#### **Sullivan Section 6.2**

Gene Quinn

A **binomial probability experiment** is an experiment that meets the following criteria:

The experiment consists of a *fixed* number n of repetitions called **trials**.

A **binomial probability experiment** is an experiment that meets the following criteria:

- The experiment consists of a fixed number n of repetitions called trials.
- Each trial has only two possible outcomes, which are mutually exclusive.
  - The two outcomes are usually referred to as **success** and **failure**.

A **binomial probability experiment** is an experiment that meets the following criteria:

- The experiment consists of a fixed number n of repetitions called trials.
- Each trial has only two possible outcomes, which are mutually exclusive.
  - The two outcomes are usually referred to as **success** and **failure**.
- Each of the trials is independent of the others.
  That is, the outcome of one trial has no effect on the other trials.

A **binomial probability experiment** is an experiment that meets the following criteria:

- The experiment consists of a fixed number n of repetitions called trials.
- Each trial has only two possible outcomes, which are mutually exclusive.
  - The two outcomes are usually referred to as **success** and **failure**.
- Each of the trials is independent of the others.
  That is, the outcome of one trial has no effect on the other trials.
- The probability of success is the same for each trial

When these conditions are met, the number of successes is denoted by X.

X is said to be a **random variable** having the **binomial distribution**.

When these conditions are met, the number of successes is denoted by X.

X is said to be a **random variable** having the **binomial distribution**.

If the probability of success is 1/2, the binomial experiment is equivalent to a series of coin tosses.

The following notation is used for the binomial probability distribution:

ullet The experiment consists of n independent trials.

The following notation is used for the binomial probability distribution:

- ullet The experiment consists of n independent trials.
- The probability of success on each trial is denoted by p.

The following notation is used for the binomial probability distribution:

- The experiment consists of n independent trials.
- ullet The probability of success on each trial is denoted by p.
- The probability of failure on each trial is 1-p.

The following notation is used for the binomial probability distribution:

- The experiment consists of n independent trials.
- ullet The probability of success on each trial is denoted by p.
- The probability of failure on each trial is 1-p.
- The total number of successes in n independent trials is denoted by X.

The following structure known as **Pascal's triangle** is useful for computing binomial probabilities when n is fairly small (n < 10).

								1							
n = 1							1		1						
n=2						1		2		1					
n = 3					1		3		3		1				
n=4				1		4		6		4		1			
n = 5			1		5		10		10		5		1		
n = 6		1		6		15		20		15		6		1	
n = 7	1		7		21		35		35		21		7		1

							1						
n = 1						1		1					
n=2					1		2		1				
n = 3				1		3		3		1			
n = 4			1		4		6		4		1		
n = 5		1		5		10		10		5		1	
n = 6	1		6		15		20		15		6		1

The entries in successive rows of Pascal's triangle are the sum of the two closest entries in the previous row.

If you think of the outcome of n trials with two outcomes, success or failure, the entire experiment can be summarized as a sequence of S's and F's with n entries.

If you think of the outcome of n trials with two outcomes, success or failure, the entire experiment can be summarized as a sequence of S's and F's with n entries.

In the row of Pascal's triangle corresponding to n trials, there are n+1 entries.

If you think of the outcome of n trials with two outcomes, success or failure, the entire experiment can be summarized as a sequence of S's and F's with n entries.

In the row of Pascal's triangle corresponding to n trials, there are n+1 entries.

The sum of the entries in the row corresponding to n trials is always  $2^n$ .

This represents the number of possible sequences of n letters where each one has to be either S or F.

There are always n+1 entries in the row corresponding to n trials.

• The first entry is the number of sequences with no S's and n F's (the first entry is always 1)

There are always n+1 entries in the row corresponding to n trials.

- The first entry is the number of sequences with no S's and n F's (the first entry is always 1)
- **●** The second entry is the number of sequences having 1 S and n-1 F's (the second entry is always n)

There are always n+1 entries in the row corresponding to n trials.

- The first entry is the number of sequences with no S's and n F's (the first entry is always 1)
- The second entry is the number of sequences having 1 S and n-1 F's (the second entry is always n)
- The third entry is the number of sequences having 2 S's and n-2 F's

There are always n+1 entries in the row corresponding to n trials.

