5 Probability Distributions (Discrete Variables)

5.3 The Binomial Probability Distribution

The Binomial Probability Distribution

Consider the following probability experiment. Your instructor gives the class a surprise four-question multiple-choice quiz.

You have not studied the material, and therefore you decide to answer the four questions by randomly guessing the answers without reading the questions or the answers.

The Binomial Probability Distribution

Answer Page to Quiz

Directions: Circle the best answer to each question.

1. a b c
2. a b c
3. a b c
4. a b c

Circle your answers before continuing.

The Binomial Probability Distribution

Before we look at the correct answers to the quiz and find out how you did, let’s think about some of the things that might happen if you answered a quiz this way.

1. How many of the four questions are you likely to have answered correctly?

2. How likely are you to have more than half of the answers correct?

3. What is the probability that you selected the correct answers to all four questions?

The Binomial Probability Distribution

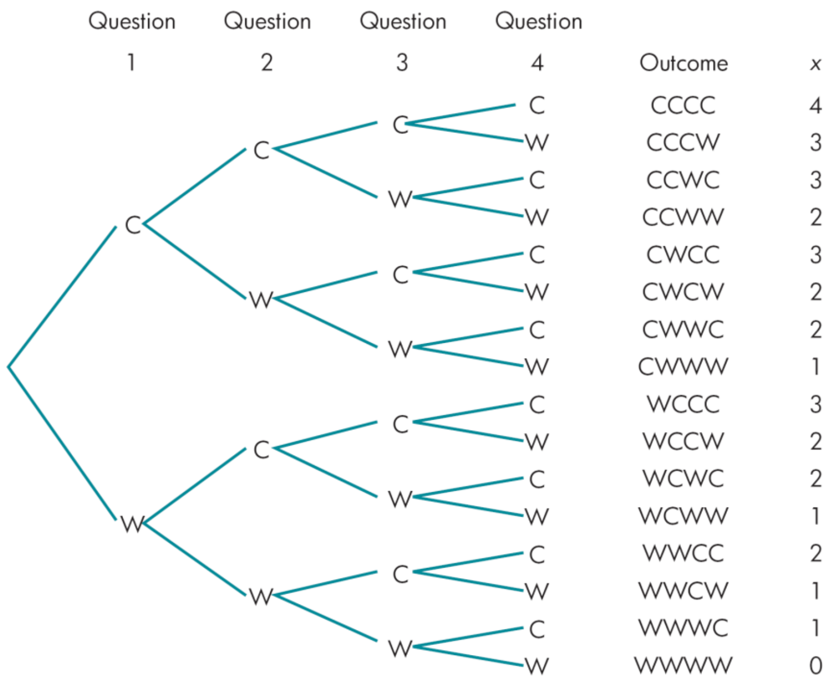
4. What is the probability that you selected wrong answers for all four questions?

5. If an entire class answers the quiz by guessing, what do you think the class “average” number of correct answers will be?

To find the answers to these questions, let’s start with a tree diagram of the sample space, showing all 16 possible ways to answer the four-question quiz.

The Binomial Probability Distribution

Each of the four questions is answered with the correct answer (C) or with a wrong answer (W). See Figure 5.4.



Tree Diagram: Possible Answers to a Four-Question Quiz

Figure 5.4

The Binomial Probability Distribution

We can convert the information on the tree diagram into a probability distribution.

Let *x* be the “number of correct answers” on one person’s quiz when the quiz was taken by randomly guessing.

The random variable *x* may take on any one of the values 0, 1, 2, 3, or 4 for each quiz. Figure 5.4 shows 16 branches representing five different values of *x*.

The Binomial Probability Distribution

Notice that the event $x = 4$, “four correct answers,” is represented by the top branch of the tree diagram, and the event $x = 0$, “zero correct answers,” is shown on the bottom branch.

The other events, “one correct answer,” “two correct answers,” and “three correct answers,” are each represented by several branches of the tree.

We find that the event $x = 1$ occurs on four different branches, event $x = 2$ occurs on six branches, and event $x = 3$ occurs on four branches.

