MATH 1700

Instructor: Mehdi Maadooliat

Chapter 5



Department of Mathematical and Statistical Sciences

CHAPTER 5



Random Variables

- Discrete
- Continuous
- Probability function
- Probability distribution
- Mean
- Variance
- Standard deviation
- Binomial
 - Probability experiment
 - Probability function
 - Mean
 - Variance

USA AND ITS AUTOMOBILES AND RANDOM VARIABLE



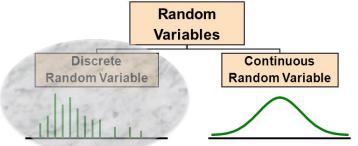
• The national average is 2.28 vehicles per household, with nearly 34% being single-vehicle and 31% being two-vehicle households. However, nearly 35% of all households have three or more vehicles.

Vehicles,x	1	2	3	4	5	6	7	8
P(x)	0.34	0.31	0.22	0.06	0.03	0.02	0.01	0.01

- Random variable A variable that assumes a unique numerical value for each of the outcomes in the sample space of a probability experiment.
 - -X = # of vehicles in a household is a random variable
- Another example: Let Y = the number of heads when we flip a coin twice.
 - Sample Space: $S = \{TT, TH, HT, HH\}$
 - **Random Variable:** $Y = \{0, 1, 2\}$

RANDOM VARIABLES - EXAMPLES





- Discrete random variable A quantitative random variable that can assume a countable number of values. Examples:
 - We toss five coins and observe the "number of heads" visible. The random variable x is the number of heads observed and may take on integer values from 0 to 5.
 - Let the "number of phone calls received" per day by a company be the random variable. Integer values ranging from zero to some very large number are possible values.
- Continuous random variable A quantitative random variable that can assume an uncountable (continuum of values) number of values. Example:
 - Let the "length of the cord" on an electrical appliance be a random variable. The random variable is a numerical value between 12 and 72 inches for most appliances.

PROBABILITY DISTRIBUTIONS OF A DISCRETE RANDOM VARIABLE



- Consider a coin-tossing experiment where two coins are tossed and no heads, one head, or two heads are observed.
- If we define the random variable x to be the number of heads observed when two coins are tossed, then

-
$$P(x = 0) = P(\text{no H}) = P(TT) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 0.25$$

- $P(x = 1) = P(\text{one H}) = P(HT \text{ or } TH) = \frac{1}{2} \times \frac{1}{2} + \frac{1}{2} \times \frac{1}{2} = \frac{1}{2} = 0.50$
- $P(x = 2) = P(\text{two H}) = P(HH) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4} = 0.25$

• Probability distribution: A distribution of the probabilities associated with each of the values of a random variable. The probability distribution is a theoretical distribution; it is used to represent populations. R

X	P(x)
0	0.25
1	0.50
2	0.25

Probability Distribution: Tossing Two Coins

PROBABILITY DISTRIBUTIONS OF A DISCRETE RANDOM VARIABLE



Another example: Rolling a die

X	1	2	3	4	5	6
P(x)	<u>1</u>	<u>1</u>	1/6	<u>1</u>	<u>1</u>	1/6

- Probability function A rule, P(x), that assigns probabilities to the values of the random variables.
 - The probability function for the experiment of rolling a die is

•
$$P(x) = \frac{1}{6}$$
, for $x = 1, 2, 3, 4, 5, 6$

•
$$P(x) = \frac{2!}{x!(2-x)!} \left(\frac{1}{2}\right)^x \left(\frac{1}{2}\right)^{2-x}$$
, for $x = 0,1,2$

0 0.25
1 0.50
2 0.25

- Properties:
 - **Property 1:** $0 \le P(x) \le 1$
 - Property 2: $\sum_{\text{all } x} P(x) = 1$

EXAMPLE 2 - DETERMINING A PROBABILITY FUNCTION

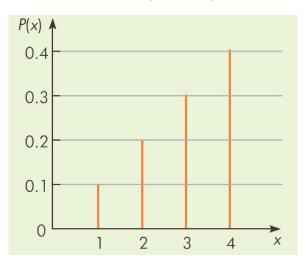


• Example: Is $P(x) = \frac{x}{10}$, for x = 1, 2, 3, 4 a probability

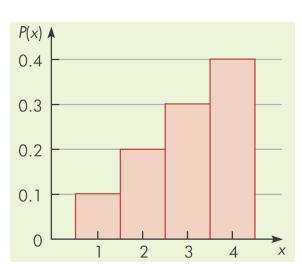
function?

