

# Using Your TI-83/84 Calculator: Normal Probability Distributions

## Elementary Statistics

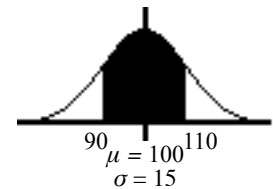
### Dr. Laura Schultz

Always start by drawing a sketch of the normal distribution that you are working with. Shade in the area (probability) that you are given or trying to find, and label the mean, standard deviation, lower bound, and upper bound that you are given or trying to find. Don't worry about making your drawing to scale; the purpose of the sketch is to get you thinking clearly about the problem you are trying to solve. For illustration purposes, let's consider the distribution of adult scores on the Weschler IQ test. These IQ scores are normally distributed with  $\mu = 100$  and  $\sigma = 15$ .

### Using the normalcdf command

The normalcdf command is used for finding an area under the normal density curve. This area corresponds to the probability of randomly selecting a value between the specified lower and upper bounds. You can also interpret this area as the percentage of all values that fall between the two specified boundaries.

- Let's find the percentage of adults who score between 90 and 110 on the Weschler IQ test. Begin by sketching the distribution and labeling the relevant information. We are ultimately trying to find the area under the normal density curve that is bounded by 90 and 110, so shade in that area on your sketch.



- Press **2nd****[VARS]** for the **[DISTR]** menu. Scroll down to 2:normalcdf( and then press **[ENTER]**.
- Enter the two  $x$  values (or  $z$  scores) that form the boundaries of the area that you are trying to find, the population mean, and the population standard deviation using the following syntax: Lowerbound, upperbound,  $\mu$ ,  $\sigma$ ) and then press **[ENTER]**. For this example, type 90, 110, 100, 15) and then press **[ENTER]**.

```

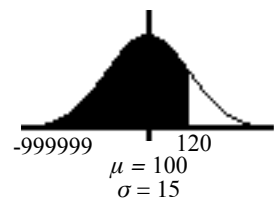
DISTR DRAW
1:normalPdf(
2:normalcdf(
3:invNorm(
4:invT(
5:tPdf(
6:tCdf(
7:χ²Pdf(
    
```

```

normalcdf(90,110
,100,15)
.495015066
    
```

- Your calculator will return the area under the normal curve bounded by 90 and 110. Thus, we find that 49.5% of adults score between 90 and 110 on the Weschler IQ test. Remember to round percentages to three significant figures.

- What is the probability that a randomly selected adult scores less than 120 on the Weschler IQ test? Problems like this are a bit trickier because your calculator requires both a lower bound and an upper bound. Ultimately, you need to specify an approximation of  $-\infty$  for the lower bound. In most cases, -999999 is a good choice for the lower bound. Hence, the syntax for problems of this sort is normalcdf(-999999, upperbound,  $\mu$ ,  $\sigma$ ).

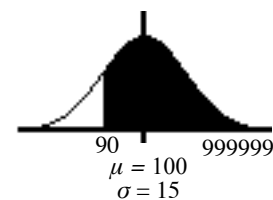


- Start by drawing a sketch. Then, press **2nd****[VARS]** for the **[DISTR]** menu. Scroll down to 2:normalcdf( and press **[ENTER]**. Type **(-)**999999, 120, 100, 15) and then press **[ENTER]** again. We find that the probability of randomly selecting an adult whose IQ is below 120 is 0.909. In symbols,  $P(x \leq 120) = 0.909$ . Remember to round probability values to 3 significant figures.

```

normalcdf(-99999
9,120,100,15)
.9087887181
    
```

- What percentage of adults score at least 90 on the Weschler IQ test? Because your calculator requires both a lower bound and an upper bound, you will need to specify an approximation of  $\infty$  for your upper bound. Using 999999 is good enough in most cases. Hence, the syntax for problems of this sort is `normalcdf(lowerbound, 999999,  $\mu$ ,  $\sigma$ )`.
- Start by drawing a sketch. Then, press `2nd``[VARS]` for the `[DISTR]` menu. Scroll down to `2:normalcdf(` and press `[ENTER]`. Type `90, 999999, 100, 15)` and then press `[ENTER]` again.
- We find that 74.8% of adults score at least 90 on the Weschler IQ test. Remember to round to 3 significant figures.



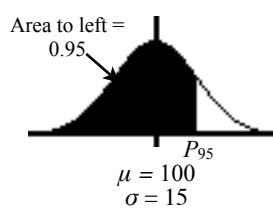
```
normalcdf(90, 999
999, 100, 15)
.747507533
```

### Using the invNorm command

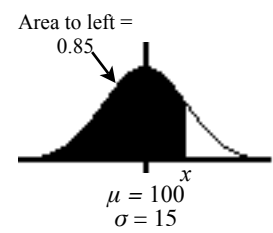
Use the `invNorm` command when you are given a probability or percentage and asked to find an  $x$  value or  $z$  score. This command is often used to find values corresponding to percentiles or quartiles. Your calculator requires that you enter the cumulative area to the left of the desired value; drawing a sketch is very useful for making sure you enter the correct area. Sometimes you will need to work with an area other than the one specified by the problem.

- Press `2nd``[VARS]` for the `[DISTR]` menu. Scroll down to `3:invNorm(` and press `[ENTER]`.
- Let's start by finding the IQ score corresponding to the 95<sup>th</sup> percentile ( $P_{95}$ ). Begin by drawing a sketch and labeling it.
- Enter the total area to the left of the desired  $x$  value (or  $z$  score) using the following syntax: `invNorm(area to left,  $\mu$ ,  $\sigma$ )` and then press `[ENTER]`. For this example, type `95, 100, 15)` and then press `[ENTER]`.
- Your calculator will return the  $x$  value (or  $z$  score). For this example, we find that the Weschler IQ score corresponding to the 95<sup>th</sup> percentile is 124.7. Round your answer to one more decimal place than what was provided for  $\mu$ .
- Let's try a trickier example. This time, find the IQ score separating the top 15% of all Weschler IQ scores from the rest. Start by drawing a sketch. Given that the total area under the normal density curve is always 1, the area to the left of the IQ score that we are seeking can be found by subtracting the given area from 1. In this case,  $1 - 0.15 = 0.85$ . Don't forget to convert percentages to their equivalent decimal values.
- Press `2nd``[VARS]` for the `[DISTR]` menu. Scroll down to `3:invNorm(` and press `[ENTER]`.
- The area to the left of the IQ score we are seeking is 0.85, so type `0.85, 100, 15)` and then press `[ENTER]`. We find that an IQ score of 115.5 separates the top 15% of all adult Weschler IQ scores from the rest.

```
0:QUIT DRAW
1:normalpdf(
2:normalcdf(
3:invNorm(
4:invT(
5:tpdf(
6:tcdf(
7:Xdpdf(
```



```
invNorm(0.95, 100
, 15)
124.6728044
```



```
invNorm(0.85, 100
, 15)
115.5465007
```

## Applying the Central Limit Theorem (Working with sample means)

- The Central Limit Theorem applies whenever you are working with a distribution of sample means ( $\bar{x}$ ), and the sample comes from a normally distributed population, and/or the sample size is at least 30 ( $n \geq 30$ ).
- Recall that the mean for a distribution of sample means is  $\mu_{\bar{x}} = \mu$ , and the standard deviation for a distribution of sample means is  $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ .
- Thus, the modified calculator commands to use when you are applying the Central Limit Theorem to work with sample means ( $\bar{x}$ ) are as follows:

$$\text{normalcdf}(\text{lowerbound}, \text{upperbound}, \mu, \frac{\sigma}{\sqrt{n}})$$

$$\text{normalcdf}(-999999, \text{upperbound}, \mu, \frac{\sigma}{\sqrt{n}})$$

$$\text{normalcdf}(\text{lowerbound}, 999999, \mu, \frac{\sigma}{\sqrt{n}})$$

$$\text{invNorm}(\text{area to left}, \mu, \frac{\sigma}{\sqrt{n}})$$

# 13.1 and 13.2 Permutations and Combinations

- Solve problems related to the Basic Counting Principle.
  - Distinguish between dependent and independent events.
    - Solve problems involving permutations or combinations.
- Distinguish between linear and circular Permutations

- I. Tree Diagram, Basic Counting Principal, Independent Vs Dependent Events (13.1)
  - Suppose that a school has a strict dress code. If a boy in the school can select from 3 different ties and 4 pairs of pants how many dress options does he have each day?

## **Tree Diagram**

Each category is represented by a branch of the tree.  
Follow the branches to determine all of the outcomes.

## **Basic Counting Principle**

Suppose one event can be chosen in  $p$  different ways, and another independent event can be chosen in  $q$  different ways. Then the two events can be chosen successively in  $p \cdot q$  ways.

## **Independent events**

Option A does not affect the choices in option B.

## **Dependent events**

Option A affects the choices in option B.



Ex 1: JT has decided to buy a new suit made either of wool or rayon. He has narrowed down the choices of available colors to gray, blue, black or tan, and the matching tie to a floral, a stripe, or a geometric pattern.

- a) Assuming that all choices make an acceptable appearance, are the choices for material, color, and tie pattern independent or dependent events.

Independent- since they are all acceptable combinations, the choice of one does not affect the choice of the other.

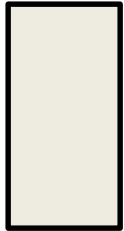
- b) How many different selections of his suit and tie are possible?

2 materials, 4 colors, 3 ties

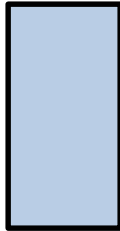
$$2 * 4 * 3 = 24 \text{ options}$$

Ex 2: Vickie works for a bookstore. Her manager asked her to arrange a set of five best-sellers for a display. The display is to be set up as shown below. The display set is made up of one book from each of 5 categories. There are 4 nonfiction, 4 science fiction, 3 history, 3 romance and 3 mystery books from which to choose.

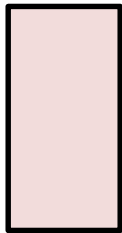
Nonfiction



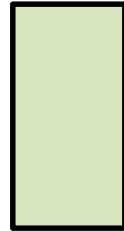
Sci Fi



History



Romance



Mystery



a) Are the choices for each book independent or dependent events?

Independent- the choice of the book for each slot does not affect the choice for the other slots.

a) How many different ways can Vickie choose the books for the display?

$$\begin{array}{cccccc} 1^{\text{st}} \text{ spot} & & 2^{\text{nd}} \text{ spot} & & 3^{\text{rd}} \text{ spot} & & 4^{\text{th}} \text{ spot} & & 5^{\text{th}} \text{ spot} \\ 4 & * & 4 & * & 3 & * & 3 & * & 3 \\ & & & & = 432 \text{ different arrangements} \end{array}$$



## II. Permutations and Combinations (13.1)

### Permutations

The arrangement of objects in a certain order.

The order of the objects is very important.