The Binomial Probability Distribution

Each individual question has only one correct answer among the three possible answers, so the probability of selecting the correct answer to an individual question is $\frac{1}{3}$.

The probability that a wrong answer is selected on an individual question is $\frac{2}{3}$.

The probability of each value of *x* can be found by calculating the probabilities of all the branches and then combining the probabilities for branches that have the same *x* values.

The Binomial Probability Distribution

The calculations follow, and the resulting probability distribution appears in Table 5.7.

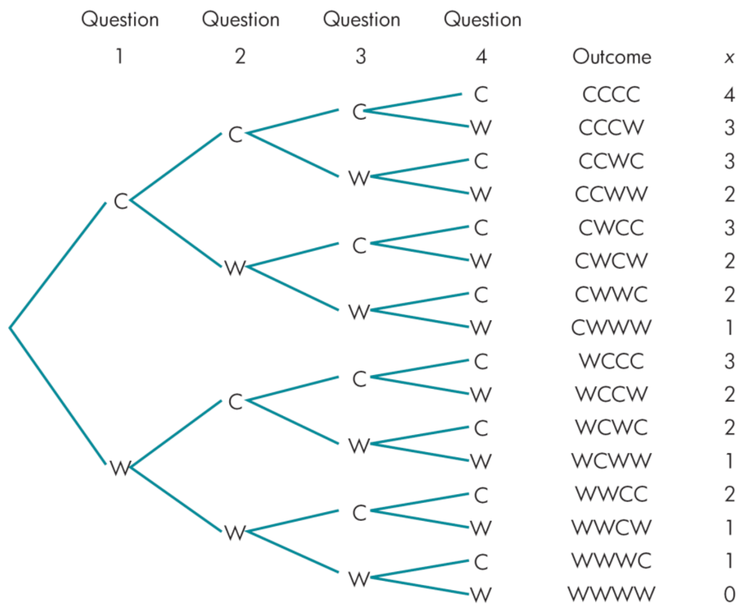
x	$P(x)$
0	0.198
1	0.395
2	0.296
3	0.099
4	0.012
	1.000

Probability Distribution for the Four-Question Quiz

Table 5.7

The Binomial Probability Distribution

$P(x = 0)$ is the probability that the correct answers are given for zero questions and the wrong answers are given for four questions (there is only one branch on Figure 5.4 where all four are wrong—WWWW):



Tree Diagram: Possible Answers to a Four-Question Quiz

Figure 5.4

The Binomial Probability Distribution

$$P(x = 0) = \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \left(\frac{2}{3}\right)^4 = \frac{16}{81} = \mathbf{0.198}$$

Note

Answering each individual question is a separate and independent event, thereby allowing us to use formula (4.7),

$$P(\text{A and B}) = P(\text{A}) \cdot P(\text{B}) \qquad \textbf{(4.7)}$$

which states that we should multiply the probabilities.

The Binomial Probability Distribution

$P(x = 1)$ is the probability that the correct answer is given for exactly one question and wrong answers are given for the other three (there are four branches on Figure 5.4 where this occurs—namely, CWWW, WCWW, WWCW, WWWC—and each has the same probability):

$$P(x = 1) = (4) \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = (4)\left(\frac{1}{3}\right)^1\left(\frac{2}{3}\right)^3 = \mathbf{0.395}$$

The Binomial Probability Distribution

$P(x = 2)$ is the probability that correct answers are given for exactly two questions and wrong answers are given for the other two (there are six branches on Figure 5.4 where this occurs—CCWW, CWCW, CWWC, WCCW, WCWC, WWCC—and each has the same probability):

$$P(x = 2) = (6) \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} = (6)\left(\frac{1}{3}\right)^2\left(\frac{2}{3}\right)^2 = \mathbf{0.296}$$

The Binomial Probability Distribution

$P(x = 3)$ is the probability that correct answers are given for exactly three questions and a wrong answer is given for the other question (there are four branches on Figure 5.4 where this occurs—CCCW,CCWC,CWCC,WCCC—and each has the same probability):

$$P(x = 3) = (4) \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = (4)\left(\frac{1}{3}\right)^3\left(\frac{2}{3}\right)^1 = \mathbf{0.099}$$

The Binomial Probability Distribution

P(*x* = 4) is the probability that correct answers are given for all four questions (there is only one branch on Figure 5.4 where all four are correct—CCCC):

$$P(x = 4) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \left(\frac{1}{3}\right)^4 = \frac{1}{81} = \mathbf{0.012}$$

The Binomial Probability Distribution

Now we can answer the five questions that were asked about the four-question quiz.

