CHAPTER 5

# Discrete Probability Distributions

## Objectives

After completing this chapter, you should be able to

1. Construct a probability distribution for a random variable.
2. Find the mean, variance, standard deviation, and expected value for a discrete random variable.
3. Find the exact probability for $X$ successes in $n$ trials of a binomial experiment.
4. Find the mean, variance, and standard deviation for the variable of a binomial distribution.
5. Find probabilities for outcomes of variables, using the Poisson, hypergeometric, and multinomial distributions.

## Outline

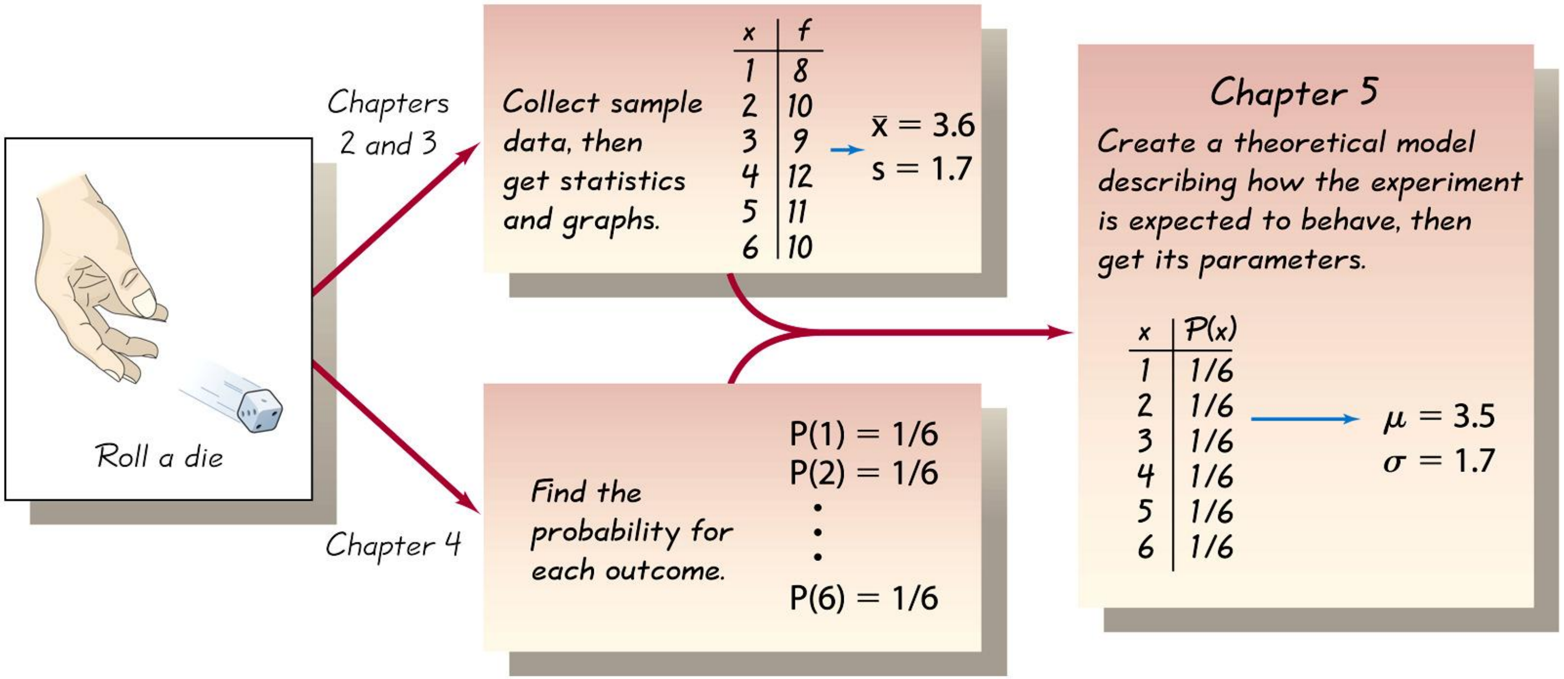
- Probability distribution describe what will probably happen instead of what actually did happen.