- The first entry is the number of sequences with no S's and n F's (the first entry is always 1)
- The second entry is the number of sequences having 1 S and n-1 F's (the second entry is always n)
- The third entry is the number of sequences having 2 S's and n-2 F's
- The fourth entry is the number of sequences having 3 S's and n-3 F's

■ The next to last entry is the number of sequences having n-1 S's and 1 F (the next to last entry is always n)

- The next to last entry is the number of sequences having n-1 S's and 1 F (the next to last entry is always n)
- The last entry is the number of sequences having n S's and 0 F's (the last entry is always 1)

The simplest case occurs when success and failure are equally likely.

If we identify "heads" with "success", the experiment corrsponds tossing a fair coin n times.

The simplest case occurs when success and failure are equally likely.

If we identify "heads" with "success", the experiment corrsponds tossing a fair coin n times.

In this case, the probability of obtaining 0, 1, 2, etc. heads in n tosses is the corresponding entry in Pascal's table, divided by the sum of the row  $(2^n)$ .

When success and failure are *not* equally likely, we need to use the following modified procedure to calculate the probabilities.

The number of trials n determines which row of Pascal's triangle is used.

Suppose the probability of success on each trial is p.

We compute the probabilities associated with each value of X,

where X represents the number of successes in n trials.

• The first entry in the row is multiplied by  $p^n$ .

Suppose the probability of success on each trial is p.

We compute the probabilities associated with each value of X,

where X represents the number of successes in n trials.

- The first entry in the row is multiplied by  $p^n$ .
- The second entry in the row is multiplied by  $p^{n-1}(1-p)$ .

Suppose the probability of success on each trial is p.

We compute the probabilities associated with each value of X,

where X represents the number of successes in n trials.

- The first entry in the row is multiplied by  $p^n$ .
- The second entry in the row is multiplied by  $p^{n-1}(1-p)$ .
- The third entry in the row is multiplied by  $p^{n-2}(1-p)^2$ .

Suppose the probability of success on each trial is p.

We compute the probabilities associated with each value of X,

where X represents the number of successes in n trials.

- The first entry in the row is multiplied by  $p^n$ .
- The second entry in the row is multiplied by  $p^{n-1}(1-p)$ .
- The third entry in the row is multiplied by  $p^{n-2}(1-p)^2$ .
- Continue in this fashion. The  $n+1^{st}$  entry is multiplied by  $(1-p)^n$

An alternative method of computing binomial probabilities makes use of mathematical entities called **combinations**.

An alternative method of computing binomial probabilities makes use of mathematical entities called **combinations**.

- We define 1! to be 1.
- We define 2! to be  $2 \cdot 1$ .

An alternative method of computing binomial probabilities makes use of mathematical entities called **combinations**.

- We define 1! to be 1.
- We define 2! to be  $2 \cdot 1$ .
- We define 3! to be  $3 \cdot 2 \cdot 1$ .

An alternative method of computing binomial probabilities makes use of mathematical entities called **combinations**.

- We define 1! to be 1.
- $\bullet$  We define 2! to be  $2 \cdot 1$ .
- We define 3! to be  $3 \cdot 2 \cdot 1$ .
- We define 4! to be  $4 \cdot 3 \cdot 2 \cdot 1$ .

An alternative method of computing binomial probabilities makes use of mathematical entities called **combinations**.

- We define 1! to be 1.
- $\blacksquare$  We define 2! to be  $2 \cdot 1$ .
- We define 3! to be  $3 \cdot 2 \cdot 1$ .
- We define 4! to be  $4 \cdot 3 \cdot 2 \cdot 1$ .
- We define 5! to be  $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ .

An alternative method of computing binomial probabilities makes use of mathematical entities called **combinations**.

First, we need to define another mathematical entity called a **factorial**, which will be designated by a number followed by an exclamation point (!).

- We define 1! to be 1.
- $\blacksquare$  We define 2! to be  $2 \cdot 1$ .
- We define 3! to be  $3 \cdot 2 \cdot 1$ .
- We define 4! to be  $4 \cdot 3 \cdot 2 \cdot 1$ .
- We define 5! to be  $5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$ .

For convenience, we define 0! to be 1.