$p(n, n) = n!$       Number of permutations of  $n$  objects, taken  $n$  at a time.

$P(n, r) = \frac{n!}{(n-r)!}$       Number of permutations of  $n$  objects, taken  $r$  at a time.

### Combinations

The arrangement of objects when the order is not important.

$C(n, r) = \frac{n!}{r!(n-r)!}$       Number of combinations of  $n$  objects taken  $r$  at a time.

What is the difference between permutations & combinations?

Lets take the following letters two at a time: A,B,C,D

Permutations

AB BA CA DA  
AC BC CB DB  
AD BD CD DC

Combinations

AB BC  
AC BD  
AD CD

Notice that for permutations, AB and BA are different because order matters. For combinations, they are the same thing because order does not matter.

1) A math teacher needs to place five books- algebra, algebra 2, geometry, trigonometry, and advanced math- on her bookshelf. Assume that all of the books are placed upright with the binded edges in view.

a) Independent or Dependent events?

Dependent

b) How many ways can the books be displayed on the shelf?

This is 5 objects, taken 5 at a time. The order matters, so it is a permutation.

$$p(n, n) = n!$$

$$p(5,5) = 5!$$

= 120 different ways

- 2) Every year a popular magazine picks the 15 top-rated television movies of the previous year. A network wants to select 3 out of these 15 movies to show during the month of August. How many groups of 3 movies can be selected?

This is 15 objects taken 3 at a time. Order is not important in each group, so this is a combination.

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

$$C(15, 3) = \frac{15!}{3!(15-3)!}$$

455 different groups of three movies can be chosen.

3) A high school honor society is composed of 7 students.

a) How many different ways can all the members stand in a straight line for a picture?

The order matters, so this is a permutation of 7 objects taken 7 at a time.

$$p(n, n) = n!$$

$$p(7, 7) = 7!$$

5040 different ways

b) In how many ways can they select a president and a vice-president, assuming that one person cannot hold more than one office?

This is a permutation of 7 people taken 2 at a time.

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

$$P(7, 2) = \frac{7!}{(7 - 2)!}$$

42 different ways

### III. Permutations with Repetitions (13.2)

The number of permutations of  $n$  objects of which  $p$  are alike and  $q$  are alike is

$$\frac{n!}{p!q!}$$

1) How many 5-letter patterns can be formed from the letters of the word *graph*?

There are no repeats, so we just use  $n!$

$$5! = 120 \text{ different combinations}$$

2) How many 5-letter patterns can be formed from the letters of the word *jello*?

There are two ls

$$\frac{n!}{p!q!} = \frac{5!}{2!} = 60 \text{ different combinations}$$

3) How many 11-letter patterns can be formed from the letters of the word *Mississippi*?

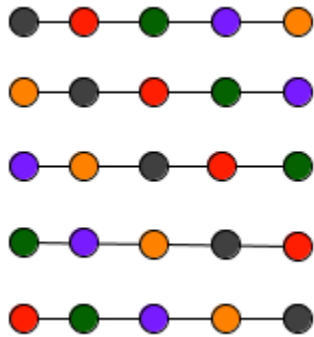
There are 4 s's, 4 i's, 2 p's

$$\frac{n!}{p!q!} = \frac{11!}{4!4!2!} = 34650 \text{ different combinations}$$

## IV. Circular Permutations

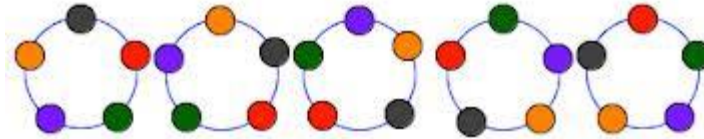
Linear:

We start counting from the first point, so we have 5 different patterns

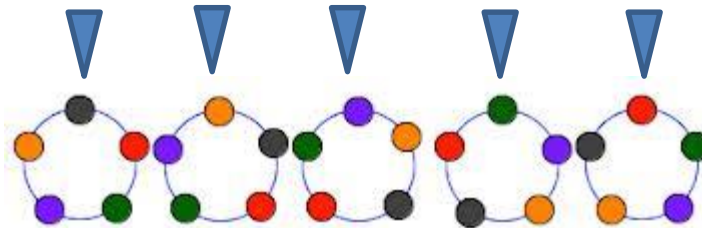


Circular: Unless otherwise stated, we don't have a set start or end point. This means that the five patterns are actually one pattern that repeats itself.

$\frac{n!}{n}$  or  $(n - 1)!$  Permutations of the  $n$  objects.



Circular with a fixed point: We have a set starting point, so this is equivalent to a linear permutation.





1) Seven people are to be seated at a round table. How many ways can the people be arranged?

Circular permutation, no reference point

$(7-1)! = 6!$  Or 720 different ways

2) Seven people are to be seated at a round table where one person is seated next to a window. How many ways can the people be arranged relative to the window?

Circular permutation with a reference point

$7!$  Or 5040 different ways

3) Martha is at a picnic that has a circular, revolving condiment server. Six condiments are placed on the table. How many ways can the condiments be arranged?

Circular permutation, no reference point

5! Or 120 different ways

# Notes: 13.3 Probability and Odds

## I. Probability

Probability: The likelihood that a given event will occur.

If an event can succeed in 's' ways and fail in 'f' ways, then the probability of success and the Probability of failure are determined as follows:

$$P_{(s)} = \frac{s}{s+f} \qquad P_{(f)} = \frac{f}{s+f}$$

## A. Simple Cases

Ex 1: Before 2001, The US had 42 Presidents including Clinton. Only two of them (George Washington and John Adams) belonged to the Federalist Party. If you were to select a president at random for a history report, what is the probability that your selection would be a Federalist?

$$P_{(s)} = \frac{s}{s + f} = \frac{2}{2 + 40} = \frac{2}{42} = \frac{1}{21}$$

Ex 2: A sock drawer contains single, unmatched socks: 19 are white, 8 are black, and 5 are blue.

a. What is the probability that a sock selected at random will be black?

$$P_{(black)} = \frac{s}{s + f} = \frac{8}{8 + 24} = \frac{8}{32} = \frac{1}{4}$$

b. What is the probability that it will NOT be white?

$$P_{(not\ white)} = \frac{s}{s + f} = \frac{13}{8 + 24} = \frac{13}{32}$$

## B. Using Counting Methods (Permutations and Combinations)

- Ex 1: A box of 60 baseball cards contains 8 cards with print errors on them. If 5 cards are selected at random, what is the probability that all 5 have print errors?

$P_{(s)} = \frac{s}{s+f}$  Think of this as all possible ways you could draw 5 error cards out of the 8 error cards, divided by all the ways you could draw 5 cards out of 60. We then have:

$$P_{(5 \text{ print errors})} = \frac{C(8,5)}{C(60,5)} = \frac{1}{97,527}$$

## C. Complements

Since an event can either succeed or fail, the probability of either occurring is 1.

$$P_{(s)} + P_{(f)} = 1$$

Because their sum is one,  $P_{(s)}$  and  $P_{(f)}$  are called **complements**.

- Ex 2: A box contains 35 machine parts, 8 of which are defective. A quality control inspector randomly selects 5 of the 35 parts for testing. What is the probability that at least one part is defective?

Just think about all the different permutations here: the question is asking for the probability of there being 1 or 2 or 3 or 4 or 5 defectives in his sample: call this  $P(A)$ . This is very time consuming and boring. It is faster to find the complement and then subtract this from 1.

$P(\text{at least 1 defective part}) = 1 - P(\text{no defective parts})$

$$= 1 - \frac{\text{\# of possible groups of 5 non - defective parts}}{\text{\# of possible groups of 5 parts}}$$

$$= 1 - \frac{C(27,5)}{C(35,5)} \approx .7513$$



## II. Odds

Odds: The odds of a successful outcome of an event is the ratio of the probability of its success to the probability of its failure.

$$\mathbf{Odds} = \frac{\mathbf{P}_{(s)}}{\mathbf{P}_{(f)}}$$

Odds is a comparison of possible outcomes, not the possibility of an outcome.

**Example:** A game consists of 24 cards that say “Move Ahead One Space” and 10 cards that say “Move Back One Space”. A card may be chosen randomly from anywhere in the deck of 34 cards.

- a. What is the probability of drawing a “Move Ahead” card?

$$P_{(s)} = \frac{s}{s + f} = \frac{24}{24 + 10} = \frac{24}{34} = \frac{12}{17}$$

- b. What are the odds that a player will move his or her token ahead one space on any given turn.

$$Odds = \frac{P_{(s)}}{P_{(f)}} \quad P_{(f)} = \frac{f}{s+f} = \frac{10}{24+10} = \frac{10}{34} = \frac{5}{17}$$

$$Odds = \frac{P_{(s)}}{P_{(f)}} = \frac{\frac{12}{17}}{\frac{5}{17}} = \frac{12}{5}$$

**Example:** Determine the probability of picking a diamond out of a standard deck of cards.

$$P_{(s)} = \frac{s}{s+f} = \frac{13}{13+39} = \frac{13}{52} = \frac{1}{4}$$

**Example:** Determine the odds of picking a diamond out of a standard deck of cards.

$$Odds = \frac{P_{(s)}}{P_{(f)}} \quad P_{(f)} = \frac{f}{s+f} = \frac{39}{13+39} = \frac{3}{4}$$

$$Odds = \frac{P_{(s)}}{P_{(f)}} = \frac{\frac{1}{4}}{\frac{3}{4}} = \frac{1}{3}$$

**Example:** On the volleyball team, there are 5 freshmen, 6 sophomores, 9 juniors, and 4 seniors. If one is selected at random, what is the probability that the student is a freshman?

$$P_{(s)} = \frac{s}{s+f} = \frac{5}{24}$$

**Example:** What are the odds that the student is a freshman?

$$\text{Odds} = \frac{P_{(s)}}{P_{(f)}} \qquad P_{(f)} = \frac{f}{s+f} = \frac{19}{24}$$

$$\text{Odds} = \frac{P_{(s)}}{P_{(f)}} = \frac{\frac{5}{24}}{\frac{19}{24}} = \frac{5}{19}$$

### III. Finding a Sample Space

- Four male and 2 female students have been selected as equal qualifiers for 2 college scholarships. If the awarded recipients are to be chosen at random, what are the odds that one will be male and one will be female?

First, determine the total number of possible groups.

$C(4, 1)$  number of groups of one male

$C(2, 1)$  number of groups of one female

Using the Basic Counting Principle, we can find the number of possible groups of 1 male and 1 female.

$$C(4, 1) \cdot C(2, 1) = 8$$

The total number of groups who qualified is  $C(6, 2)$  or 15.

So, the number of groups that do not have 1 male and 1 female is 7.