-To compare theoretical probabilities to actual results is order to determine whether outcomes are unusual.

-The very core of inferential statistics are based on some knowledge of probability distributions.

# Combining Descriptive Methods and Probabilities

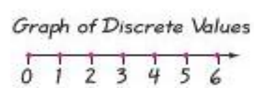
In this chapter we will construct probability distributions by presenting possible outcomes along with the relative frequencies we expect.



# 5.1 Probability Distributions

- A **random variable** is a variable whose values are determined by chance (x).
- A **discrete variable** is a variable can assume only a specific number of values; can be counted.


(a) Discrete Random Variable: Count of the number of movie patrons.




- A **continuous variable** is a variable that can assume all values in the interval between any two given values; can be measured.

# 5.1 Probability Distributions

A **discrete probability distribution** consists of the values a random variable can assume and the corresponding probabilities of the values.

Two requirement for a probability distribution:

1. The sum of the **probabilities** of all events in a sample space add up to 1.

2. Each probability is between 0 and 1, inclusively.

# Chapter 5
# Discrete Probability Distributions

## Section 5-1

Example 5-1

Page #254

# Example 5-1: Rolling a Die

Construct a probability distribution for rolling a single die.

| **Outcome** $X$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| **Probability** $P(X)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

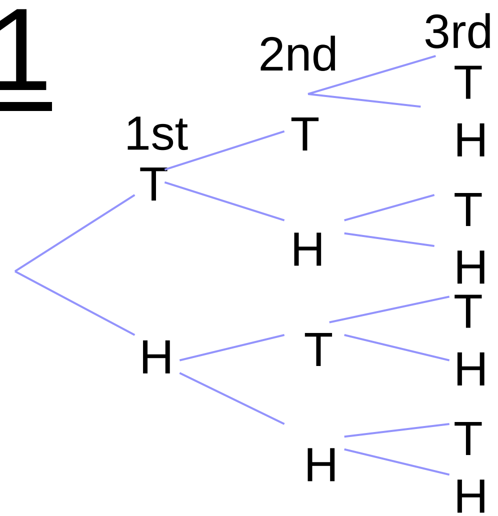
# Chapter 5
# Discrete Probability Distributions

Represent graphically the probability distribution for the sample space for tossing three coins.