**Definition:** The number of **combinations** of n objects taken r at a time is denoted by either

$${}_{n}C_{r}$$
 or  ${n \choose r}$ 

and is defined to be:

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

**Example:** Find the number of **combinations** of 4 objects taken 2 at a time.

That is, find

$$_4C_2$$
 or  $\begin{pmatrix} 4 \\ 2 \end{pmatrix}$ 

**Example:** Find the number of **combinations** of 4 objects taken 2 at a time.

That is, find

$$_4C_2$$
 or  $\binom{4}{2}$ 

By definition,

$$\binom{4}{2} = \frac{4!}{2!(4-2)!} = \frac{4 \cdot 3 \cdot 2 \cdot 1}{(2 \cdot 1) \cdot (2 \cdot 1)} = \frac{24}{(2)(2)} = 6$$

The general formula for computing the probability of k successes in a binomial experiment with n trials when the probability of success on each trial is p is:

$$P(k \text{ successes}) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

The general formula for computing the probability of k successes in a binomial experiment with n trials when the probability of success on each trial is p is:

$$P(k \text{ successes}) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \dots, n$$

or, equivalently,

$$P(k \text{ successes}) =_n C_k p^k (1-p)^{n-k}, \quad k = 0, 1, ..., n$$

While it's good to know how these probabilities are computed, you don't have to do it by hand.

While it's good to know how these probabilities are computed, you don't have to do it by hand.

Spreadsheets have a function, usually called **BINOMDIST**, that computes them.

While it's good to know how these probabilities are computed, you don't have to do it by hand.

Spreadsheets have a function, usually called **BINOMDIST**, that computes them.

Most people find it better to use a spreadsheet, for convenience and accuracy

Keep in mind we are trying to compute the probability of a given number of successes in a binomial experiment.

Keep in mind we are trying to compute the probability of a given number of successes in a binomial experiment.

We have to know the number of trials in the experiment, n

Keep in mind we are trying to compute the probability of a given number of successes in a binomial experiment.

We have to know the number of trials in the experiment, n

We also need to know the probability of success on each trial,  $\boldsymbol{p}$ 

If *Cumulative* is FALSE, then BINOMDIST returns the probability of **exactly** x successes in n trials.

If *Cumulative* is FALSE, then BINOMDIST returns the probability of **exactly** x successes in n trials.

So, to find the probability that exactly 4 out of 10 trials are successes when p=0.6, the formula would be:

BINOMDIST(4,10,0.6,FALSE)

If *Cumulative* is FALSE, then BINOMDIST returns the probability of **exactly** x successes in n trials.

So, to find the probability that exactly 4 out of 10 trials are successes when p=0.6, the formula would be:

BINOMDIST(4,10,0.6,FALSE)

To find the probability that exactly 7 out of 10 trials are successes when p=0.6, the formula would be:

BINOMDIST(7,10,0.6,FALSE)

Sometimes we are interested in, say, the probability of 4 successes **or fewer** in 10 trials.

In this case, code Cumulative=TRUE in the BINOMDIST function.

Sometimes we are interested in, say, the probability of 4 successes **or fewer** in 10 trials.

In this case, code Cumulative=TRUE in the BINOMDIST function.

To find the probability that 4 **or fewer** out of 10 trials are successes when p = 0.6, the formula would be:

=BINOMDIST(4,10,0.6,TRUE)

Sometimes we are interested in, say, the probability of 4 successes **or fewer** in 10 trials.

In this case, code Cumulative=TRUE in the BINOMDIST function.

To find the probability that 4 **or fewer** out of 10 trials are successes when p=0.6, the formula would be:

```
=BINOMDIST(4,10,0.6,TRUE)
```

To find the probability that 7 or fewer out of 10 trials are successes when p = 0.6, the formula would be:

BINOMDIST(7,10,0.6,TRUE)

The compliment of the event "7 successes or fewer in 10 trials"

is the event "at least 8 successes in 10 trials"

The compliment of the event "7 successes or fewer in 10 trials"

is the event "at least 8 successes in 10 trials"

To find the probability that **at least** 8 out of 10 trials are successes when p=0.6, the formula would be:

```
=1-BINOMDIST(7,10,0.6,TRUE)
```

The compliment of the event "7 successes or fewer in 10 trials"

is the event "at least 8 successes in 10 trials"

To find the probability that **at least** 8 out of 10 trials are successes when p=0.6, the formula would be:

```
=1-BINOMDIST(7,10,0.6,TRUE)
```

To find the probability of at least 8 successes, we add the probabilities of  $1, 2, 3, \ldots, 7$  successes and subtract the total from 1.

In summary, for a binomial experiment with n trials and probability of success p, the probabilities of some common events are:

exactly $k$ successes	=BINOMDIST(k,n,p,FALSE)
k or fewer successes	=BINOMDIST(k,n,p,TRUE)
at least $k$ successes	=1-BINOMDIST(k-1,n,p,TRUE)
more than $k$ successes	=1-BINOMDIST(k,n,p,TRUE)
fewer than $k$ successes	=BINOMDIST(k-1,n,p,TRUE)
fewer than $j$ or	=1+BINOMDIST(j-1,n,p,TRUE)
more than $k$ successes	-BINOMDIST(k,n,p,TRUE)
Between $j$ and $k$	=BINOMDIST(k,n,p,TRUE)
successes (inclusive)	-BINOMDIST(j-1,n,p,TRUE)

Suppose the criteria for a binomial probability experiment are met.

