Chapter 5
Discrete Probability Distributions

- Random Variables
- Developing Discrete Probability Distributions
- Expected Value and Variance
- Bivariate distributions and Covariance
- Financial Portfolios
- Binomial Probability Distribution
- Poisson Probability Distribution
- Hypergeometric Probability Distribution

Random Variables (1 of 2)

- A random variable is a numerical description of the outcome of an experiment.
  - A discrete random variable may assume either a finite number of values or an infinite sequence of values.
  - A continuous random variable may assume any numerical value in an interval or collection of intervals.

Discrete Random Variable with a Finite Number of Values

<u>Example:</u> An accountant taking CPA examination

The examination has four parts.

Let random variable x = the number of parts of the CPA examination passed

x may assume the finite number of values 0,1,2,3 or 4.

Discrete Random Variable with an Infinite Number of Values

Example: Cars arriving at a toll booth

Let *x* = number of cars arriving in one day,

where *x* can take on the values 0, 1, 2, . . .

We can count the customers arriving, but there is no finite upper limit on the number that might arrive.

Random Variables (2 of 2)

Random Experiment	Random Variable (*x*)	Possible Values for the Random Variable
Flip a coin	Face of coin showing	1 if heads; 0 if tails
Roll a die	Number of dots showing on top of die	1, 2, 3, 4, 5, 6
Contact five customers	Number of customers who place an order	0, 1, 2, 3, 4, 5
Operate a health care clinic for one day	Number of patients who arrive	0, 1, 2, 3, ...
Offer a customer the choice of two products	Product chosen by customer	0 if none; 1 if choose product A; 2 if choose product B

Discrete Probability Distributions (1 of 8)

- The probability distribution for a random variable describes how probabilities are distributed over the values of the random variable.

- We can describe a discrete probability distribution with a table, graph, or formula.

Discrete Probability Distributions (2 of 8)

Two types of discrete probability distributions:

- First type: uses the rules of assigning probabilities to experimental outcomes to determine probabilities for each value of the random variable.
- Second type: uses a special mathematical formula to compute the probabilities for each value of the random variable.

Discrete Probability Distributions (3 of 8)

- The probability distribution is defined by a probability function, denoted by *f(x)*, that provides the probability for each value of the random variable.
- The required conditions for a discrete probability function are:

$$f(x) \geq 0 \text{ and } \sum f(x) = 1$$

Discrete Probability Distributions (4 of 8)

- There are three methods for assigning probabilities to random variables:
    - Classical method
    - Subjective method
    - Relative frequency
- The use of the relative frequency method to develop discrete probability distributions leads to what is called an empirical discrete distribution.

Discrete Probability Distributions (5 of 8)