## Section 5-1

Example 5-2

Page #254



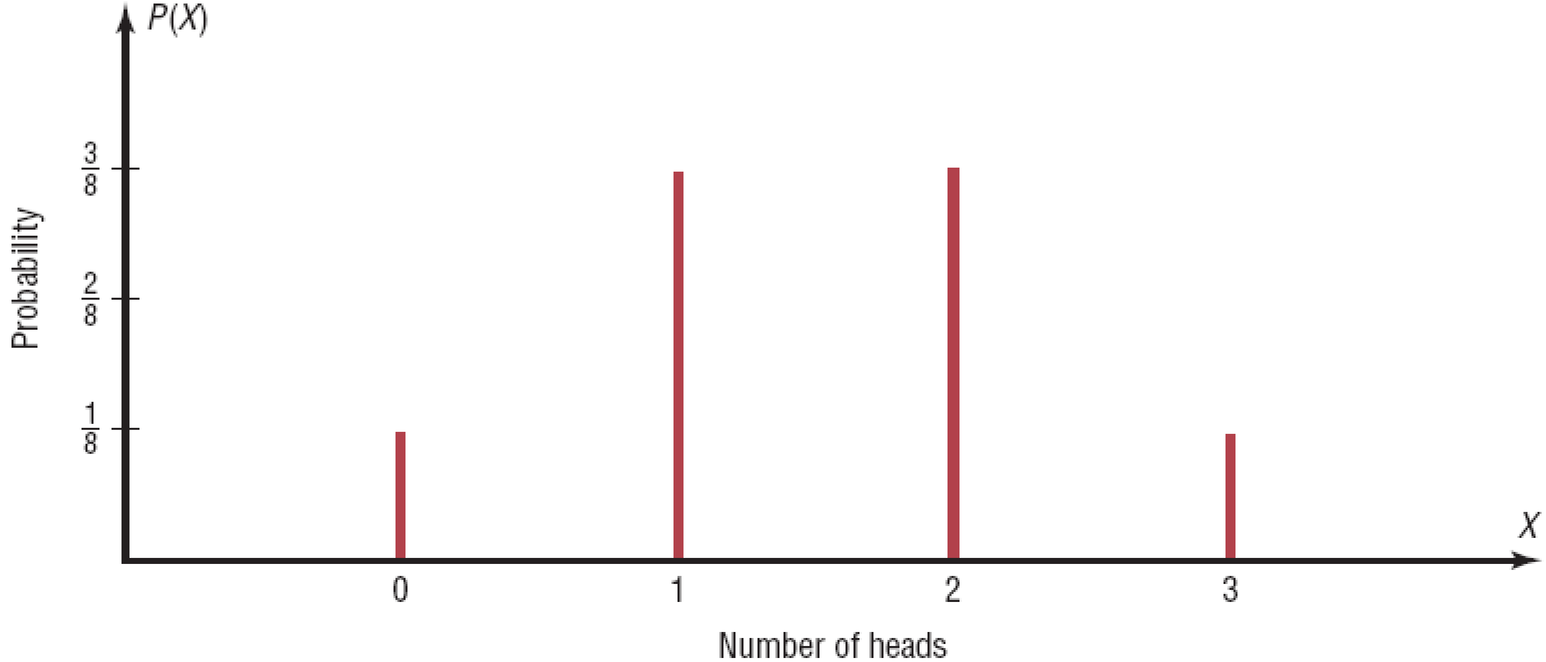

| Sample space | X | P(X) |
|---|---|---|
| TTT | 0 head | 1/8 |
| TTH | 1 head | 3/8 |
| THT | | |
| THH | 2 head | 3/8 |
| HTT | | |
| HTH | | |
| HHT | | |
| HHH | 3 head | 1/8 |

# Example 5-2: Tossing Coins

Represent graphically the probability distribution for the sample space for tossing three coins.