- Solution:
 - **Property 1:** 0 ≤ P(x) ≤ 1
 - Property 2: $\sum_{\text{all } x} P(x) = 1$
- Question: What is P(x = 5)?

$$- P(x = 5) = 0$$



Line Representation: Probability Distribution for $P(x) = \frac{x}{10}$, for x = 1, 2, 3, 4



Histogram: Probability Distribution for $P(x) = \frac{x}{10}$, for x = 1, 2, 3, 4

x	P(x)
1	$\frac{1}{10} = 0.1 \checkmark$
2	$\frac{2}{10} = 0.2 \checkmark$
3	$\frac{3}{10} = 0.3 \checkmark$
4	$\frac{4}{10} = 0.4 \checkmark$
	$\frac{10}{10} = 1.0$

Probability Distribution for $P(x) = \frac{x}{10}$, for x = 1, 2, 3, 4



- Remember that:
 - We used **sample statistics** to describe a sample.
 - \bar{x} is the sample mean
 - s^2 is the sample variance (s is the sample standard deviation).
- Probability distributions may be used to represent theoretical populations, the counterpart to samples.
- We use population parameters (mean, variance, and standard deviation) to describe these probability distributions.
 - $-\mu$ (the Greek letter lower case mu) is the population mean.
 - σ^2 (the Greek letter lower case sigma) is the population variance
 - $\sigma = \sqrt{\sigma^2}$ is the population standard deviation.
 - $-\mu$, σ^2 and σ are called *population parameters*. (A parameter is a constant, and typically unknown value in real statistics problems.)



- Mean of a discrete random variable (expected value): The mean, μ , of a discrete random variable x is found by multiplying each possible value of x by its own probability and then adding all the products together:
 - **mean of x:** mu = sum of (each x multiplied by its own probability)

$$- \mu = \sum_{i=1}^{n} [x_i P(x_i)]$$

$$- \mu = x_1 P(x_1) + x_2 P(x_2) + \dots + x_n P(x_n)$$

• For the # of H when we flip a coin twice discrete distribution:

$$- \mu = x_1 P(x_1) + x_2 P(x_2) + x_3 P(x_3)$$

$$- = 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4}$$

$$- = 1$$

$$x$$
 $P(x)$
 $x_1 \longrightarrow 0$ 0.25 $P(x_1)$
 $x_2 \longrightarrow 1$ 0.50 $P(x_2)$
 $x_3 \longrightarrow 2$ 0.25 $P(x_3)$



Variance of a discrete random variable:

The variance, σ^2 , of a discrete random variable x is found by multiplying each possible value of the squared deviation from the mean, $(x - \mu)^2$, by its own probability and then adding all the products together:

- variance: sigma squared = sum of (squared deviation times probability)
- $\sigma^2 = \sum_{i=1}^n [(x_i \mu)^2 P(x_i)]$

$$- \sigma^2 = (x_1 - \mu)^2 P(x_1) + (x_2 - \mu)^2 P(x_2) + \dots + (x_n - \mu)^2 P(x_n)$$

• For the # of H when we flip a coin twice discrete distribution:

$$- \sigma^{2} = (x_{1} - \mu)^{2} P(x_{1}) + (x_{2} - \mu)^{2} P(x_{2}) + (x_{3} - \mu)^{2} P(x_{3})$$

$$- (0 - 1)^{2} \times \frac{1}{4} + (1 - 1)^{2} \times \frac{1}{2} + (2 - 1)^{2} \times \frac{1}{4}$$

$$- = \frac{1}{2} = 0.5$$

$$x_{1} \longrightarrow 0 \quad 0.25 \longrightarrow P(x_{1})$$

$$x_{2} \longrightarrow 1 \quad 0.50 \longrightarrow P(x_{2})$$

$$x_{3} \longrightarrow 2 \quad 0.25 \longrightarrow P(x_{3})$$

Alternative formula: $\sigma^2 = \sum_{i=1}^n [x_i^2 P(x_i)] - \mu^2$



- Likewise, standard deviation of a random variable is calculated in the same manner as is the standard deviation of sample data.
- Standard deviation of a discrete random variable: The positive square root of variance.
 - standard deviation: $\sigma = \sqrt{\sigma^2}$
- For the # of H when we flip a coin twice discrete distribution:

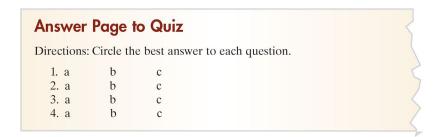
$$-\sigma = \sqrt{0.5}$$

$$- = 0.707$$

$$- = 0.71$$



- Consider the following probability experiment. I give you 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.
- Here are some questions for you?