Answer 1: The most likely occurrence would be to get one answer correct; it has a probability of 0.395. Zero, one, or two correct answers are expected to result approximately 89% of the time (0.198 + 0.395 + 0.296 = 0.889).

Answer 2: Having more than half correct is represented by $x = 3$ or 4; their total probability is 0.099 + 0.012 = 0.111. (You will pass this quiz only 11% of the time by random guessing.)

The Binomial Probability Distribution

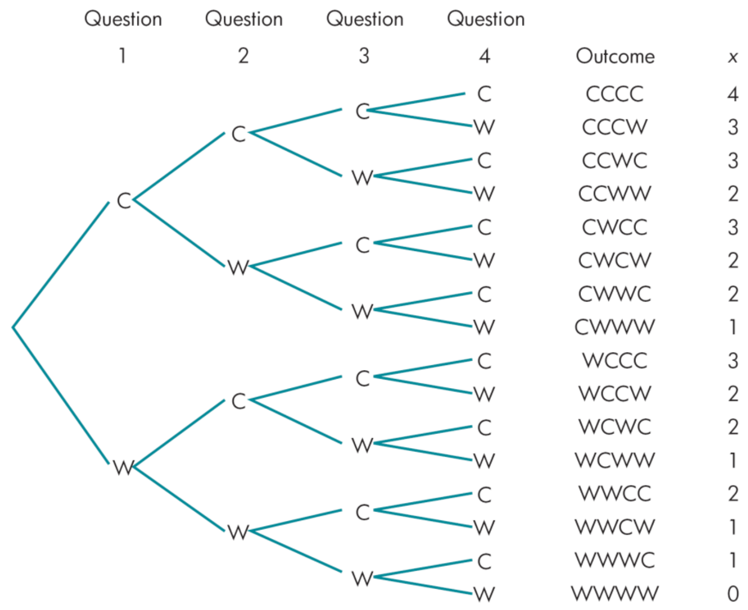
Answer 3: P(all four correct) = $P(x = 4) = 0.012$. (All correct occurs only 1% of the time.)

Answer 4: P(all four wrong) = $P(x = 0) = 0.198$. (That's almost 20% of the time.)

Answer 5: The class average is expected to be $\frac{1}{3}$ of 4, or 1.33 correct answers.

The Binomial Probability Distribution

The correct answers to the quiz are b, c, b, a. How many correct answers did you have? Which branch of the tree in Figure 5.4 represents your quiz results?



Tree Diagram: Possible Answers to a Four-Question Quiz

Figure 5.4

The Binomial Probability Distribution

You might ask several people to answer this same quiz by guessing the answers. Then construct an observed relative frequency distribution and compare it with the distribution shown in Table 5.7.

x	$P(x)$
0	0.198
1	0.395
2	0.296
3	0.099
4	0.012
	1.000

Probability Distribution for the Four-Question Quiz

Table 5.7

The Binomial Probability Distribution

Many experiments are composed of repeated trials whose outcomes can be classified into one of two categories: **success** or **failure**.

Examples of such experiments are coin tosses, right/wrong quiz answers, and other, more practical experiments such as determining whether a product did or did not do its prescribed job and whether a candidate gets elected or not.

The Binomial Probability Distribution

There are experiments in which the trials have many outcomes that, under the right conditions, may fit this general description of being classified in one of two categories. For example, when we roll a single die, we usually consider six possible outcomes.