Finally, determine the odds.

$$\text{Odds} = \frac{P_{(s)}}{P_{(f)}} = \frac{\frac{8}{15}}{\frac{7}{15}} = \frac{8}{7}$$

More Practice Problems:

**One card is drawn at random from an ordinary deck of cards. Find the probability and odds of each event.**

**a) a red face card**

**b) 6 of clubs**

**c) a king**

**d) not a king**

**Example:** Two dice are thrown. Find the probability that both dice show the same number.

**A number wheel is divided in 12 equal sectors, numbered 1 – 12.**

**Find the probability and odds of each event:**

**a) It is a 5**

**b) It is a 2-digit number**

**c) It is a factor of 12**

**d) It is an odd number**

**e) It is a negative number**

**f) It is a 3 or a 6 or a 9**



**Suppose two letters are picked at random from the word 'computer'. What is the probability that both of the letters are consonants?**

**What are the odds that both letters are consonants?**

**A letter is selected at random from those in the word  
RECTANGLE. Find the probability and odds of each  
event.**

**a) It is a vowel**

**b) It is a consonant**

**c) It is between D and M in the alphabet**

**d) It is either C or G**

# Notes 13.4 Probability of Compound Events

## I. Independent and Dependent Events

### A. Definitions

**Probability of Two Independent Events:** If two events, A and B, are both independent, then the probability of both events occurring is...

$$P(A \text{ and } B) = P(A) * P(B)$$

**Probability of Two Dependent Events:** If two events, A and B, are dependent, then the probability of both events occurring is...

$$P(A \text{ and } B) = P(A) * P(B \text{ following } A)$$

## B. Examples of Independent and Dependent Events

**Example of Independent Events:** What is the probability of rolling a 4 on a die 3 times in a row?

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

$$P(\text{three 4s}) = \begin{matrix} \text{Roll 1} & \text{Roll 2} & \text{Roll 3} \\ P(4) * & P(4) * & P(4) \end{matrix}$$

$$= \left(\frac{1}{6}\right)^3$$

$$= \frac{1}{216}$$

**Example of Dependent Events:** What is the probability of drawing 6 hearts from a deck of cards without replacement?

$$\begin{aligned} P(\text{six hearts}) &= \begin{array}{cccccc} \text{Draw 1} & \text{Draw 2} & \text{Draw 3} & \text{Draw 4} & \text{Draw 5} & \text{Draw 6} \\ \frac{13}{52} & * & \frac{12}{51} & * & \frac{11}{50} & * & \frac{10}{49} & * & \frac{9}{48} & * & \frac{8}{47} \end{array} \\ &= \frac{1235520}{14658134400} \\ &= \frac{3861}{4580667} \end{aligned}$$

## C. More Examples

**Example 1:** Suppose the odds of the Sixers beating the Kings in Basketball was 5 : 2. What is the probability of the Sixers beating the Kings 4 times in a row?

$$\text{Odds} = \frac{5}{2} \quad \frac{P(\text{Sixers win})}{P(\text{Kings win})} = \frac{\frac{5}{7}}{\frac{2}{7}} \quad P(\text{Sixers win}) = \frac{5}{7}$$

$$\begin{aligned} P(4 \text{ wins}) &= \begin{array}{cccc} \text{Game 1} & \text{Game 2} & \text{Game 3} & \text{Game 4} \\ \frac{5}{7} & * & \frac{5}{7} & * & \frac{5}{7} & * & \frac{5}{7} \end{array} \\ &= \left(\frac{5}{7}\right)^4 = \frac{625}{2401} \end{aligned}$$

**Example 2:** A particular bag of marbles contains 4 red, 6 green, 2 blue, and 5 white marbles. What is the probability of picking a red, white, and blue marble, in that order?

$$P(r,w,b) = \begin{array}{c} \textit{Pick 1} \\ \frac{4}{17} \end{array} * \begin{array}{c} \textit{Pick 2} \\ \frac{5}{16} \end{array} * \begin{array}{c} \textit{Pick 3} \\ \frac{2}{15} \end{array} \quad P(r,w,b) = \frac{1}{102}$$

What would the probability be with replacement?

$$P(r,w,b) = \begin{array}{c} \textit{Pick 1} \\ \frac{4}{17} \end{array} * \begin{array}{c} \textit{Pick 2} \\ \frac{5}{17} \end{array} * \begin{array}{c} \textit{Pick 3} \\ \frac{2}{17} \end{array} \quad P(r,w,b) = \frac{40}{4913}$$

## II. Mutually Exclusive and Inclusive Events

### A. Definitions

**Mutually Exclusive Events:** If two events, A and B, are mutually exclusive, then that means that if A occurs, then B cannot, and vice versa.

**Probability of Mutually Exclusive Events:** If two events, A and B, are mutually exclusive, then the probability that either A **OR** B occurs is...

$$P(A \text{ or } B) = P(A) + P(B)$$



**Inclusive Events:** If two events, A and B, are inclusive, then that means that if A occurs, B could also occur, and vice versa.

**Probability of Inclusive Events:** If two events, A and B, are inclusive, then the probability that either A or B occurs is...

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

## B. Examples

**Example 1:** A particular bag of marbles contains 4 red, 6 green, 2 blue, and 5 white marbles. If 3 marbles are picked, what is the probability of picking all reds or all greens?

mutually exclusive event

$$P(\text{red or greens}) = P(\text{red}) + P(\text{green})$$

$$= \left[ \frac{4}{17} * \frac{3}{16} * \frac{2}{15} \right] + \left[ \frac{6}{17} * \frac{5}{16} * \frac{4}{15} \right]$$

$$= \frac{24}{4080} + \frac{120}{4080}$$

$$= \frac{144}{4080} = \frac{3}{85}$$

$$= \frac{C_{(4,3)}}{C_{(17,3)}} + \frac{C_{(6,3)}}{C_{(17,3)}}$$

$$= \frac{4}{680} + \frac{20}{680}$$

$$= \frac{24}{680} = \frac{3}{85}$$

**Example 2:** Slips of paper numbered 1 to 15 are placed in a box. A slip of paper is drawn at random. What is the probability that the number picked is either a multiple of 5 or an odd number?

inclusive event

$$P(\text{mult of 5 or odd}) = P(\text{mult of 5}) + P(\text{odd}) - P(5 \text{ and odd})$$

$$= \frac{3}{15} + \frac{8}{15} - \frac{2}{15}$$

$$= \frac{9}{15} = \frac{3}{5}$$

**Example 3:** Two cards are picked out of a standard deck.

What is the probability of both cards being either face cards or clubs?

inclusive event

$$P(\text{face or clubs}) = P(\text{face}) + P(\text{club}) - P(\text{face and club})$$

$$= \left[ \frac{12}{52} * \frac{11}{51} \right] + \left[ \frac{13}{52} * \frac{12}{51} \right] - \left[ \frac{3}{52} * \frac{2}{51} \right]$$

$$= \frac{132}{2652} + \frac{156}{2652} - \frac{6}{2652}$$

$$= \frac{282}{2652} = \frac{47}{442}$$

Ex 4. In a particular group of hospital patients, the probability of having high blood pressure is  $\frac{3}{8}$ , the probability of having arteriosclerosis is  $\frac{5}{12}$ , and the probability of having both is  $\frac{1}{4}$ .

- a) Determine whether the events are mutually exclusive or mutually inclusive.

Mutually inclusive

- a) What is the probability that a patient in this group has either high blood pressure or arteriosclerosis?

$$\frac{3}{8} + \frac{5}{12} - \frac{1}{4} = \frac{13}{24}$$

Probability of Two Independent Events:

$$P(A \text{ and } B) = P(A) * P(B)$$

Probability of Two Dependent Events:

$$P(A \text{ and } B) = P(A) * P(B \text{ following } A)$$

Probability of Mutually Exclusive Events:

$$P(A \text{ or } B) = P(A) + P(B)$$

Probability of Inclusive Events:

$$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$$

## Notes 13.5: Conditional Probability

# Notes 13.5: Conditional Probability

## I. Definitions and Formulas

The probability of an event given that some other event has occurred. i.e. a reduced sample space.

It is written as  $P(B | A)$  - the probability of B given A.

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

E.g. Given that the toss of a die is even, what is the probability that it is divisible by three?

Method 1: Reduced sample space is  $\{2,4,6\}$   
 $P(\text{Divisible by 3} | \text{toss is even}) = \frac{1}{3}$

Method 2: Using  $P(B | A) = \frac{P(A \cap B)}{P(A)}$   
 $= \frac{1/6}{1/2} = \frac{1}{3}$



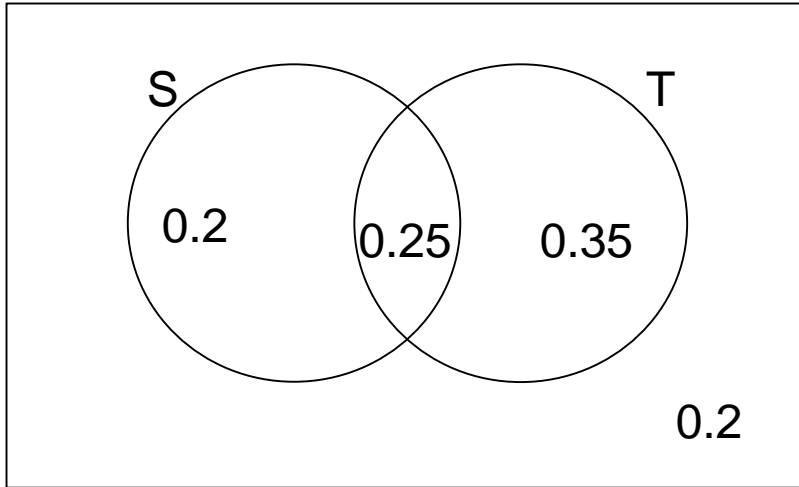
# If Independent.....

- If events A and B are independent, then event B will not be influenced by whether event A has occurred, and so  $P(B | A) = P(B)$

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

## II. Basic Examples

Ex 1:



$$P(S) = .2 + .25 = .45$$

$$P(T) = .35 + .25 = .6$$

$$P(S \cap T) = .25$$

$$P(S | T) =$$

$$P(T | S) =$$

$$P(S | T') =$$

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

$$= \frac{.25}{.6}$$

$$\approx .42$$

$$= \frac{.25}{.45}$$

$$\approx .56$$

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

$$= .5$$

Ex 2: A frog climbing out of a well is affected by the weather. When it rains, he falls back down the well with a probability of  $1/10$ . In dry weather, he only falls back down with probability of  $1/25$ . The probability of rain is  $1/5$ .