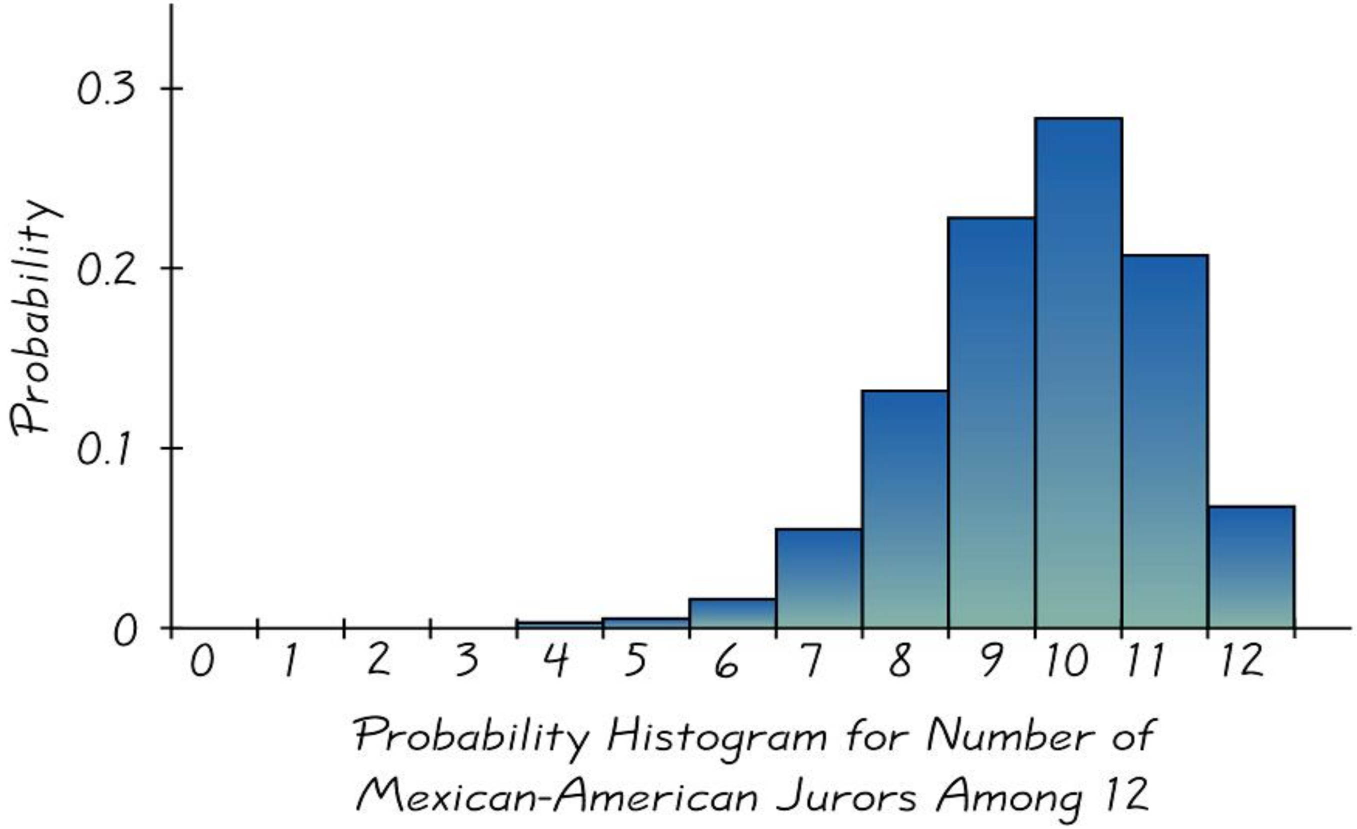
.

| Number of heads $X$ | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| Probability $P(X)$ | $\frac{1}{8}$ | $\frac{3}{8}$ | $\frac{3}{8}$ | $\frac{1}{8}$ |



# Graphs

The probability histogram is very similar to a relative frequency histogram, but the vertical scale shows probabilities.



# 5-2 Mean, Variance, Standard Deviation, and Expectation

MEAN: $\mu = \sum X \cdot P(X)$

Rationale for formulas:

μ = ∑(f * x) / N = ∑ [f * x / N] = ∑ [x * f / N] = ∑ [x * P(x)]

VARIANCE:

$$\sigma^2 = \sum \left[ X^2 \cdot P(X) \right] - \mu^2$$

# Mean, Variance, Standard Deviation, and Expectation

**Rounding Rule**

The mean, variance, and standard deviation should be rounded to one more decimal place than the outcome *X*.

When fractions are used, they should be reduced to lowest terms.

# Chapter 5
# Discrete Probability Distributions

## Section 5-2

Example 5-5

Page #260

# Example 5-5: Rolling a Die

Find the mean of the number of spots that appear when a die is tossed.

| Outcome $X$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Probability $P(X)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

$$\mu = \sum X \cdot P(X)$$

$$= 1 \cdot \tfrac{1}{6} + 2 \cdot \tfrac{1}{6} + 3 \cdot \tfrac{1}{6} + 4 \cdot \tfrac{1}{6} + 5 \cdot \tfrac{1}{6} + 6 \cdot \tfrac{1}{6}$$

$$= \tfrac{21}{6} = \boxed{3.5}$$

# Chapter 5
# Discrete Probability Distributions

## <u>Section 5-2</u>

Example 5-8

Page #261

# Example 5-8: Trips of 5 Nights or More

The probability distribution shown represents the number of trips of five nights or more that American adults take per year. (That is, 6% do not take any trips lasting five nights or more, 70% take one trip lasting five nights or more per year, etc.) Find the mean.