The possible outcomes of the experiment, and the probabilities associated with each outcome are completely determined by two numbers:

The number of trials in the experiment, denoted by n

Suppose the criteria for a binomial probability experiment are met.

The possible outcomes of the experiment, and the probabilities associated with each outcome are completely determined by two numbers:

The number of trials in the experiment, denoted by n

The probability of success on each trial, denoted by p

**Example:** Suppose the criteria for a binomial probability experiment are met, and we are told that the experiment consists of 6 trials, each with probability of success equal to 0.6.

**Example:** Suppose the criteria for a binomial probability experiment are met, and we are told that the experiment consists of 6 trials, each with probability of success equal to 0.6.

Immediately, we know that the random variable X defined to be the number of successes obtained in the experiment must have one of the following values:

0, 1, 2, 3, 4, 5, 6

**Example:** Suppose the criteria for a binomial probability experiment are met, and we are told that the experiment consists of 6 trials, each with probability of success equal to 0.6.

Immediately, we know that the random variable X defined to be the number of successes obtained in the experiment must have one of the following values:

Furthermore, we know that for k = 0, 1, ..., 6, the probability that exactly k successes are obtained is given by the formula:

$$P(X=k) = {}_{6}C_k \cdot p^k (1-p)^{n-k}$$

In particular, we know that:

• The probability of 0 successes is  ${}_{6}C_{0} \cdot (0.6)^{0}(0.4)^{6}$ 

- The probability of 0 successes is  ${}_{6}C_{0}\cdot(0.6)^{0}(0.4)^{6}$
- The probability of 1 success is  ${}_{6}C_{1} \cdot (0.6)^{1}(0.4)^{5}$

- The probability of 0 successes is  ${}_{6}C_{0}\cdot(0.6)^{0}(0.4)^{6}$
- The probability of 1 success is  ${}_{6}C_{1} \cdot (0.6)^{1}(0.4)^{5}$
- The probability of 2 successes is  ${}_{6}C_{2} \cdot (0.6)^{2}(0.4)^{4}$

- The probability of 0 successes is  ${}_{6}C_{0} \cdot (0.6)^{0}(0.4)^{6}$
- The probability of 1 success is  ${}_{6}C_{1} \cdot (0.6)^{1}(0.4)^{5}$
- The probability of 2 successes is  ${}_{6}C_{2} \cdot (0.6)^{2}(0.4)^{4}$
- The probability of 3 successes is  ${}_{6}C_{3} \cdot (0.6)^{3}(0.4)^{3}$

- The probability of 0 successes is  ${}_{6}C_{0} \cdot (0.6)^{0}(0.4)^{6}$
- The probability of 1 success is  ${}_{6}C_{1} \cdot (0.6)^{1}(0.4)^{5}$
- The probability of 2 successes is  ${}_{6}C_{2} \cdot (0.6)^{2}(0.4)^{4}$
- The probability of 3 successes is  ${}_{6}C_{3} \cdot (0.6)^{3}(0.4)^{3}$
- The probability of 4 successes is  ${}_{6}C_{4} \cdot (0.6)^{4}(0.4)^{2}$

- The probability of 0 successes is  ${}_{6}C_{0} \cdot (0.6)^{0}(0.4)^{6}$
- The probability of 1 success is  ${}_{6}C_{1} \cdot (0.6)^{1}(0.4)^{5}$
- The probability of 2 successes is  ${}_{6}C_{2} \cdot (0.6)^{2}(0.4)^{4}$
- The probability of 3 successes is  ${}_{6}C_{3} \cdot (0.6)^{3}(0.4)^{3}$
- The probability of 4 successes is  ${}_{6}C_{4} \cdot (0.6)^{4}(0.4)^{2}$
- The probability of 5 successes is  ${}_{6}C_{5} \cdot (0.6)^{5}(0.4)^{1}$