**Find the probability that given he falls it was a rainy day.**

Let's start by drawing the tree diagram of these events:

**We want:** 
$$P(\text{rains} \mid \text{falls}) = \frac{P(\text{rains and falls})}{P(\text{falls})}$$

**From the tree we can see:** 
$$P(\text{rains and falls}) = \frac{1}{50}$$

$$\text{and that } P(\text{falls}) = \frac{1}{50} + \frac{4}{125}$$

**Therefore:** 
$$P(\text{rains} \mid \text{falls}) = \frac{\frac{1}{50}}{\left(\frac{1}{50} + \frac{4}{125}\right)} = \frac{5}{13}$$

Ex 3: A math teacher gave her class two tests. 25% of the class passed both tests and 42% of the class passed the first test. What percent of those who passed the first test also passed the second test?

$$P(\text{Second}|\text{First}) = \frac{P(\text{First and Second})}{P(\text{First})} = \frac{0.25}{0.42} = 0.60 = 60\%$$

### III. Using tables

Conditional probabilities can be found simply from data in tables, as illustrated by the following.

The table opposite shows the choices of language and the gender of the 200 students choosing those languages.

	French	German	Total
Male	40	40	80
Female	90	30	120
Total	130	70	200

A student is choosing at random, find the probability of that student,

a) studying French,

$$\frac{130}{200}$$

b) being male,

$$\frac{80}{200}$$

c) being male and studying German,

$$\frac{40}{200}$$

d) being female, given he/she studies French.

$$\frac{90}{130}$$

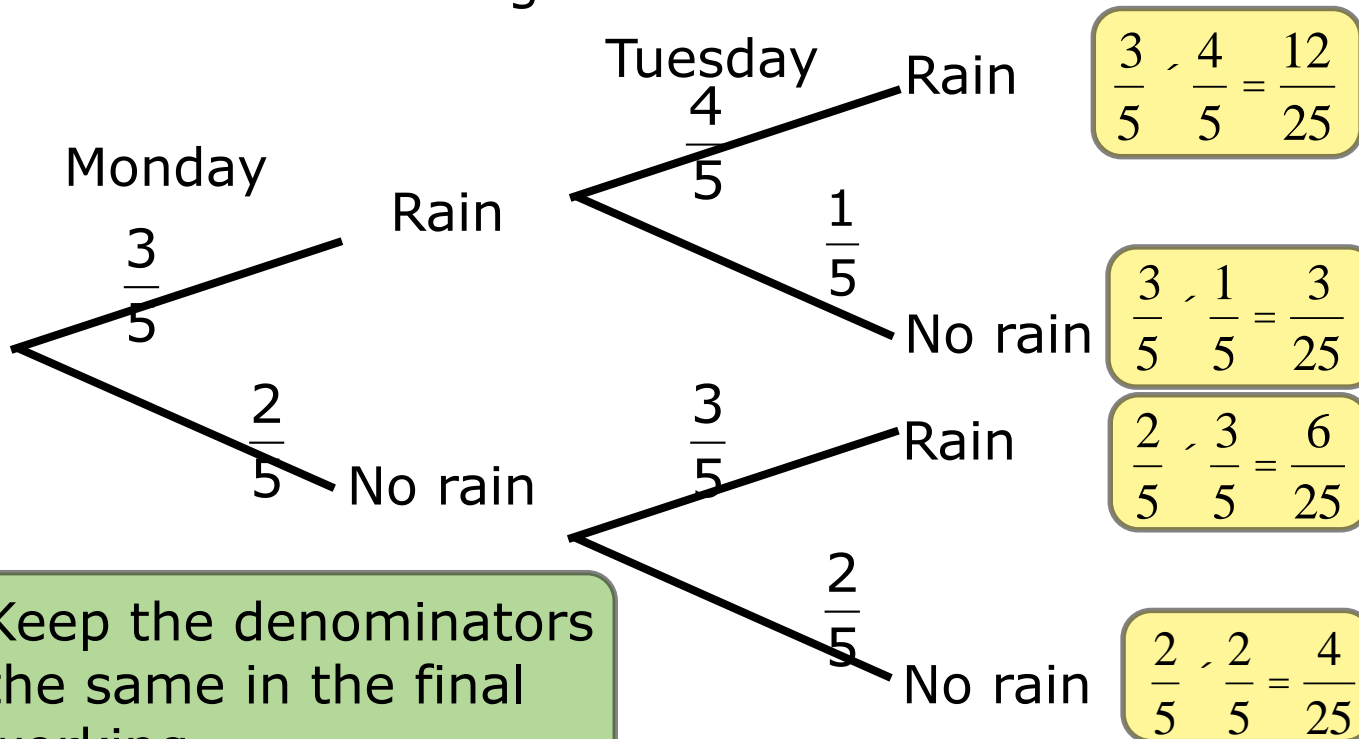
e) studying German, given that he is male.

$$\frac{40}{80}$$

## IV. Constructing a tree diagram

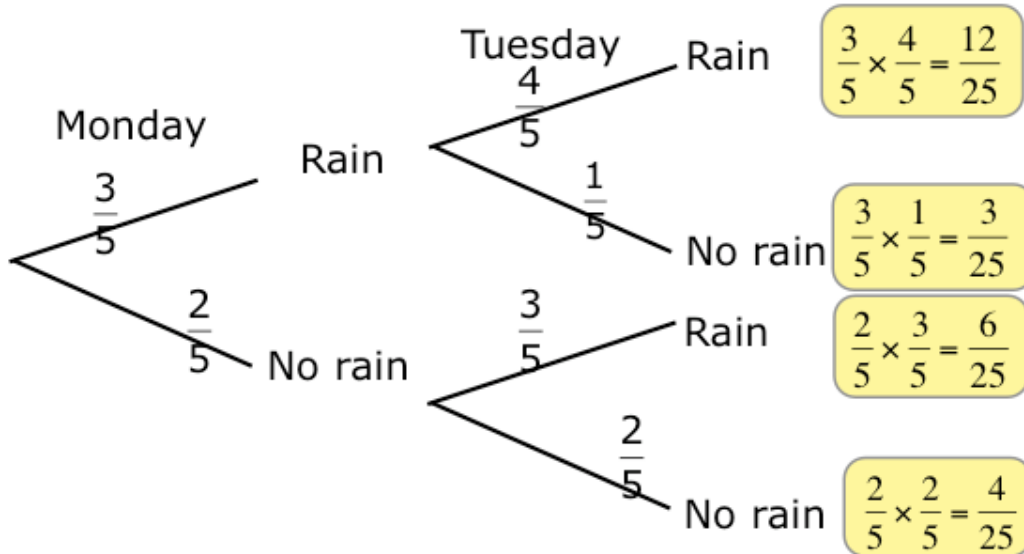
The probability of rain in Manchester on any Monday is  $\frac{3}{5}$ . If it rains on a Monday, then the probability of rain on Tuesday is  $\frac{4}{5}$ . If it does not rain on Monday the probability of rain on Tuesday is  $\frac{3}{5}$ .

Construct a tree diagram to show this information.



Keep the denominators the same in the final working.

# Probabilities



Calculate the probability that,

a) it rains at least once,

$$\frac{21}{25}$$

b) it rains one day only,

$$\frac{9}{25}$$

c) it rains on one day only, given it rains at least once.

$\frac{\text{rain on one day only}}{\text{rained at least once}}$

$$\frac{9}{9+12} = \frac{9}{21}$$

It is possible to ignore the denominators here only if they are the same at the end of your tree diagram.

# Additional Example

A teacher oversleeps with a probability of 0.3. If she oversleeps then the probability that she eats breakfast is 0.2, otherwise it will be 0.6.

a) Construct a tree diagram to show this information.

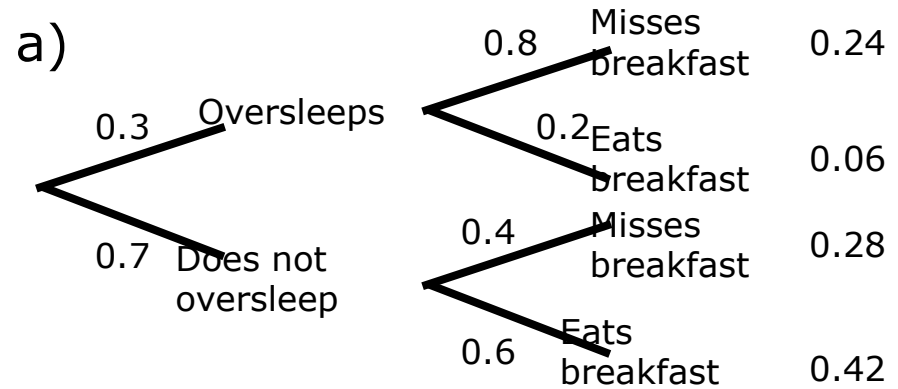
Use your diagram to find the probability that,

b) she oversleeps and does not eat breakfast,

c) she does not have breakfast,

d) she overslept, given he has breakfast,

e) she overslept, given he does not have breakfast.



b) 0.24

c) 0.52

d)  $\frac{6}{48}$

e)  $\frac{24}{52}$



# Monty Hall Problem

[http://www.nytimes.com/2008/04/08/science/08monty.html?\\_r=0](http://www.nytimes.com/2008/04/08/science/08monty.html?_r=0)

# Notes 13.6: Probability and Binomial Expansions

## I. Binomial Experiment

### A. Definition, Requirements, Formula

- Some probability problems can be solved using binomial expansions. This is called a ‘binomial experiment’.
- A binomial experiment is an experiment which satisfies these four conditions
  - 1. A fixed number of trials
  - 2. Each trial is independent of the others
  - 3. There are only two outcomes
  - 4. The probability of each outcome remains constant from trial to trial.

## **Examples of binomial experiments**

- Tossing a coin 20 times to see how many tails occur.
- Asking 200 people if they watch ABC news.
- Rolling a die to see if a 5 appears.

## **Examples which are NOT binomial experiments**

- Rolling a die until a 6 appears (not a fixed number of trials)
- Asking 20 people how old they are (not two outcomes)
- Drawing 5 cards from a deck for a poker hand (done without replacement, so not independent)

## I. Example of Probability and Binomial Expansion

- About 70% of the questions created by a random computer test generator for a science program are multiple choice. What is the probability that at least 3 of 5 questions generated for a short quiz are multiple choice?

We can represent the probability of each number of successes using a binomial expansion.

Let S = probability of a question being multiple choice. Let F = probability of a question not being multiple choice. We can use Pascal's triangle and the patterns we learned last unit to expand the binomial.

$$(S+F)^5 = S^5 + 5S^4 F + 10 S^3F^2 + 10S^2F^3 + 5SF^4 + F^5$$

Term	Rep the Prob that:
$S^5$	All 5 are MC
$5S^4F$	4 are MC and 1 is not
$10 S^3F^2$	3 are MC and 2 are not
$10S^2F^3$	2 are MC and 3 are not
$+5SF^4$	1 is MC and 4 are not
$F^5$	zero are MC and 5 are not

What is the probability that at least 3 of 5 questions generated for a short quiz are multiple choice?