| **Number of trips** $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| **Probability** $P(X)$ | 0.06 | 0.70 | 0.20 | 0.03 | 0.01 |

# Example 5-8: Trips of 5 Nights or More

| Number of trips $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| Probability $P(X)$ | 0.06 | 0.70 | 0.20 | 0.03 | 0.01 |

$$\mu = \sum X \cdot P(X)$$

$$= 0(0.06) + 1(0.70) + 2(0.20)$$

$$+ 3(0.03) + 4(0.01)$$

$$= \boxed{1.2}$$

# Chapter 5
# Discrete Probability Distributions

## Section 5-2

Example 5-9

Page #262

# Example 5-9: Rolling a Die

Compute the variance and standard deviation for the probability distribution in Example 5–5.

| Outcome $X$ | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| Probability $P(X)$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{6}$ |

$$\sigma^2 = \sum \left[ X^2 \cdot P(X) \right] - \mu^2$$

$$\sigma^2 = 1^2 \cdot \tfrac{1}{6} + 2^2 \cdot \tfrac{1}{6} + 3^2 \cdot \tfrac{1}{6} + 4^2 \cdot \tfrac{1}{6} )$$

$$+ 5^2 \cdot \tfrac{1}{6} + 6^2 \cdot \tfrac{1}{6} - (3.5)^2$$

$$\sigma^2 = \boxed{2.9}, \quad \sigma = \boxed{1.7}$$

# Chapter 5
# Discrete Probability Distributions

## Section 5-2

Example 5-11

Page #263

# Example 5-11: On Hold for Talk Radio

A talk radio station has four telephone lines. If the host is unable to talk (i.e., during a commercial) or is talking to a person, the other callers are placed on hold. When all lines are in use, others who are trying to call in get a busy signal. The probability that 0, 1, 2, 3, or 4 people will get through is shown in the distribution. Find the variance and standard deviation for the distribution.

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(X)$ | 0.18 | 0.34 | 0.23 | 0.21 | 0.04 |

# Example 5-11: On Hold for Talk Radio

| $X$ | 0 | 1 | 2 | 3 | 4 |
|---|---|---|---|---|---|
| $P(X)$ | 0.18 | 0.34 | 0.23 | 0.21 | 0.04 |

$$\mu = 0(0.18) + 1(0.34) + 2(0.23) + 3(0.21) + 4(0.04) = 1.6$$

$$\sigma^2 = 0^2(0.18) + 1^2(0.34) + 2^2(0.23) + 3^2(0.21) + 4^2(0.04) - (1.6)^2$$

$$\sigma^2 = \boxed{1.2}, \quad \sigma = \boxed{1.1}$$

Identifying unusual results with the range rule of thumb:

Maximum usual value = $\mu + 2\sigma$

Minimum usual value = $\mu - 2\sigma$

Caution:
Know that the use of the number 2 in the range rule of thumb is somewhat arbitrary, and this rule is a guideline, not an absolutely rigid rule.

# Example 5-11: On Hold for Talk Radio

A talk radio station has four telephone lines. If the host is unable to talk (i.e., during a commercial) or is talking to a person, the other callers are placed on hold. When all lines are in use, others who are trying to call in get a busy signal.

Should the station have considered getting more phone lines installed?

# Example 5-11: On Hold for Talk Radio

No, the four phone lines should be sufficient.

The mean number of people calling at any one time is 1.6.

Since the standard deviation is 1.1, most callers would be accommodated by having four phone lines because $\mu + 2\sigma$ would be

$$1.6 + 2(1.1) = 1.6 + 2.2 = 3.8.$$

Very few callers would get a busy signal since at least 75% of the callers would either get through or be put on hold. (See Chebyshev's theorem in Section 3–2.)