However, if we are interested only in knowing whether a "one" shows or not, there are really only two outcomes: the "one" shows or "something else" shows.

The experiments just described are called *binomial probability experiments*.

The Binomial Probability Distribution

Binomial probability experiment An experiment that is made up of repeated trials that possess the following properties:

1. There are n repeated identical independent trials.

2. Each trial has two possible outcomes (success or failure).

3. P(success) = p, P(failure) = q, and $p + q = 1$.

4. The **binomial random variable** x is the count of the number of successful trials that occur; x may take on any integer value from zero to n.

The Binomial Probability Distribution

Notes

1. Properties 1 and 2 describe the two basic characteristics of any binomial experiment.

2. **Independent trials** mean that the result of one trial does not affect the probability of success on any other trial in the experiment. In other words, the probability of success remains constant throughout the entire experiment.

3. Property 3 gives the algebraic notation for each trial.

The Binomial Probability Distribution

4. Property 4 concerns the algebraic notation for the complete experiment.

5. It is of utmost importance that both *x* and *p* be associated with "success."

The four-question quiz qualifies as a binomial experiment made up of four trials when all four of the answers are obtained by random guessing.

Property 1: A **trial** is the **answering of one question**, and it is repeated ***n*** **= 4** times. The trials are **independent** because the probability of a correct answer on any one question is not affected by the answers on other questions.

The Binomial Probability Distribution

Property 2: The two possible outcomes on each trial are **success = C**, correct answer, and **failure = W**, wrong answer.

Property 3: For each trial (each question): ***p* = *P*(*correct*) =** $\frac{1}{3}$ and ***q* = *P*(*incorrect*) =** $\frac{2}{3}$. $[p + q = 1]$

Property 4: For the total experiment (the quiz): ***x* = number of correct answers** and can be any integer value from zero to $n = 4$.

The Binomial Probability Distribution

The key to working with any probability experiment is its probability distribution. All binomial probability experiments have the same properties, and therefore the same organization scheme can be used to represent all of them.

The *binomial probability function* allows us to find the probability for each possible value of x.

The Binomial Probability Distribution

Binomial probability function For a binomial experiment, let p represent the probability of a "success" and q represent the probability of a "failure" on a single trial. Then $P(x)$, the probability that there will be exactly x successes in n trials, is

$$P(x) = \binom{n}{x}(p^x)(q^{n-x}) \qquad \text{for } x = 0, 1, 2, \ldots, n \qquad \textbf{(5.5)}$$

The Binomial Probability Distribution

When you look at the probability function, you notice that it is the product of three basic factors:

1. The number of ways that exactly x successes can occur in n trials, $\binom{n}{x}$

2. The probability of exactly x successes, p^x

3. The probability that failure will occur on the remaining (n – x) trials, q^{n-x}

The Binomial Probability Distribution

The number of ways that exactly *x* successes can occur in a set of *n* trials is represented by the symbol $\binom{n}{x}$, which must always be a positive integer.

This term is called the **binomial coefficient** and is found by using the formula

$$\binom{n}{x} = \frac{n!}{x!(n-x)!} \qquad \textbf{(5.6)}$$

The Binomial Probability Distribution

Notes

1. *n*! (*"n factorial"*) is an abbreviation for the product of the sequence of integers starting with *n* and ending with one. For example, 3! = 3 • 2 • 1 = 6 and 5! = 5 • 4 • 3 • 2 • 1 = 120. There is one special case, 0!, that is defined to be 1. For more information about **factorial notation**, see the *Student Solutions Manual*.

2. The values for *n*! and $\binom{n}{x}$ can be readily found using most scientific calculators.

The Binomial Probability Distribution

3. The binomial coefficient $\binom{n}{x}$ is equivalent to the number of combinations ${}_nC_x$, the symbol most likely on your calculator.

4. See the *Student Solutions Manual* for general information on the binomial coefficient.

The Binomial Probability Distribution

A coin is tossed three times and we observe the number of heads that occur in the three tosses. This is a binomial experiment because it displays all the properties of a binomial experiment:

1. There are n = **3** repeated **independent** trials (each coin toss is a separate trial, and the outcome of any one trial has no effect on the probability of another trial).