We need to substitute in the probabilities and add together all the terms that represent 3, 4, or 5 questions being multiple choice.

$$(S+F)^5 = S^5 + 5S^4 F + 10 S^3 F^2 + 10S^2 F^3 + 5SF^4 + F^5$$

In this case,  $S=.7$ , and  $F = .3$

$$P(\text{at least 3 are MC}) = S^5 + 5S^4 F + 10 S^3 F^2$$

$$P(\text{at least 3 are MC}) = (.7)^5 + 5(.7)^4(.3) + 10 (.7)^3(.3)^2$$

$$P(\text{at least 3 are MC}) = .16807 + .36015 + .3087$$

$$P(\text{at least 3 are MC}) = .83692 \text{ or } \approx 84\%$$

## II. Using the Binomial Theorem:

- Remember that we can also use the Binomial Theorem to expand binomials.

$$(x + y)^n = \sum_{r=0}^n \frac{n!}{r!(n-r)!} x^{n-r} y^r$$

- Part of this should look like something we just learned:

$\frac{n!}{r!(n-r)!}$  is the formula for combinations  ${}_n C_r$

- To make this simpler to use, we can rewrite this as

$$(x + y)^n = \sum_{r=0}^n {}_n C_r x^{n-r} y^r$$

### III. Using Only One Term

- For binomial experiments, we are often looking for one term. Remember that when we are only looking for one term, we do not need the summation portion of the formula. We only need:

$${}_n C_r x^{n-r} y^r$$

Where,

x is the probability of success P(s)

y is the probability of failure P(f)

n is the number of trials

r is the number of failures (or think of it as the number subtracted from n to get the correct number of failures).

n-r is the number of successes

# Examples

- Ex 1: Eight out of every 10 people who contract a certain viral infection can recover. If a group of 9 people become infected, what is the probability that exactly 5 people will recover?
- There are 9 people involved and there are only two outcomes: recover or don't recover. The events are independent of each other. So this is a binomial experiment. We are looking for the term where 5 people recover (which means 4 people do NOT recover), so  $x^5y^4$ .  $x = \frac{8}{10}$  or  $\frac{4}{5}$  (*probability of recover*),  $y = \frac{2}{10}$  or  $\frac{1}{5}$  (*probability of not recovering*),  $n = 9$  (9 people, or trials),  $r = 4$  (the number of failures).

$${}^nC_r x^{n-r} y^r$$

$${}^9C_4 \left(\frac{4}{5}\right)^5 \left(\frac{1}{5}\right)^4$$

$$= .066 \text{ or } 6.6\%$$



- Ex 2: Find each probability if a number cube is tossed five times:

a. P(only one 4)

$${}_5C_4 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^4$$

$$\frac{3125}{7776}$$

b. P(at least three 4s)

P(three 4s) + P(four 4s) + P(five 4s)

$${}_5C_2 \left(\frac{1}{6}\right)^3 \left(\frac{5}{6}\right)^2 + {}_5C_1 \left(\frac{1}{6}\right)^4 \left(\frac{5}{6}\right)^1 + {}_5C_0 \left(\frac{1}{6}\right)^5 \left(\frac{5}{6}\right)^0$$

$$\frac{23}{648}$$

c. P(exactly five 4s)

$${}_5C_0 \left(\frac{1}{6}\right)^5 \left(\frac{5}{6}\right)^0$$

$$\frac{1}{7776}$$

b. P(no more than two 4s)

P(no 4s) + P(one 4) + P(two 4s)

$${}_5C_5 \left(\frac{1}{6}\right)^0 \left(\frac{5}{6}\right)^5 + {}_5C_4 \left(\frac{1}{6}\right)^1 \left(\frac{5}{6}\right)^4 + {}_5C_3 \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^3$$

$$\frac{625}{648}$$

# Notes: Sets and Venn Diagrams

## I. Sets

### A. Definition and Type

Set - collection of objects called members or elements

3 ways to describe a set:

- words:  $N$  is the set of all natural numbers

- list:  $N = \{1, 2, 3, \dots\}$

- set-builder notation:  $N = \{x \mid x \in N\}$

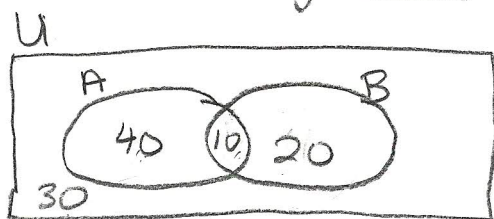
Types of sets:

- finite: a limited number of members

- infinite: unlimited number of members

- well-defined set: has a universe of objects which are allowed into consideration and any object in the universe is either an element of the set or it is not.

## II. Venn Diagrams



$U$ : universe or universal set: includes everything in the box

$A$ : Everything inside left circle.

$B$ : Everything inside right circle.

$U$ : Union - includes everything in both sets

$\cap$ : Intersection - includes the overlap of the sets

$A'$  or  $A^c$ : complement of  $A$  - everything that is NOT  $A$ .

Ex 1: Determine the number of elements in each.

$$U: 30 + 40 + 10 + 20 = \boxed{100}$$

$$A': 50$$

$$A: 40 + 10 = \boxed{50}$$

$$B': 70$$

$$B: 10 + 20 = \boxed{30}$$

$$(A \cup B)': 30$$

$$A \cup B: 40 + 10 + 20 = \boxed{70}$$

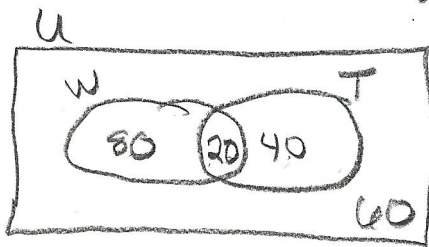
$$(A \cap B)': 90$$

$$A \cap B: \boxed{10}$$

$$U': 0$$

Ex 2: of 200 people interviewed, 100 people are Star Wars fans, 60 people are Star Trek fans, and 20 people are fans of both.

1. Draw a Venn Diagram to represent this information.



2. Determine the number of people in each of the following and describe in words.

U: 200 people interviewed.

W: 100 people who like SW.

$W \cup T$ : 140 people like SW or ST.

$W \cap T$ :

# **Notes Stats Day 1: Introduction to Probability Distributions**

# I. The Basics

## A. Random Variable

- A random variable  $x$  takes on a defined set of values with different probabilities.
  - For example, if you roll a die, the outcome is random (not fixed) and there are 6 possible outcomes, each of which occur with probability one-sixth.
  - For example, if you poll people about their voting preferences, the percentage of the sample that responds “Yes on Proposition 100” is also a random variable (the percentage will be slightly differently every time you poll).
- Roughly, probability is how frequently we expect different outcomes to occur if we repeat the experiment over and over (“frequentist” view) .

## B. Discrete Vs Continuous

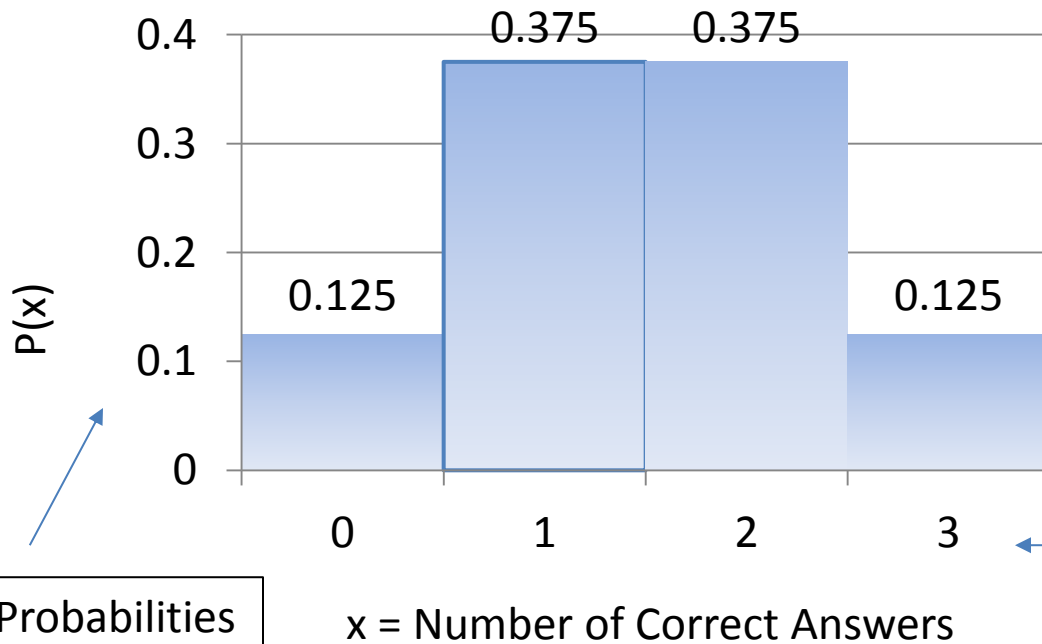
- **Discrete** random variables have a countable number of outcomes
  - Examples: Dead/alive, treatment/placebo, dice, counts, etc.
- **Continuous** random variables have an infinite continuum of possible values.
  - Examples: blood pressure, weight, the speed of a car, the real numbers from 1 to 6.

# C. Sample Space and Probability

- The Sample Space is where we list all possible numerical outcomes.
- For example, if we are talking about the sample space for rolling a die, it would be 1, 2, 3, 4, 5, 6. The probability of each outcome would be  $\frac{1}{6}$ .
- Or, think about a student guessing on a 3-question multiple choice test. The number of questions he could get correct would be the sample space (0, 1, 2, 3). Using the Binomial Theorem, we know that the probabilities are  $\frac{1}{8}$ ,  $\frac{3}{8}$ ,  $\frac{3}{8}$ , and  $\frac{1}{8}$  respectively.
- These are both discrete examples.

# Histograms

- We can represent a situation's sample space and probabilities with a histogram. Below is a histogram representing the probability distribution of randomly guessing on a 3-question multiple choice quiz.



What is the probability we get 2 or more questions correct?

$$0.375 + 0.125 = .5$$

Sample Space

Probabilities



## E. Expected Value

- The expected value is what it sounds like- what we expect to get. It is a type of average that is weighted for each probability. For a single discrete variable.
- To find the expected value, multiply each member of the sample space by its probability, and add all of this together. For the previous example of a 3-question quiz, the expected value is

$$0 \left(\frac{1}{8}\right) + 1 \left(\frac{3}{8}\right) + 2 \left(\frac{3}{8}\right) + 3 \left(\frac{1}{8}\right) \text{ or } 1.5.$$

Notice that this value is not always part of the sample space. Also notice that it is near the center of the histogram.

# Theoretical vs. Experimental Probability

- So far, we have been working almost exclusively with theoretical probability. Theoretical probability comes from our reasoning and calculation. We know that a die has 6 sides, so we reason that the probability of rolling a 5 is  $\frac{1}{6}$ .
- Experimental probability, on the other hand, comes from an actual experiment. It is the ratio of the number of times an event occurs to the total number of trials or times the activity is performed.

# Frequency Tables with Probability

- When doing an experiment, we often use frequency tables to organize and keep track of our results. We can add observed probabilities to our frequency tables. If we are doing an experiment, we use observed probabilities, not theoretical.