# Expectation

- The **expected value**, or **expectation**, of a discrete random variable of a probability distribution is the theoretical average of the variable.
- The expected value is, by definition, the mean of the probability distribution.

$$E(X) = \mu = \sum X \cdot P(X)$$

# Chapter 5
# Discrete Probability Distributions

## Section 5-2

Example 5-13

Page #265

# Example 5-13: Winning Tickets

One thousand tickets are sold at \$1 each for four prizes of \$100, \$50, \$25, and \$10. After each prize drawing, the winning ticket is then returned to the pool of tickets. What is the expected value if you purchase two tickets?

| **Gain *X*** | \$98 | \$48 | \$23 | \$8 | - \$2 |
|---|---|---|---|---|---|
| **Probability *P(X)*** | $\frac{2}{1000}$ | $\frac{2}{1000}$ | $\frac{2}{1000}$ | $\frac{2}{1000}$ | $\frac{992}{1000}$ |

$$E(X) = \$98 \cdot \frac{2}{1000} + \$48 \cdot \frac{2}{1000} + \$23 \cdot \frac{2}{1000}$$

$$+ \$8 \cdot \frac{2}{1000} + (-\$2) \cdot \frac{992}{1000} = \boxed{-\$1.63}$$

# Example 5-13: Winning Tickets

One thousand tickets are sold at $1 each for four prizes of $100, $50, $25, and $10. After each prize drawing, the winning ticket is then returned to the pool of tickets. What is the expected value if you purchase two tickets? *Alternate Approach*

| **Gain *X*** | $100 | $50 | $25 | $10 | $0 |
|---|---|---|---|---|---|
| **Probability *P(X)*** | $\frac{2}{1000}$ | $\frac{2}{1000}$ | $\frac{2}{1000}$ | $\frac{2}{1000}$ | $\frac{992}{1000}$ |

$$E(X) = \$100 \cdot \frac{2}{1000} + \$50 \cdot \frac{2}{1000} + \$25 \cdot \frac{2}{1000}$$

$$+ \$10 \cdot \frac{2}{1000} + \$0 \cdot \frac{992}{1000} - \$2 = \boxed{-\$1.63}$$

# 5-3 The Binomial Distribution

- Many types of probability problems have only two possible outcomes or they can be reduced to two outcomes.
- Examples include: when a coin is tossed it can land on heads or tails (by chance: <u>50/50</u>), when a baby is born it is either a boy or girl, etc.
- It involve proportions used with methods of inferential statistics.

# The Binomial Distribution

The **binomial experiment** is a probability experiment that satisfies these requirements:

1. Each trial can have only two possible outcomes—success or failure.
2. There must be a fixed number of trials.
3. The outcomes of each trial must be independent of each other.
4. The probability of success must remain the same for each trial.

# Notation for the Binomial Distribution

| | |
|---|---|
| $P(S)$ | The symbol for the probability of success |
| $P(F)$ | The symbol for the probability of failure |
| $p$ | The numerical probability of success |
| $q$ | The numerical probability of failure |

$P(S) = p$ and $P(F) = 1 - p = q$

| | |
|---|---|
| $n$ | The number of trials |
| $X$ | The number of successes, note that $X = 0, 1, 2, 3, \ldots, n$ |
| $P(X)$ | The probability of getting exactly x successes among the n trial. |

# The Binomial Distribution

In a binomial experiment, the probability of exactly $X$ successes in $n$ trials is

$$P(X) = \frac{n!}{(n-X)!X!} \cdot p^X \cdot q^{n-X}$$

or

$$P(X) = \underbrace{{}_nC_x}_{\text{number of possible desired outcomes}} \cdot \underbrace{p^X \cdot q^{n-X}}_{\text{probability of a desired outcome}}$$