2. Each trial (each toss of the coin) results in one of two possible outcomes: success = **heads** (what we are counting) or failuare = **tails**.

The Binomial Probability Distribution

3. The probability of success is $p = P(\text{H}) = \mathbf{0.5}$, and the probability of failure is
   $q = P(\text{T}) = \mathbf{0.5}$ [$p + q = 0.5 + 0.5 = 1$]

4. The random variable x is the **number of heads** that occur in the three trials. x will assume exactly one of the values **0, 1, 2, or 3** when the experiment is complete.

The Binomial Probability Distribution

The binomial probability function for the tossing of three coins is

$$P(x) = \binom{n}{x}(p^x)\ (q^{n-x}) = \binom{3}{x}(0.5)^x(0.5)^{3-x} \quad \text{for} \quad x = 0, 1, 2, 3$$

Let's find the probability *x* = 1 of using the preceding binomial probability function:

$$P(x = 1) = \binom{3}{1}(0.5)^1(0.5)^2 = 3(0.5)(0.25) = \mathbf{0.375}$$

Example 9 – *Binomial Probability of “Bad Eggs”*

The manager of Steve’s Food Market guarantees that none of his cartons of a dozen eggs will contain more than one bad egg.

If a carton contains more than one bad egg, he will replace the whole dozen and allow the customer to keep the original eggs.

If the probability that an individual egg is bad is 0.05, what is the probability that the manager will have to replace a given carton of eggs?

Example 9 – *Solution*

At first glance, the manager's situation appears to fit the properties of a binomial experiment if we let x be the number of bad eggs found in a carton of a dozen eggs, let $p = P(bad) = 0.05$, and let the inspection of each egg be a trial that results in finding a "bad" or "not bad" egg.

There will be $n = 12$ trials to account for the 12 eggs in a carton.

Example 9 – *Solution*

cont'd

However, trials of a binomial experiment must be independent; therefore, we will assume that the quality of one egg in a carton is independent of the quality of any of the other eggs. (This may be a big assumption! But with this assumption, we will be able to use the binomial probability distribution as our model.)

Now, based on this assumption, we will be able to find/estimate the probability that the manager will have to make good on his guarantee.

Example 9 – *Solution*

cont'd

The probability function associated with this experiment will be:

$$P(x) = \binom{12}{x}(0.05)^x(0.95)^{12-x} \qquad \text{for } x = 0, 1, 2, \ldots, 12$$

The probability that the manager will replace a dozen eggs is the probability that $x = 2, 3, 4, \ldots, 12$.

We know that $\Sigma P(x) = 1$; that is,

$$\boldsymbol{P(0)} + \boldsymbol{P(1)} + P(2) + \ldots + P(12) = 1$$

Example 9 – *Solution*

cont'd

$P(\text{replacement}) = P(2) + P(3) + \ldots + P(12)$

$= 1 - [\boldsymbol{P(0)} + \boldsymbol{P(1)}]$

It is easier to find the probability of replacement by finding $P(x = 0)$ and $P(x = 1)$ and subtracting their total from 1 than by finding all of the other probabilities. We have

$$P(x) = \binom{12}{x}(0.05)^x(0.95)^{12-x}$$

$$P(0) = \binom{12}{0}(0.05)^0(0.95)^{12} = \mathbf{0.540}$$

Example 9 – *Solution*

cont'd

$$P(1) = \binom{12}{1}(0.05)^1(0.95)^{11}$$

$$= \mathbf{0.341}$$

$$P(\text{replacement}) = 1 - (0.540 + 0.341)$$

$$= \mathbf{0.119}$$

If p = 0.05 is correct, then the manager will be busy replacing cartons of eggs.

If he replaces 11.9% of all the cartons of eggs he sells, he certainly will be giving away a substantial proportion of his eggs. This suggests that he should adjust his guarantee (or market better eggs).