Outcome						
Observed Frequency						
Probability						

Depending on our experiment and our data, we may need to create class intervals. If a variable takes a large number of values, then it is easier to present and handle the data by grouping the values into class intervals. Continuous variables are more likely to be presented in class intervals, while discrete variables can be grouped into class intervals or not. In deciding on the width of the class intervals, you will have to find a compromise between having intervals short enough so that not all of the observations fall in the same interval, but long enough so that you do not end up with only one observation per interval.

Example: Lets say we surveyed 20 graduate students and asked them how much in student loans they owed. We got the following responses (in thousands): 0, 0, 13, 25, 18, 22, 23, 24, 24, 25, 28, 29, 30, 32, 33, 36, 39, 41, 47, 49.

First we need to decide on class intervals:

Outcome						
Observed Frequency						
Probability						

Example: Lets say we surveyed 20 graduate students and asked them how much in student loans they owed. We got the following responses (in thousands): 0, 0, 13, 18, 25, 22, 23, 24, 24, 25, 28, 29, 30, 32, 33, 36, 39, 41, 47, 49.

Outcome	0-9	10-19	20-29	30-39	40-49
Observed Frequency	2	2	8	5	3
Probability	2/20 Or .1	2/20 or .1	8/20 or .4	5/20 or .25	3/20 or .15

# Stats Day 2: Describing Distributions

## I. Mean, Median, Mode

The MEAN is the numerical average of the data set.

**The mean is found by adding all the values in the set, then dividing the sum by the number of values.**

The MEDIAN is the number that is in the middle of a set of data

- 1. Arrange the numbers in the set in order from least to greatest.**
- 2. Then find the number that is in the middle.**
- 3. If there is no number exactly in the middle, find the mean of the two middle numbers.**

Ex 1: These are Abby's science test scores. Find the mean and median.

86

97

84

73

63

88

97

100

95



Lets find Abby's  
**MEAN** science test  
score?

97

84

88

100

95

63

73

86

97

Add all  
the values.

Divide the sum  
by the number  
of values.

+

---

783

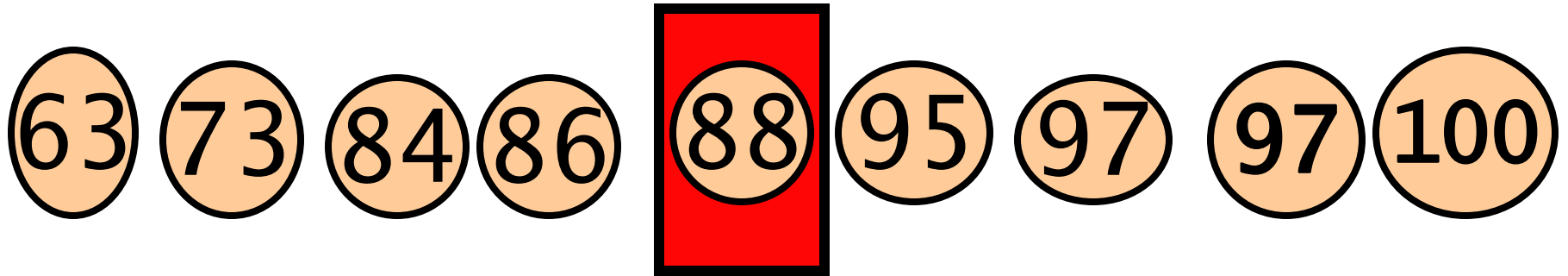
783

÷

9

The mean is **87**

Arrange values from  
least to greatest.



Find the number that is in the middle.

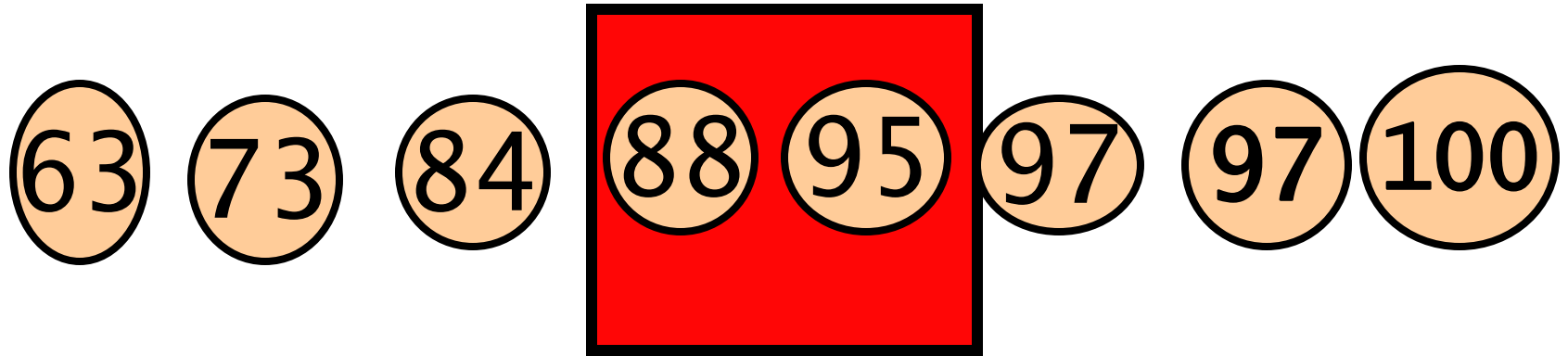
The median is **88**.

Half the numbers are  
less than the median.

Half the numbers are  
greater than the median.

Ex 2: Find the median.

Arrange values from least to greatest.



There are two numbers in the middle.

Add the 2 numbers.

$$88 + 95 = 183$$

Divide by 2.

$$183 \div 2$$

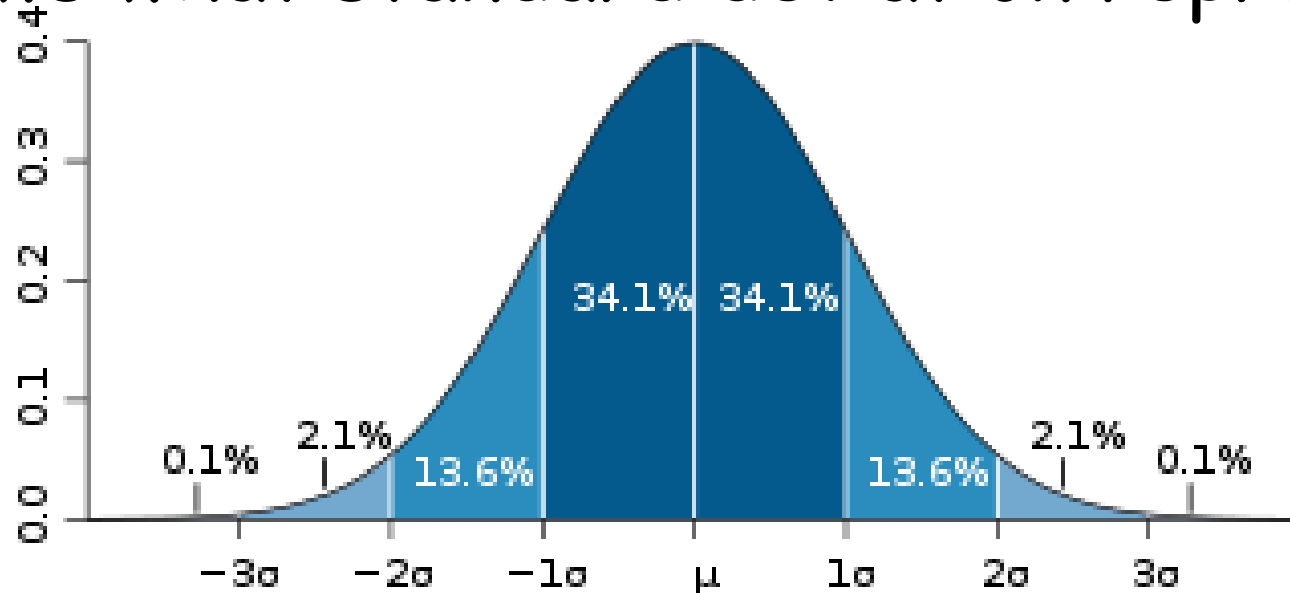
The median is  
**91.5**

# II. Standard Deviation

## A. Definition and Notation

**Standard Deviation** shows the variation in data. If the data is close together, the standard deviation will be small. If the data is spread out, the standard deviation will be large.

B. Bell Curve: The **bell curve**, which represents a normal distribution of data, shows what standard deviation represents.



One standard deviation away from the mean ( $\mu$ ) in either direction on the horizontal axis accounts for around **68 percent** of the data. Two standard deviations away from the mean accounts for roughly **95 percent** of the data with three standard deviations representing about **99 percent** of the data.

## C. Steps to Finding Standard Deviation

- 1) Find the **mean** of the data.
- 2) Subtract the mean from each value.
- 3) Square each deviation of the mean.
- 4) Find the sum of the squares.
- 5) Divide the total by the number of items, minus 1.
- 6) Take the square root.

# D. Standard Deviation Formula

The standard deviation formula can be represented using Sigma Notation:

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

The expression under the radical is called the 'variance'.

The standard deviation formula is the square root of the variance.

\*\*Note that you sometimes see the formula with only  $n$ , as opposed to  $n-1$ , in the denominator. Using only  $n$  is typical when working with populations. When working with other types of samples, it is typical to use  $n-1$ . This is also the formula that your graphing calculator will most likely use.

# Ex 1: Find the standard deviation

The math test scores of five students are: 92, 88, 80, 68 and 52.

1) Find the **mean**:  $(92+88+80+68+52)/5 = 76$ .

2) Find the **deviation from the mean**:

$$92-76=16$$

$$88-76=12$$

$$80-76=4$$

$$68-76=-8$$

$$52-76=-24$$



3) Square the deviation from the mean:

$$(16)^2 = 256$$

$$(12)^2 = 144$$

$$(4)^2 = 16$$

$$(-8)^2 = 64$$

$$(-24)^2 = 576$$

4) Find the sum of the squares of the deviation from the mean:

$$256+144+16+64+576= 1056$$

5) Divide by the number of data items, minus 1 ( $5-1 = 4$ ):

$$1056/4 = 264$$

6) Find the square root of the variance:  $\sqrt{264} \approx 16.35$

Thus the **standard deviation** of the test scores is **16.25**.

