quartile is the 25th percentile.)

5.5 – Sampling Concepts

- In the previous sections, we have asked probability questions about *individual* items.

What is the probability that...

- ...a man is at least 72 inches tall? $P(X > 72)$
- ...a car is travelling less than 40 *mph*? $P(X < 40)$

In this section we consider probability questions about *averages*.

What is the probability that...

- ...the average of 30 men's heights is at least 72"? $P(\bar{X} > 72)$
- ...the average of 30 cars is less than 40 *mph*? $P(\bar{X} < 40)$

- A review of some terms:
 - Population** – The set of *ALL* elements that have a characteristic of interest.
 - Sample** – A subset of a population, usually chosen so that it is representative of population.
 - Simple Random Sample (SRS)** – Chosen by making sure that (1) each item selected comes from the same population, (2) each item is selected independently of the others, and (3) each item has an equal chance to be in the sample.
 - Reasons for sampling** – Want information about a population. Thus, we can take a:
 - Census - exact(?), timely, costly
 - Sample - inexact, quicker, less costly, can be very accurate.

Statistics (as we will see in the next chapter) allows us to quantify (measure) the inexactness of results from a sample when drawing conclusions about a population. Now, we need to understand a bit of the probability

theory upon which this idea is built.

3. **Sampling distribution of a sample average** – In order to answer questions about probabilities concerning averages, we must consider what the probability distribution of an average is. In other words, every time we take a sample from a population, we get a different value. It is very important to realize that *sample averages vary!* In statistics we use the sample average frequently. To make intelligent use of a sample average we must know *how much it varies*.
4. **Theorem** – Consider a simple random sample from a normal distribution: $X_1, X_2, \dots, X_n \sim N(\mu, \sigma)$. A very use fact is that the distribution of the sample average is *also normally distributed*:

$$\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

We notice that the mean is the same as for individuals, μ , and the standard deviation changes to σ/\sqrt{n} . The standard deviation for an average is *smaller* than the standard deviation for an individual item.

5. **Example** – Consider an SRS $X_1, X_2, \dots, X_9 \sim N(80, 15)$. Then, from the theorem above:

$$\bar{X} \sim N\left(80, \frac{15}{\sqrt{9}} = 5\right)$$

Again, the important point to see here is that *averages vary*. Each time you take a sample from a population you are going to get a *different* average. Think about it. Take a sample 20 people from campus and average their weights. Then pick another sample of 20 people and average their weights. The averages will be different. Thus, averages vary. This means that when we calculate a single average, we would like to know how much it varies. This tells us how *accurate* \bar{x} is, how much *error* there is in \bar{x} . In other words, if I obtained a sample average of 48.3, would *another* sample be more likely to have an average of 62 (large variability) or 47 (small variability).

6. **Standard error of the mean** – $\frac{\sigma}{\sqrt{n}}$ is called the *standard error of the mean* and is the *standard deviation of the sample average*. In other words, it measures the variability in a sample average.
7. Now we can calculate probabilities about averages. The only difference between asking questions about individuals and about averages is that with averages, we divide the standard deviation by \sqrt{n} .
8. **Example** – A company that owns and services a fleet of cars for its sales force has found that the service lifetime of disc brake pads varies from car to car according to a *normal distribution* with mean 55,000 miles and standard deviation 4500 miles. Find the probability that the average lifetime for 8 randomly selected cars is less than 51,800 miles.

$$\bar{X}_8 \sim N\left(55,000, \frac{4500}{\sqrt{8}} = 1590.99\right)$$

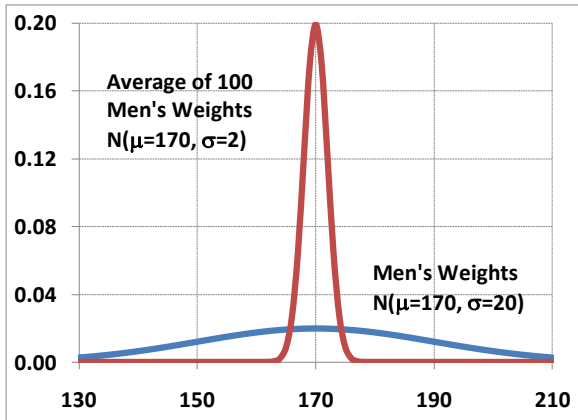
$$P(\bar{X}_8 < 51,800) = \text{normalcdf}(-1E99, 51800, 55000, 1590.99) = 0.0221 \approx \frac{2}{100}$$

9. **Example** – Men’s weights are normally distributed with mean 170 pounds and standard deviation 20 pounds. A sample of 100 men’s weights is taken.

a. What is the distribution of the sample average?

$$\bar{X}_{100} \sim N\left(170, \frac{20}{\sqrt{100}}\right) = N(170, 2)$$

b. Graph $f(x)$, the density function of men’s weights, and $f(\bar{x})$, the density function of the average of 100 men’s weights.



c. Find the probability that a randomly chosen man weighs less than 166 lbs.

$$P(X < 166) = \text{normalcdf}(-1E99, 166, 170, 20) = 0.4207 \approx 0.42 = 42\%$$

d. Suppose that you take a SRS of size 100 men. Find the standard error of the mean.

$$\sigma_{\bar{X}} = \frac{20}{\sqrt{100}}$$

e. Find the probability that if you take a SRS of size 100 men, their average weight will be less than 166 lbs.

$$P(\bar{X}_{100} < 166) = \text{normalcdf}(-1E99, 166, 170, 2) = 0.0228 \approx 2\%$$

Homework 5.4

1. Total sales in a store per day is normally distributed with a mean of \$6500 and standard deviation of \$1200. Suppose that sales from day-to-day are independent of one another.