- The probability of 0 successes is  ${}_{6}C_{0} \cdot (0.6)^{0}(0.4)^{6}$
- The probability of 1 success is  ${}_{6}C_{1} \cdot (0.6)^{1}(0.4)^{5}$
- The probability of 2 successes is  ${}_{6}C_{2} \cdot (0.6)^{2}(0.4)^{4}$
- The probability of 3 successes is  ${}_{6}C_{3} \cdot (0.6)^{3}(0.4)^{3}$
- The probability of 4 successes is  ${}_{6}C_{4}\cdot(0.6)^{4}(0.4)^{2}$
- The probability of 5 successes is  ${}_{6}C_{5} \cdot (0.6)^{5}(0.4)^{1}$
- The probability of 6 successes is  ${}_{6}C_{6} \cdot (0.6)^{6}(0.4)^{0}$

If we think of a large collection of binomial experiments producing a population of outcomes, the **population mean**  $\mu_X$  will be given by the formula:

$$\mu_X = np$$

If we think of a large collection of binomial experiments producing a population of outcomes, the **population mean**  $\mu_X$  will be given by the formula:

$$\mu_X = np$$

The **population standard deviation**  $\sigma_X$  is given by the formula:

$$\sigma_X = \sqrt{n \cdot p \cdot (1-p)}$$

**Example:** If X represents the number of successes in 100 trials in a binomial experiment with probability of success equal to 0.6, what is the mean  $\mu_X$  and standard deviation  $\sigma_X$  of X?

**Example:** If X represents the number of successes in 100 trials in a binomial experiment with probability of success equal to 0.6, what is the mean  $\mu_X$  and standard deviation  $\sigma_X$  of X?

The mean  $\mu_X$  is given by

$$\mu_X = n \cdot p = 100 \cdot 0.6 = 60$$

**Example:** If X represents the number of successes in 100 trials in a binomial experiment with probability of success equal to 0.6, what is the mean  $\mu_X$  and standard deviation  $\sigma_X$  of X?

The mean  $\mu_X$  is given by

$$\mu_X = n \cdot p = 100 \cdot 0.6 = 60$$

The standard deviation  $\sigma_X$  is given by

$$\sigma_X = \sqrt{n \cdot p \cdot (1-p)} = \sqrt{100 \cdot 0.6 \cdot 0.4} = 4.90$$

One of the properties of the binomial probability distribution is that the distribution is bell shaped when n is reasonably large.

One of the properties of the binomial probability distribution is that the distribution is bell shaped when n is reasonably large.

How large is a "reasonably large" value of n? It depends on p.

A commonly used rule of thumb states that the binomial distribution will be approximately bell shaped provided that

$$n \geq \frac{10}{p \cdot (1-p)}$$

Earlier we found that for a binomial experiment with 100 trials each having a probability of 0.6 of success, the mean and standard deviation were:

$$\mu_X = 60$$
 and  $\sigma_X = \sqrt{100 \cdot 0.6 \cdot 0.4} = 4.90$ 

Earlier we found that for a binomial experiment with 100 trials each having a probability of 0.6 of success, the mean and standard deviation were:

$$\mu_X = 60$$
 and  $\sigma_X = \sqrt{100 \cdot 0.6 \cdot 0.4} = 4.90$ 

- approximately 68% of the time X will fall in the range 55.1 to 64.9
- approximately 95% of the time X will fall in the range 50.2 to 69.8

Earlier we found that for a binomial experiment with 100 trials each having a probability of 0.6 of success, the mean and standard deviation were:

$$\mu_X = 60$$
 and  $\sigma_X = \sqrt{100 \cdot 0.6 \cdot 0.4} = 4.90$ 

- ullet approximately 68% of the time X will fall in the range 55.1 to 64.9
- approximately 95% of the time X will fall in the range 50.2 to 69.8
- ullet approximately 99.7% of the time X will fall in the range 45.3 to 74.7

Earlier we found that for a binomial experiment with 100 trials each having a probability of 0.6 of success, the mean and standard deviation were:

$$\mu_X = 60$$
 and  $\sigma_X = \sqrt{100 \cdot 0.6 \cdot 0.4} = 4.90$ 

- approximately 68% of the time X will fall in the range 55.1 to 64.9
- approximately 95% of the time X will fall in the range 50.2 to 69.8
- ullet approximately 99.7% of the time X will fall in the range 45.3 to 74.7

Earlier we found that for a binomial experiment with 100 trials each having a probability of 0.6 of success, the mean and standard deviation were:

$$\mu_X = 60$$
 and  $\sigma_X = \sqrt{100 \cdot 0.6 \cdot 0.4} = 4.90$ 

- approximately 68% of the time X will fall in the range 55.1 to 64.9
- approximately 95% of the time X will fall in the range 50.2 to 69.8
- ullet approximately 99.7% of the time X will fall in the range 45.3 to 74.7