- You can also do this using your calculator:

- Press [STAT] [ENTER] to enter the statistics list editor
- Enter the numbers into the list by pressing each number and [ENTER]
- Press [STAT] [right cursor] to enter the CALC menu
- Select 1:1-Var Stats, 1-Var Stats will be typed on the home screen, press [ENTER]
- The statistics will be displayed (The cursor keys can be used to scroll through the entire list of results)

```
1-Var Stats
Σx=55.55555556
Mx=500
Σx²=80988
Sx=81.55536633
σx=76.89113677
↓n=9
```

$\bar{x}$  = mean

Sx = Standard Deviation for a sample (n-

$\sigma x$  = *Standard deviation for population (n)*

# III. Box Plot

## A. What is it?

- Also called a 'Box and Whiskers' plot
  - A 5-numbered summary of data:
    - Lower extreme
    - Lower quartile
    - Median
    - Upper quartile
    - Upper extreme
- To draw a Box Plot, we need to find all 5 of these numbers
- Each whisker, and each small box, contain 25% of the data.

## B. Steps to Creating a Box Plot

1. Order the numbers smallest to largest
2. Find the 5 numbers- median, lower and upper extremes, lower and upper quartiles
- 3 Draw the box plot- draw a number line, draw and label the parts

## C. Examples

### Example 1:

12, 13, 5, 8, 9, 20, 16, 14, 14, 6, 9, 12, 12

Step 1: Order the numbers from smallest to largest

5, 6, 8, 9, 9, 12, 12, 12, 13, 14, 14, 16, 20

## Step 2 – Find the Median

5, 6, 8, 9, 9, 12, 12, 12, 13, 14, 14, 16, 20



Median: 12

2. Find the **median**. The median is the middle number. If the data has two middle numbers, find the mean of the two numbers. What is the median?

## Step 2 – Find the Lower and Upper Extremes

5, 6, 8, 9, 9, 12, 12, 12, 13, 14, 14, 16, 20

5

Median: 12

20

2. Find the smallest and largest numbers



# Step 2 – Upper & Lower Quartiles

5, 6, 8, 9, 9, 12, 12, 12, 13, 14, 14, 16, 20

5

lower  
quartile:  
8.5

Median:  
12

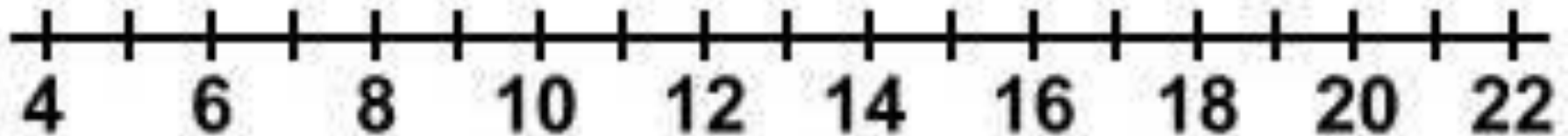
upper  
quartile:  
14

20

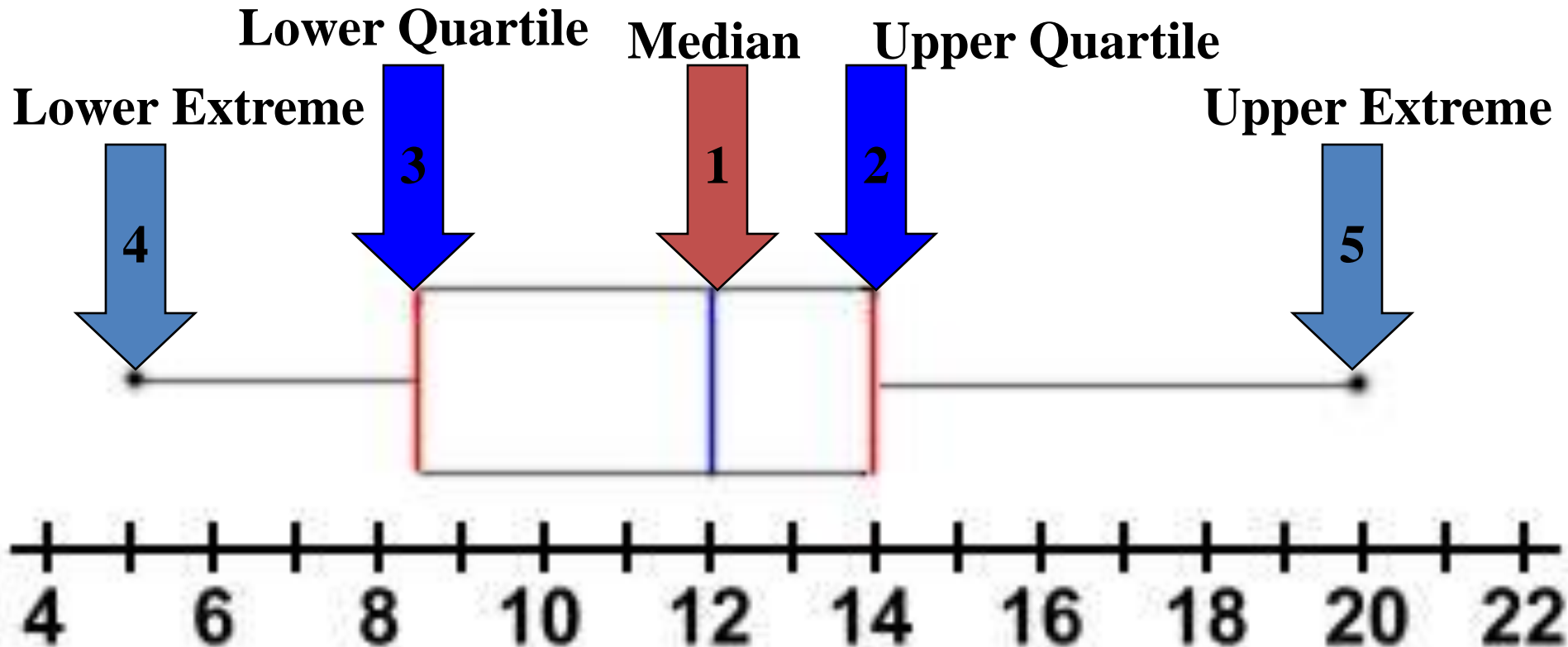
3. Find the **lower and upper medians or quartiles**. These are the middle numbers on each side of the median. What are they?

## Step 3 – Draw the Box Plot

Now you are ready to construct the actual box & whisker plot. First you will need to draw an ordinary number line that extends far enough in both directions to include all the numbers in your data:



# Name the parts of a Box-and-Whisker Plot



# IV. Range and Interquartile Range

The range is the difference between the highest and lowest numbers. In the previous example, the range was  $20 - 5$ , or 15.

The interquartile range is the difference between the upper quartile and the lower quartile.

$$14 - 8.5 = 5.5$$

# V. Outlier

## A. What is it?

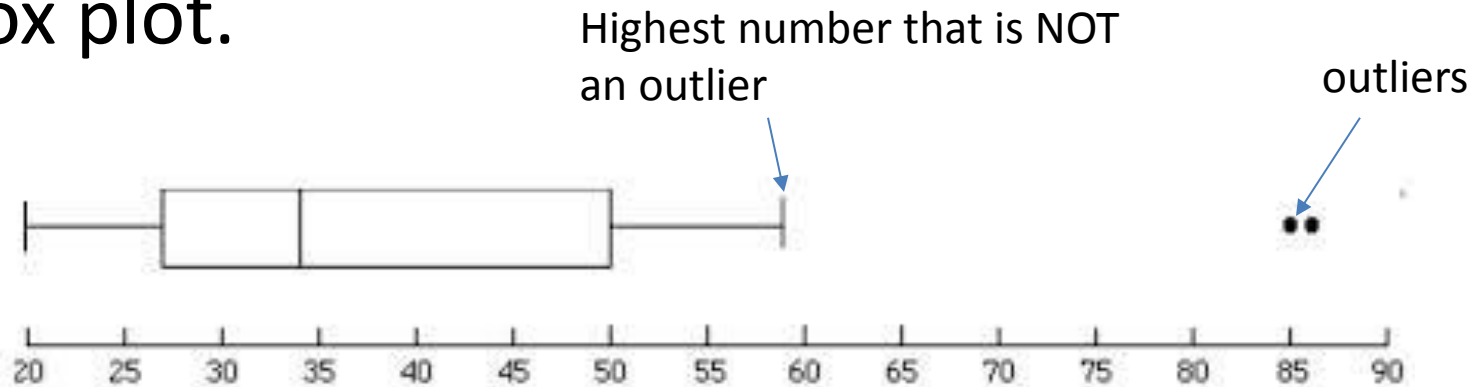
- An outlier is a data point that is distinctly separate from the rest of the data. One definition of outlier is any data point more than 1.5 interquartile ranges (IQRs) below the first quartile or above the third quartile.
- Note: The IQR definition given here is widely used but is not the last word in determining whether a given number is an outlier.

# Example of 1.5 IQR Rule

- For the data 2, 5, 6, 9, 12, we have the following five-number summary:
- minimum = 2  
first quartile = 3.5  
median = 6  
third quartile = 10.5  
maximum = 12
- IQR =  $10.5 - 3.5 = 7$ , so  $1.5 \cdot \text{IQR} = 10.5$ .
- To determine if there are outliers we must consider the numbers that are  $1.5 \cdot \text{IQR}$  or 10.5 beyond the quartiles.
- first quartile  $- 1.5 \cdot \text{IQR} = 3.5 - 10.5 = -7$   
third quartile  $+ 1.5 \cdot \text{IQR} = 10.5 + 10.5 = 21$
- Since none of the data are outside the interval from  $-7$  to  $21$ , there are no outliers.

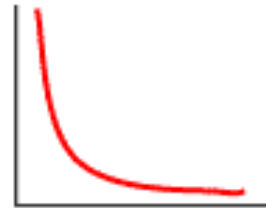
# Outliers on a Box Plot

- If there are outliers, you will need to adjust your box plot accordingly. Your whisker should end at the extreme values that are NOT outliers. Your outliers will be represented by dots that are not connected to the rest of the box plot.

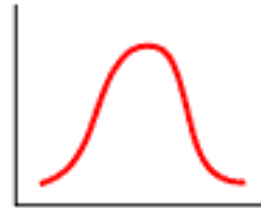


Number of Hysterectomies Performed by Male Doctors

# VI. Types of Distributions



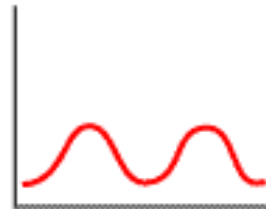
**J-shaped**



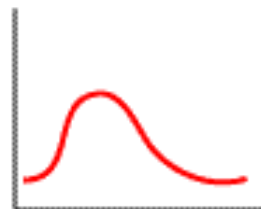
**Normal**



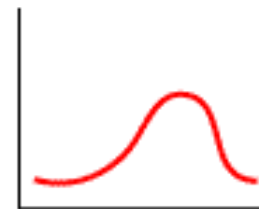
**Rectangular**



**Bimodal**



**Positive (right) skew**



**Negative (left) skew**

These are the most commonly used names, but you will also see other names depending on your source.