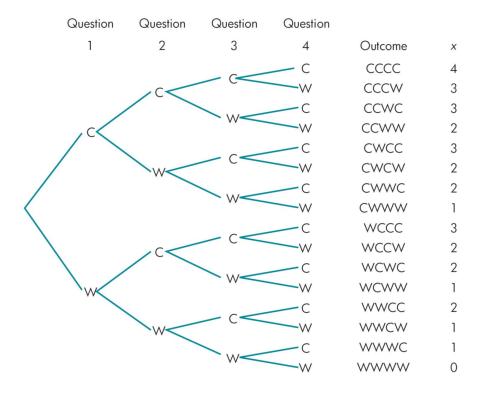


- 1. How many of the four questions are you likely to have answered correctly?
- 2. How likely are you to have more than half of the answers correct?
- 3. What is the probability that you selected the correct answers to all four questions?
- 4. What is the probability that you selected wrong answers 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

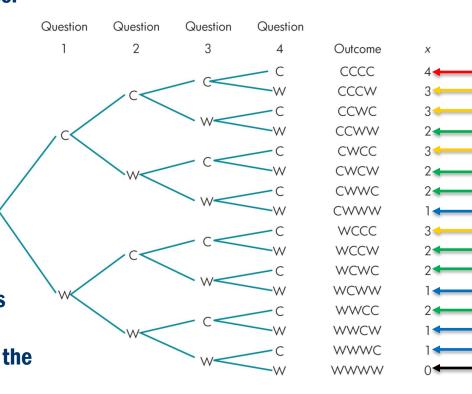


- Each of the four questions is answered with the correct answer (C) or with a wrong answer (W).
- x is the "number of correct answers" on one person's quiz when the quiz was taken by randomly guessing.



Notice that:

- The event x = 4, "four correct answers," is shown on the top branch.
- The event x = 0, "zero correct answers," is shown on the bottom branch.
- The event x = 1 occurs on four different branches.
- The event x = 2 occurs on six branches.
- The event x = 3 occurs on four branches.
- Each individual question has only one correct answer.
- The probability of selecting the correct answer to each question is $\frac{1}{3}$.
- The probability that a wrong answer is selected 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.



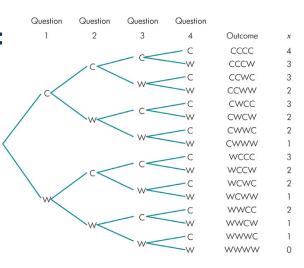


• P(x = 0) is the probability that the correct answers are given for zero questions.

-
$$P(x = 0) = \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \left(\frac{2}{3}\right)^4 = \frac{16}{81} = \mathbf{0}.\mathbf{198}$$

- Note: Answering each individual question is a separate and independent event, thereby we can use:
- $P(A \text{ and } B) = P(A) \cdot P(B)$
- P(x = 4) is the probability that correct answers are given for all four questions.

-
$$P(x = 4) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \left(\frac{1}{3}\right)^4 = \frac{1}{81} = \mathbf{0}.012$$





• 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: CWWW, WCWW, WWCW, WWWC—and each has the same probability):

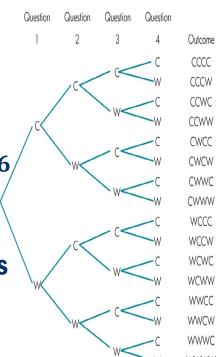
-
$$P(x = 1) = 4 \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = 4 \times \frac{1}{3} \times \left(\frac{2}{3}\right)^3 = \mathbf{0}.395$$

• 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):

-
$$P(x = 2) = 6 \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} = 6 \times \left(\frac{1}{3}\right)^2 \times \left(\frac{2}{3}\right)^2 = \mathbf{0}.296$$