Example 9 – *Solution*

cont'd

For example, if he were to replace a carton of eggs only when four or more were found to be bad, he would expect to replace only 3 out of 1000 cartons
[1.0 – (0.540 + 0.341 + 0.099 + 0.017)], or 0.3% of the cartons sold.

Notice that the manager will be able to control his "risk" (probability of replacement) if he adjusts the value of the random variable stated in his guarantee.

The Binomial Probability Distribution

Note
The value of many binomial probabilities for values of $n \leq 15$ and common values of p are found in Table 2 of Appendix B. In this example, we have $n = 12$ and $p = 0.05$, and we want the probabilities for $x = 0$ and 1.

We need to locate the section of Table 2 where $n = 12$, find the column headed $p = 0.05$, and read the numbers across from $x = 0$ and $x = 1$.

The Binomial Probability Distribution

We find .540 and .341, as shown in Table 5.8. (Look up these values in Table 2 in Appendix B.)

		p													
n	*x*	0.01	0.05	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	0.95	0.99	*x*
	⋮														
12	0	.886	.540	.282	.069	.014	.002	0+	0+	0+	0+	0+	0+	0+	0
	1	.107	.341	.377	.206	.071	.017	.003	0+	0+	0+	0+	0+	0+	1
	2	.006	.099	.230	.283	.168	.064	.016	.002	0+	0+	0+	0+	0+	2
	3	0+	.017	.085	.236	.240	.142	.054	.012	.001	0+	0+	0+	0+	3
	4	0+	.002	.021	.133	.231	.213	.121	.042	.008	.001	0+	0+	0+	4
	⋮														

Excerpt of Table 2 in Appendix B, Binomial Probabilities

Table 5.8

The Binomial Probability Distribution

Note

A convenient notation to identify the binomial probability distribution for a binomial experiment with $n = 12$ and $p = 0.05$ is $B(12, 0.05)$. $B(12, 0.05)$, read *"binomial distribution for $n = 12$ and $p = 0.05$,"* represents the entire distribution or "block" of probabilities shown in purple in Table 5.8.

When used in combination with the $P(x)$ notation, $P(x = 1|B(12, 0.05))$ indicates the probability of $x = 1$ from this distribution, or 0.341, as shown on Table 5.8.

Mean and Standard Deviation of the Binomial Distribution

Mean and Standard Deviation of the Binomial Distribution

The mean and standard deviation of a theoretical binomial probability distribution can be found by using these two formulas:

Mean of Binomial Distribution

$$\mu = np \qquad \textbf{(5.7)}$$

and

Standard Deviation of Binomial Distribution

$$\sigma = \sqrt{npq} \qquad \textbf{(5.8)}$$

Mean and Standard Deviation of the Binomial Distribution

The formula for the mean, μ, seems appropriate: the number of trials multiplied by the probability of "success." [We know that the mean number of correct answers on the binomial quiz was expected to be $\frac{1}{3}$ of 4, $4(\frac{1}{3})$, or *np*.]

The formula for the standard deviation, σ, is not as easily understood.

Mean and Standard Deviation of the Binomial Distribution

A coin is tossed three times. Let the “number of heads” that occur in those three tosses be the random variable, *x*. *x* is the number of heads in three coin tosses, $n = 3$, and $p = \frac{1}{2} = 0.5$. Using formula (5.7), we find the mean of *x* to be

$$\mu = np = (3)(0.5) = 1.5$$

Using formula (5.8), we find the standard deviation of *x* to be

$$\sigma = \sqrt{npq} = \sqrt{(3)(0.5)(0.5)} = \sqrt{0.75} = 0.866 = 0.87$$

Example 11 – *Calculating the Mean and Standard Deviation of a Binomial Distribution*

Find the mean and standard deviation of the binomial distribution when n = 20 and p = $\frac{1}{5}$ (or 0.2, in decimal form).