# Chapter 5
# Discrete Probability Distributions

## Section 5-3

Example 5-16

Page #272

# Example 5-16: Survey on Doctor Visits

A survey found that one out of five Americans say he or she has visited a doctor in any given month. If 10 people are selected at random, find the probability that exactly 3 will have visited a doctor last month. (binomial experiment?)

$$P(X)=\frac{n!}{(n-X)!X!}\cdot p^{X}\cdot q^{n-X}$$

$$n=10,\text{"one out of five"}\rightarrow p=\tfrac{1}{5}, X=3$$

$$P(3)=\frac{10!}{7!3!}\cdot\left(\frac{1}{5}\right)^{3}\cdot\left(\frac{4}{5}\right)^{7}=\boxed{0.201}$$

# Chapter 5
# Discrete Probability Distributions

## Section 5-3

Example 5-17

Page #273

# Example 5-17: Survey on Employment

A survey from Teenage Research Unlimited (Northbrook, Illinois) found that 30% of teenage consumers receive their spending money from part-time jobs. If 5 teenagers are selected at random, find the probability that at least 3 of them will have part-time jobs. (binomial experiment?)

$$n = 5,\ p = 0.30,\ \text{"at least 3"} \rightarrow X = 3, 4, 5$$

$$P(3) = \frac{5!}{2!3!} \cdot (0.30)^3 \cdot (0.70)^2 = 0.132$$

$$P(4) = \frac{5!}{1!4!} \cdot (0.30)^4 \cdot (0.70)^1 = 0.028 = 3$$

$$P(5) = \frac{5!}{0!5!} \cdot (0.30)^5 \cdot (0.70)^0 = 0.002$$

$$P(X \geq 3) = 0.132 + 0.028 + 0.002 = \boxed{0.162}$$

# Chapter 5
# Discrete Probability Distributions

## Section 5-3

Example 5-18

Page #273

# Example 5-18: Tossing Coins

A coin is tossed 3 times. Find the probability of getting exactly two heads, using Table B.

$$n = 3, p = \tfrac{1}{2} = 0.5, X = 2 \rightarrow P(2) = \boxed{0.375}$$

| | | | | | | | p | | p = 0.5 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| n | X | 0.05 | 0.1 | 0.2 | 0.3 | 0.4 | 0.5 | 0.6 | 0.7 | 0.8 | 0.9 | 0.95 |
| 2 | 0 | | | | | | | | | | | |
| | 1 | | | | | | | | | | | |
| | 2 | | | | | | | | | | | |
| 3 | 0 | | | | | | 0.125 | | | | | |
| n = 3 | 1 | | | | | | 0.375 | | | | | |
| | 2 | | | | | | 0.375 | | | | | |
| X = 2 | 3 | | | | | | 0.125 | | | | | |

# The Binomial Distribution

The mean, variance, and standard deviation of a variable that has the *binomial distribution* can be found by using the following formulas.

Mean: $\mu = np$

Variance: $\sigma^2 = npq$

Standard Deviation: $\sigma = \sqrt{npq}$

# Chapter 5
# Discrete Probability Distributions

## Section 5-3

Example 5-23

Page #276

# Example 5-23: Likelihood of Twins

The *Statistical Bulletin* published by Metropolitan Life Insurance Co. reported that 2% of all American births result in twins. If a random sample of 8000 births is taken, find the mean, variance, and standard deviation of the number of births that would result in twins. (binomial experiment?)

$$\mu = np = 8000(0.02) = \boxed{160}$$

$$\sigma^2 = npq = 8000(0.02)(0.98) = 156.8 = \boxed{157}$$

$$\sigma = \sqrt{npq} = \sqrt{8000(0.02)(0.98)} = 12.5 = \boxed{13}$$

# 5-4 Other Types of Distributions

- The **multinomial distribution** is similar to the binomial distribution but has the advantage of allowing one to compute probabilities when there are more than two outcomes.

$$P(X)=\frac{n!}{X_1!X_2!X_3!\cdots X_k!}\cdot p_1^{X_1}\cdot p_2^{X_2}\cdot p_3^{X_3}\cdots p_k^{X_k}$$

- The binomial distribution is a special case of the multinomial distribution.