- Consider the next 9 days. What is the probability that the average sales will be less than \$6000?
- Find the standard error of the mean for problem a.
- Graph the distribution of sales and the distribution of average sales over 9-day periods.
- Consider the next 25 days. What is the probability that the average sales will be less than \$7000?

5.6 – The Central Limit Theorem

1. **The Central Limit Theorem (CLT)** is the basis for most of the statistical procedures that we will learn in this class. It states that:

If you take an SRS of size n from ANY distribution with mean μ and standard deviation σ , then the sampling distribution of \bar{X} is *approximately* normal with mean μ and standard deviation σ/\sqrt{n} as the sample size gets large ($n \geq 30$).

2. Comparison to earlier theorem:

Earlier: $X_1, X_2, \dots, X_n \sim N(\mu, \sigma)$ means $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

CLT: $X_1, X_2, \dots, X_n \sim \text{???}(\mu, \sigma)$ means $\bar{X} \sim N\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$

3. Interpretation

In other words, the CLT says that if the distribution of a random variable is unknown or not given, then you can assume that distribution of \bar{X} is normal if you have a *large sample*. This basically says, “who cares what the distribution of the data is as long as we have a *large sample*.” This means that the data could be clearly skewed as long as we have at least 30 observations and the resulting *average of* 30 observations will be approximately normally distributed. For the purposes of our class, we will consider *large sample* to mean 30 or more data values. In practice, we may require larger samples.

4. Example – An electrical component is designed to provide a mean service life of 3000 hours with a standard deviation of 800 hours. A customer purchases a batch of 50 components. What is the probability that the mean life for the sample of 50 components will be at least 2750 hours?

Because of the Central Limit Theorem, $\bar{X} \sim N\left(3000, \frac{800}{\sqrt{50}}=113.137\right)$

$$P(\bar{X} > 2750) = 1 - \text{normalcdf}(-1E99, 2750, 3000, 113.137) = 0.9864 \approx 99\%$$

5. One hundred small bolts are packed in a box. Each bolt has mean weight 1 ounce and standard deviation 0.01 ounce. Find the probability that the average weight of the bolts in the box is less than 1.0015 ounces.

Because of the Central Limit Theorem, $\bar{X} \sim N\left(1, \frac{0.01}{\sqrt{100}}=0.001\right)$

$$P(\bar{X} < 1.0015) = \text{normalcdf}(-1E99, 1.0015, 1, 0.001) = 0.9332 \approx 93\%$$

6. The scores of students on the ACT college entrance examination in a recent year had a $\mu = 18.6$ and $\sigma = 5.9$. The average score of 76 students at Northside High who took the test was $\bar{x} = 20.4$. What is the probability that the mean score for 76 students chosen randomly from all who took the test nationally is 20.4 or higher?

If you choose 76 students at random from the entire population of people who took the ACT, then we would expect, because of the Central Limit Theorem

$$\bar{X}_{76 \text{ Chosen Randomly}} \sim N\left(18.6, \frac{5.9}{\sqrt{76}}=0.677\right)$$

Thus,

$$P(\bar{X} > 20.4) = 1 - \text{normalcdf}(-1E99, 20.4, 18.6, 0.677) =$$

$$0.0039 \approx 0.004 = \frac{4}{1000}$$

What does this value tell us?

7. A company that owns and services a fleet of cars for its sales force has found that the service lifetime of disc brake pads varies from car to car according to a normal distribution with mean 55,000 miles and standard deviation 4500 miles. The company installs a new brand of brake pads on 8 cars. The average life of the pads on these 8 cars turns out to be 51,800 miles. What is the probability that the sample mean lifetime is 51,800 miles or less if the lifetime distribution is unchanged? What does this say about the lifetime distribution of the new brake pads?

$$\text{If } \bar{X}_{Old} = \bar{X}_{New} \text{ then } \bar{X}_{New} \sim N\left(55,000, \frac{4500}{\sqrt{8}}=1590.9\right)$$

$$P(\bar{X}_{New} < 51,800) = \text{normalcdf}(-1E99, 51800, 1590.9) = 0.0222 \approx \frac{2}{100}$$

What if the average of the 8 cars had been 54,074?

$$P(\bar{X}_{New} < 54.074) = \text{normalcdf}(-1E99, 54074, 1590.9) = 0.2699 \approx \frac{3}{10}$$

Homework 5.5

1. The number of calories in a serving of soup has a mean of 400 and standard deviation of 18. Consider a sample of 36 servings of soup.
 - a. What is the probability that the average number of calories will be less than 405?
 - b. What theorem allows you to solve this problem?
 - c. What is the probability that the average number of calories will be between 397 and 403?