We know that the "binomial distribution where and " has the probability function

$$P(x) = \binom{20}{x}(0.2)^{x}(0.8)^{20-x} \qquad \text{for } x = 0, 1, 2, \ldots , 20$$

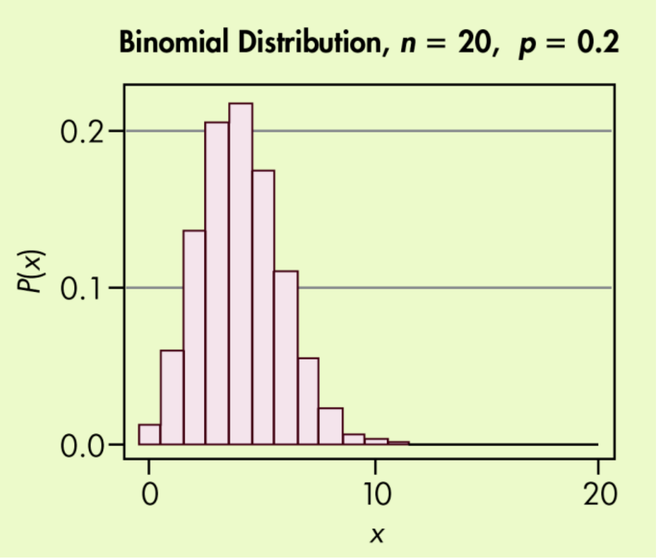
and a corresponding distribution with 21 *x* values and 21 probabilities.

As shown in the distribution chart, Table 5.9, and on the histogram in Figure 5.5.

x	*P*(*x*)
0	0.012
1	0.058
2	0.137
3	0.205
4	0.218
5	0.175
6	0.109
7	0.055
8	0.022
9	0.007
10	0.002
11	0+
12	0+
13	0+
•	•
:	:
20	0+

Binomial Distribution: n = 20, p = 0.2

Table 5.9



Histogram of Binomial Distribution *B*(20, 0.2)

Figure 5.5

Example 11 – *Calculating the Mean and Standard Deviation of a Binomial Distribution*

cont'd

Let's find the mean and the standard deviation of this distribution of *x* using formulas (5.7) and (5.8):

$$\mu = np$$

$$= (20)(0.2)$$

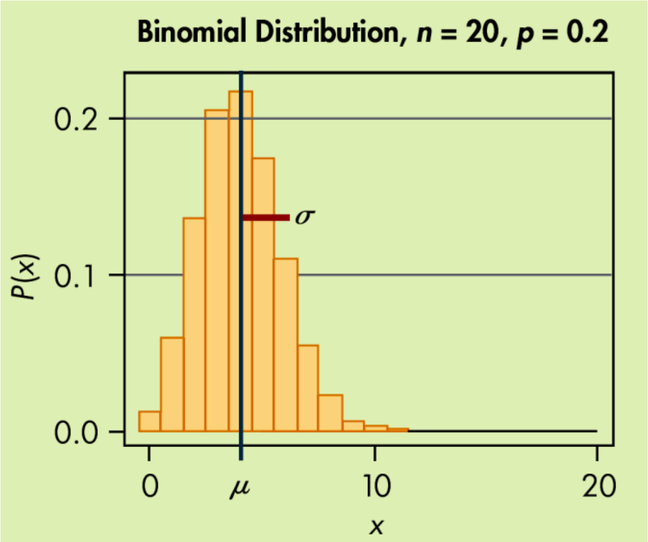
$$= \mathbf{4.0}$$

$$\sigma = \sqrt{npq}$$

$$= \sqrt{(20)(0.2)(0.8)}$$

$$= \sqrt{3.2}$$

$$= \mathbf{1.79}$$



Histogram of Binomial Distribution *B*(20, 0.2)

Figure 5.6

Figure 5.6 shows the mean, $\mu = 4$ (shown by the location of the vertical blue line along the *x*-axis), relative to the variable *x*. This 4.0 is the mean value expected for *x*, the number of successes in each random sample of size 20 drawn from a population with $p = 0.2$.

Figure 5.6 also shows the size of the standard deviation, $\sigma = 1.79$ (as shown by the length of the horizontal red line segment).

It is the expected standard deviation for the values of the random variable *x* that occur in samples of size 20 drawn from this same population.