# Chapter 5
# Discrete Probability Distributions

## Section 5-4

Example 5-24

Page #283

# Example 5-24: Leisure Activities

In a large city, 50% of the people choose a movie, 30% choose dinner and a play, and 20% choose shopping as a leisure activity. If a sample of 5 people is randomly selected, find the probability that 3 are planning to go to a movie, 1 to a play, and 1 to a shopping mall.

$$P(X) = \frac{n!}{X_1!X_2!X_3!\cdots X_k!} \cdot p_1^{X_1} \cdot p_2^{X_2} \cdot p_3^{X_3} \cdots p_k^{X_k}$$

$$P(X) = \frac{5!}{3!1!1!} \cdot (0.50)^3 (0.30)^1 (0.20)^1 = \boxed{0.15}$$

# Other Types of Distributions

- The **Poisson distribution** is a distribution useful when *n* is large and *p* is small and when the independent variables occur over a period of time.
- The Poisson distribution can also be used when a density of items is distributed over a given area or volume, such as the number of plants growing per acre or the number of defects in a given length of videotape.

# Other Types of Distributions

## Poisson Distribution

The probability of $X$ occurrences in an interval of time, volume, area, etc., for a variable, where $\lambda$ (Greek letter lambda) is the mean number of occurrences per unit (time, volume, area, etc.), is

$$P(X;\lambda)=\frac{e^{-\lambda}\lambda^{X}}{X!} \quad \text{where } X=0,1,2,\ldots$$

The letter $e$ is a constant approximately equal to 2.7183.

# Chapter 5
# Discrete Probability Distributions

## Section 5-4

Example 5-27

Page #285

# Example 5-27: Typographical Errors

If there are 200 typographical errors randomly distributed in a 500-page manuscript, find the probability that a given page contains exactly 3 errors.

First, find the mean number $\lambda$ of errors. With 200 errors distributed over 500 pages, each page has an average of $\lambda = \frac{200}{500} = 0.4$ errors per page.

$$P(X;\lambda) = \frac{e^{-\lambda}\lambda^{X}}{X!} = \frac{e^{-0.4}(0.4)^{3}}{3!} = \boxed{0.0072}$$

Thus, there is less than 1% chance that any given page will contain exactly 3 errors.

# Other Types of Distributions

- The **hypergeometric distribution** is a distribution of a variable that has two outcomes when sampling is done without replacement.

# Other Types of Distributions

## Hypergeometric Distribution

Given a population with only two types of objects (females and males, defective and nondefective, successes and failures, etc.), such that there are *a* items of one kind and *b* items of another kind and *a*+*b* equals the total population, the probability *P(X)* of selecting without replacement a sample of size *n* with *X* items of type *a* and *n*-*X* items of type *b is*

$$P(X) = \frac{{}_{a}C_{X} \cdot {}_{b}C_{n-X}}{{}_{a+b}C_{n}}$$

# Chapter 5
# Discrete Probability Distributions

## Section 5-4

Example 5-31

Page #288

# Example 5-31: House Insurance

A recent study found that 2 out of every 10 houses in a neighborhood have no insurance. If 5 houses are selected from 10 houses, find the probability that exactly 1 will be uninsured.

$$a = 2, a + b = 10 \to b = 8, \quad X = 1, n = 5 \to n - X = 4$$

$$P(X) = \frac{{}_{a}C_{X} \cdot {}_{b}C_{n-X}}{{}_{a+b}C_{n}}$$

$$P(X) = \frac{{}_{2}C_{1} \cdot {}_{8}C_{4}}{{}_{10}C_{5}} = \frac{2 \cdot 70}{252} = \frac{140}{252} = \boxed{\frac{5}{9}}$$