• P(x=3) is the probability that correct answers are given for exactly three questions and wrong answers are given for the other one(there are four branches):

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



THE BINOMIAL

PROBABILITY DISTRIBUTION



•
$$P(x = 0) = \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = \left(\frac{2}{3}\right)^4 = \frac{16}{81} = \mathbf{0}.\mathbf{198}$$

•
$$P(x = 1) = 4 \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} \times \frac{2}{3} = 4 \times \frac{1}{3} \times \left(\frac{2}{3}\right)^3 = \mathbf{0}.395$$

•
$$P(x = 2) = 6 \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} \times \frac{2}{3} = 6 \times \left(\frac{1}{3}\right)^2 \times \left(\frac{2}{3}\right)^2 = \mathbf{0}.296$$

•
$$P(x = 3) = 4 \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{2}{3} = 4 \times \left(\frac{1}{3}\right)^3 \times \frac{2}{3} = \mathbf{0}.099$$

•
$$P(x = 4) = \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} \times \frac{1}{3} = \left(\frac{1}{3}\right)^4 = \frac{1}{81} = \mathbf{0}.\mathbf{012}$$

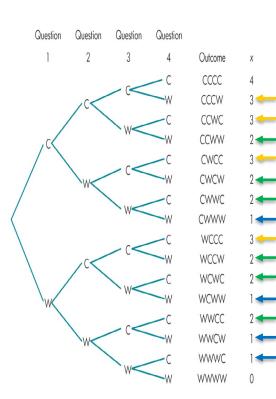
In general:

•
$$P(x = k) = \frac{4!}{k!(4-k)!} \times \left(\frac{1}{3}\right)^k \times \left(\frac{2}{3}\right)^{4-k}$$
, for $k = 0,1,2,3,4$

• Probability distribution:

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

Probability Distribution for the Four-Question Quiz





- Now we can answer those five questions.
- 1. How many of the four questions are you likely to have answered correctly?
- The most likely occurrence would be to get one answer correct; it has a probability of 0.395.

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

Probability Distribution for the Four-Question Quiz

- 2. How likely are you to have more than half of the answers correct?
- 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 with 11% chance by random guess.)
- 3. What is the probability that you selected the correct answers to all four questions?
- ightharpoonup P(all four correct) = P(x=4) = 0.012. (All correct occurs only 1% of the time.)
- 4. What is the probability that you selected wrong answers for all four questions?
- ightharpoonup P(all four wrong) = P(x = 0) = 0.198. (That's almost 20% of the time.)
- 5. If an entire class answers the quiz by guessing, what do you think the class "average" number of correct answers will be?
- \blacktriangleright The class average is expected to be $\frac{1}{3}$ of 4, or 1.33 correct answers.



- 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.
- 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.



- 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.
- 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
- $P(x) = \binom{n}{x} (p^x) (q^{n-x})$, for $x = 0, 1, 2, \dots, n$



•
$$P(x) = \binom{n}{x} (p^x)(q^{n-x})$$
, for $x = 0, 1, 2, \dots, n$

- 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, $\binom{n}{x}$
 - This term is called the **binomial coefficient** and is found by using the formula:

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

- 2. The probability of exactly x successes, p^x
- 3. The probability of failure on the remaining (n-x) trials, q^{n-x}



- 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:
 - 1. There are n = 3 repeated independent trials.
 - 2. Each trial (each toss of the coin) results in one of two possible outcomes: success = **heads** or failure = **tails**.
 - 3. The probability of success is p = P(H) = 0.5, and the probability of failure is q = P(T) = 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.
- The binomial probability function for the tossing of three coins is:

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

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

-
$$P(x = 1) = {3 \choose 1} (0.5)^1 (0.5)^{3-1} = 3(0.5)(0.25) = \mathbf{0}.375$$

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?

Solution:

- The manager's situation appears to fit the properties of a binomial experiment. Let x be the number of bad eggs found in a carton of a dozen eggs, therefore p = P(bad) = 0.05, and q = P(not bad) = 0.95.
- There will be n=12 trials to account for the 12 eggs in a carton.