# Stats Day 3: Normal Distribution

## Objectives

The student will be able to:

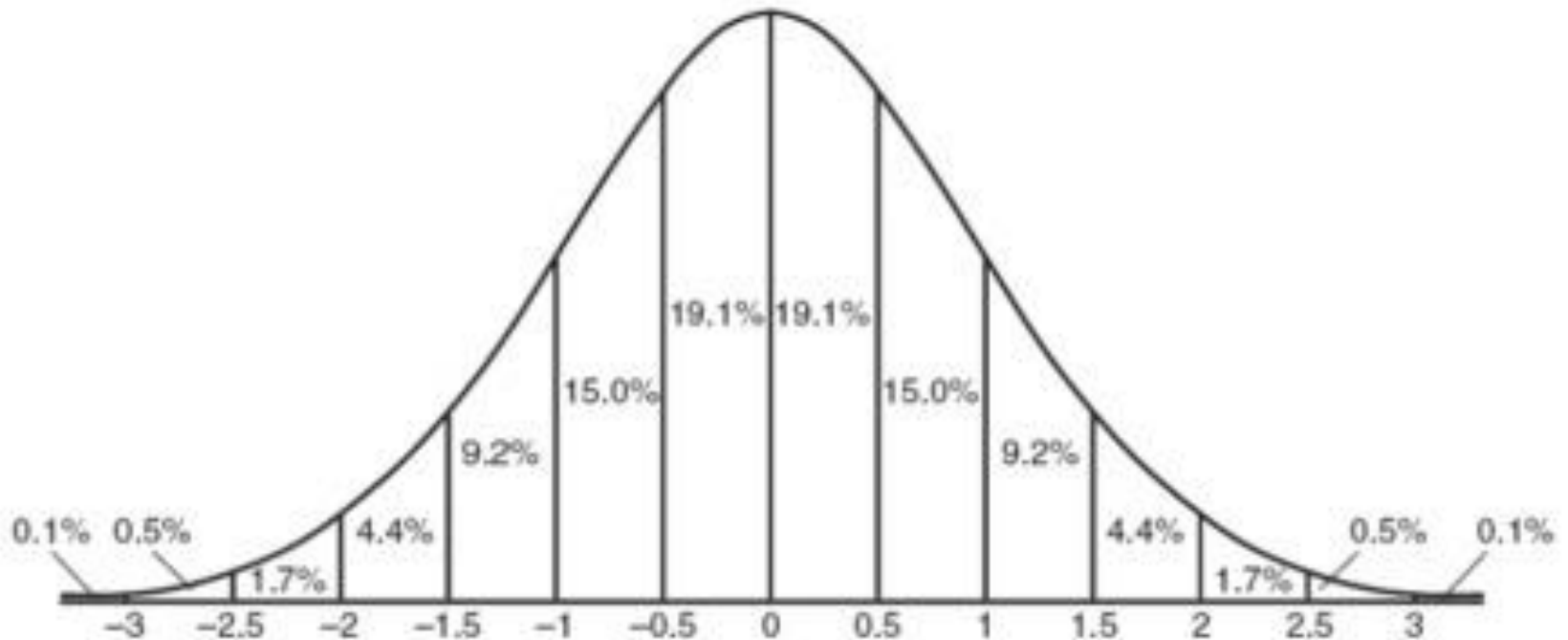
- ❖ identify properties of normal distribution
- ❖ apply mean, standard deviation, and  $z$ -scores to the normal distribution graph
- ❖ determine probabilities based on  $z$ -scores

# I. Normal Distribution Curve

A normal distribution curve is symmetrical, bell-shaped curve defined by the mean and standard deviation of a data set.

The normal curve is a probability distribution with a total area under the curve of 1.

## Normal Curve Standard Deviation



One standard deviation away from the mean ( $\mu$ ) in either direction on the horizontal axis accounts for around **68 percent** of the data. Two standard deviations away from the mean accounts for roughly **95 percent** of the data with three standard deviations representing about **99.7 percent** of the data. Copy this down. You will need to know these percentages.

The mean of the data in a standard normal distribution is 0 and the standard deviation is 1. A standard normal distribution is the set of all  $z$ -scores.

## *II.* $z$ -scores

When a set of data values are normally distributed, we can standardize each score by converting it into a  **$z$ -score**. This is the number of standard deviations the value is away from the mean.

**$z$ -scores** make it easier to compare data values measured on different scales.

The  **$z$ -score** is positive if the data value lies above the mean and negative if the data value lies below the mean.

## z-score formula

$$z = \frac{x - \mu}{\sigma}$$

Where  $x$  represents an element of the data set, the **mean** is represented by  $\mu$  and **standard deviation** by  $\sigma$ .

## Ex 1: Analyzing the data

Suppose SAT scores among college students are normally distributed with a mean of 500 and a standard deviation of 100. If a student scores a 700, what would be her  $z$ -score?

$$z = \frac{700 - 500}{100} = 2$$

Her  $z$ -score would be 2 which means her score is two standard deviations above the mean.

Ex 2: A set of math test scores has a mean of 70 and a standard deviation of 8. A set of English test scores has a mean of 74 and a standard deviation of 16. For which test would a score of 78 have a higher standing?

To solve: Find the  $z$ -score for each test.

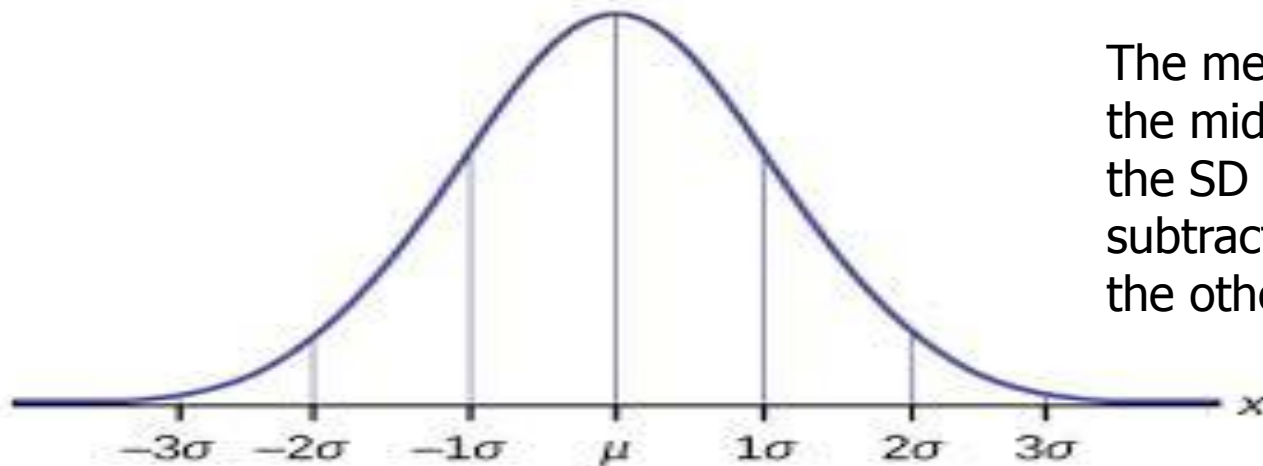
$$\text{math } z\text{-score} = \frac{78-70}{8} = 1 \qquad \text{English } z\text{-score} = \frac{76-74}{16} = .25$$

The math score would have the highest standing since it is 1 standard deviation above the mean while the English score is only .25 standard deviation above the mean.



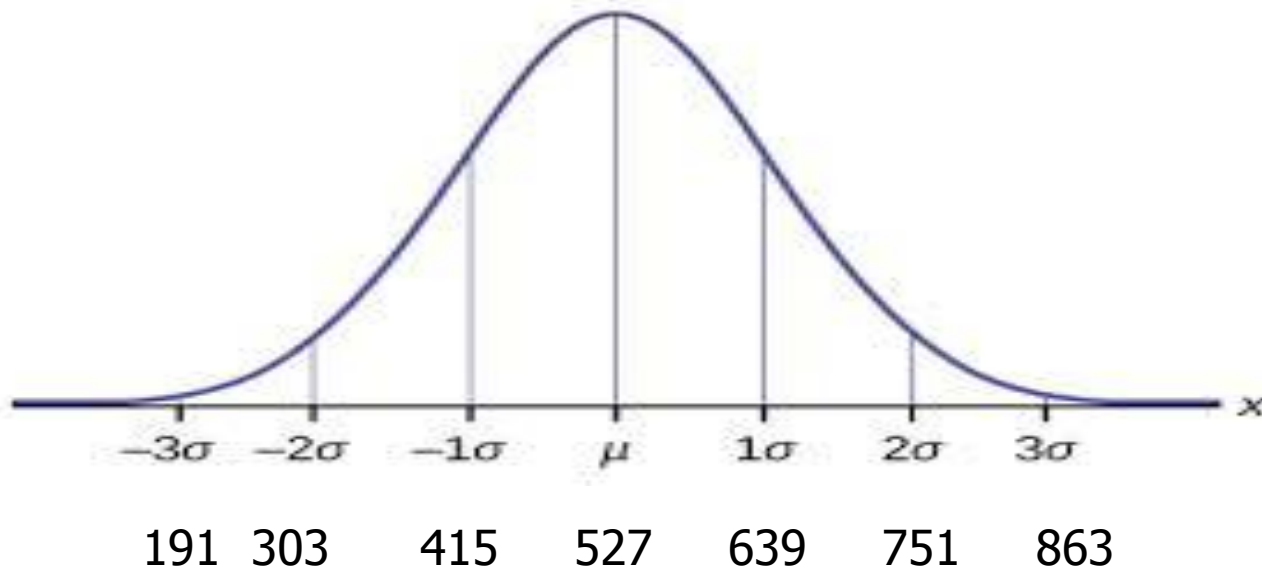
### III. Using the Normal Distribution Curve to Find Probabilities

- Most graduate schools of business require applicants for admission to take the Graduate Management Admission Council's GMAT examination. Scores on the GMAT are roughly normally distributed with a mean of 527 and a standard deviation of 112.
- a. Draw a Normal curve showing the mean and  $\pm 1$ ,  $\pm 2$ ,  $\pm 3$  standard deviations.



The mean goes in the middle, and the SD is added or subtracted to find the other numbers.

191 303 415 527 639 751 863



- B. What percentage of people get higher than an 527 on the GMAT?
- C. What percentage of people get between a 415 and a 639?
- D. What percentage of people get between a 359 and a 527?

Answers:

B. If we add up all the percentages on the right side of 527, they add up to 50%.

C. 415 is -1 standard deviation away from the mean, and 636 is 1 standard deviation away from the mean. The percentages between 415 and 639 add up to 68.1%.

D. 359 is -1.5 standard deviations away from the mean, and 527 is the mean. The total percentages add up to 43.3 percent.

## IV. Using Your Calculators to Graph Normal Distribution Probabilities

- [http://users.rowan.edu/~schultzl/TI/normal\\_distributions.pdf](http://users.rowan.edu/~schultzl/TI/normal_distributions.pdf)