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

EXAMPLE 9 - SOLUTION



Solution Cont'd:

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



- The probability that the manager will replace a dozen eggs is the probability that x = 2, 3, 4, ..., 12.
- We know that $\sum_{\text{all } x} P(x) = 1$; that is,

$$- P(0) + P(1) + P(2) + \dots + P(12) = 1$$

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

$$- = 1 - [P(0) + P(1)]$$

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

-
$$P(1) = {12 \choose 1} (0.05)^1 (0.95)^{11} = \mathbf{0.341}$$

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

$$= 0.119 = 11.9\%$$



• Note: The value of many binomial probabilities for values of $n \le 15$ 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.

								р							
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
	÷		V												
12	0	.886	.540	.282	.069	.014	.002	\bigcirc +	$\bigcirc +$	\bigcirc +	$\bigcirc +$	\bigcirc +	$\bigcirc +$	$\bigcirc +$	0
	1	.107	.341	.377	.206	.071	.017	.003	0+	0+	$\bigcirc +$	$\bigcirc +$	$\bigcirc +$	$\bigcirc +$	1
	2	.006	.099	.230	.283	.168	.064	.016	.002	0+	$\bigcirc +$	$\bigcirc +$	$\bigcirc +$	$\bigcirc +$	2
	3	0+	.017	.085	.236	.240	.142	.054	.012	.001	$\bigcirc +$	$\bigcirc +$	$\bigcirc +$	$\bigcirc +$	3
	4	0+	.002	.021	.133	.231	.213	.121	.042	.008	.001	\bigcirc +	$\bigcirc +$	$\bigcirc +$	4
	÷														

Excerpt of Table 2 in Appendix B, Binomial Probabilities

- TI-83 Calculator:
 - Binomialpdf(n, p, x) = P(x)
- Shiny App:
 - Binomial Calculator

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

•
$$\mu = \sum_{x=0}^{n} [xP(x)] = \sum_{i=1}^{n} [x {n \choose x} (p^x) (q^{n-x})]$$

$$\mu = np$$

Variance and Standard Deviation of Binomial Distribution

•
$$\sigma^2 = \sum_{x=0}^n [(x-\mu)^2 P(x)] = \sum_{i=1}^n \left[(x-\mu)^2 \binom{n}{x} (p^x) (q^{n-x}) \right]$$

$$\sigma^2 = npq$$

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

EXAMPLE 11 - CALCULATING THE MEAN MARQUET AND STANDARD DEVIATION OF A BINOMIAL Be The Difference. 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

•
$$P(x) = {20 \choose x} (0.2)^x (0.8)^{20-x}$$
 for $x = 0, 1, 2, ..., 20$

• and a corresponding distribution with $21\ x$ values and $21\ \text{probabilities}.$

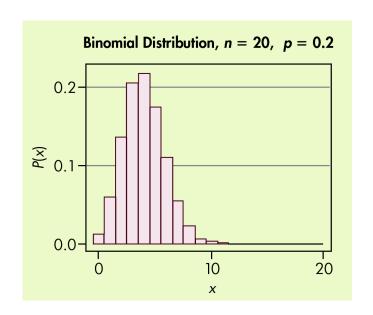
EXAMPLE 11 - CALCULATING THE MEAN MARQUET AND STANDARD DEVIATION OF A BINOMIAL Be The Difference. DISTRIBUTION

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

х	P(x)
0 1 2 3 4 5 6 7 8 9 10 11 12 13	0.012 0.058 0.137 0.205 0.218 0.175 0.109 0.055 0.022 0.007 0.002 0+ 0+ 0+ 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 MARQUET AND STANDARD DEVIATION OF A BINOMIAL Be The Difference. DISTRIBUTION

• Let's find the mean and the standard deviation of this distribution of x:

•
$$\mu = np$$

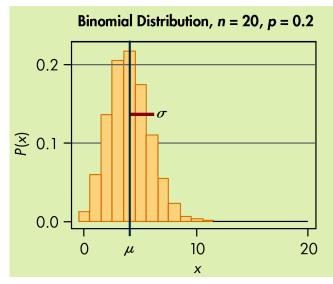
$$= (20)(0.2)$$

$$\bullet = 4.0$$

•
$$\sigma = \sqrt{npq}$$

$$\bullet = \sqrt{20(0.2)(0.8)} = \sqrt{3.2}$$

$$\bullet = 1.79$$



Histogram of Binomial Distribution B(20, 0.2)

• It is the expected standard deviation for the values of the random variable \boldsymbol{x} that occur in samples of size 20 drawn from this same population

QUESTIONS?



ANY QUESTION?