Example: DiCarlo Motors

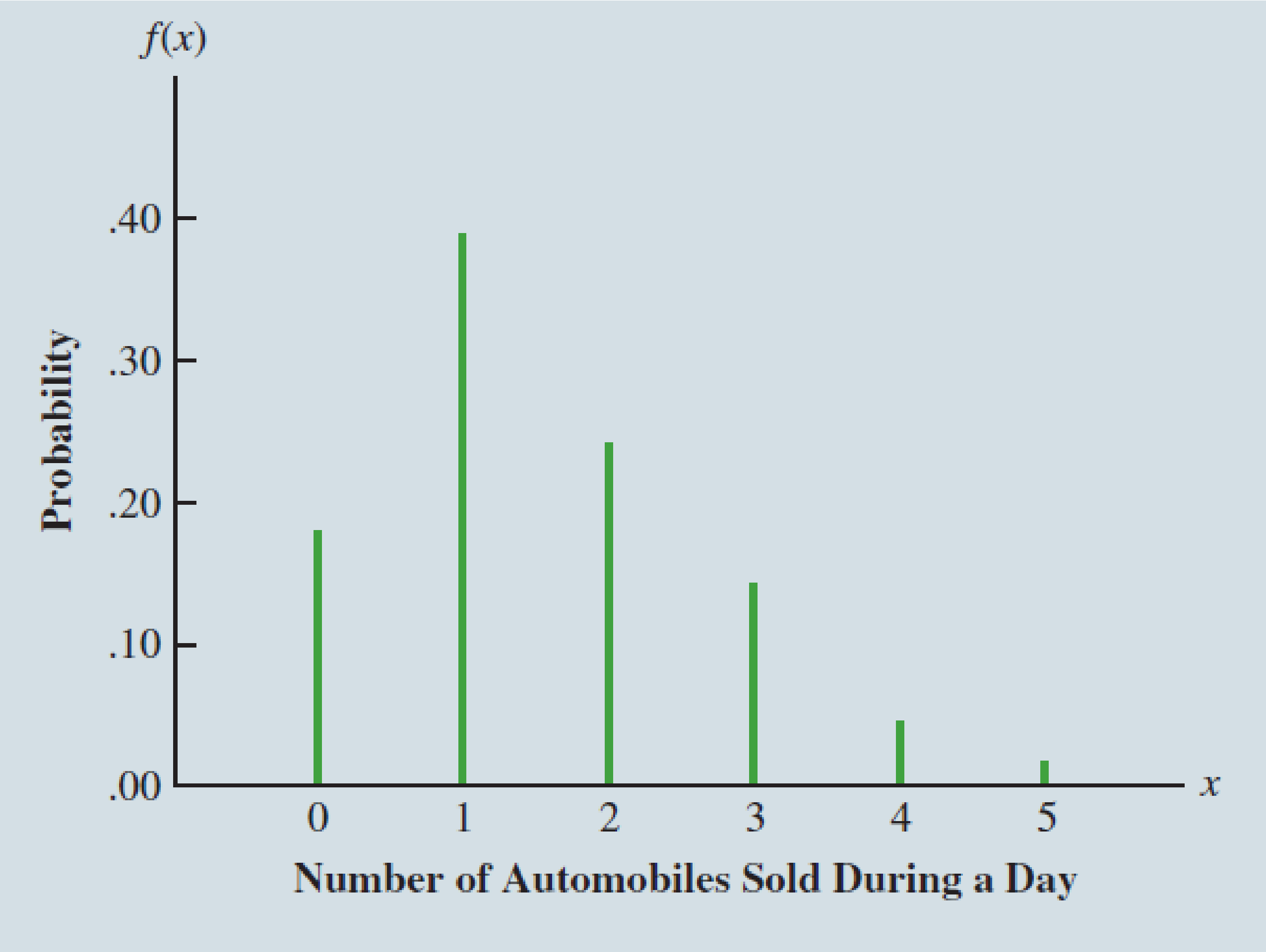
Using past data on daily car sales for 300 days, a tabular representation of the probability distribution for sales was developed.

Number of cars sold	Number of days	x	$f(x)$
0	54	0	.18
1	117	1	.39
2	72	2	.24
3	42	3	.14
4	12	4	.04
5	3	5	.01
Total	**300**		**1.00**

Discrete Probability Distributions (6 of 8)

Example: DiCarlo Motors

Graphical representation of the probability distribution



Discrete Probability Distributions (7 of 8)

- In addition to tables and graphs, a formula that gives the probability function, $f(x)$, for every value of x is often used to describe the probability distributions.
- Some of the discrete probability distributions specified by formulas are
    - Discrete—uniform distribution
    - Binomial distribution
    - Poisson distribution
    - Hypergeometric distribution

Discrete Probability Distributions (8 of 8)

- The discrete uniform probability distribution is the simplest example of a discrete probability distribution given by a formula.
- The discrete uniform probability function is

$$f(x) = 1/n$$

where: n = the number of values the random variable may assume

The values of the random variable are equally likely.

Expected Value

- The expected value, or mean, of a random variable is a measure of its central location.

$$E(x) = \mu = \sum xf(x)$$

- The expected value is a weighted average of the values the random variable may assume. The weights are the probabilities.
- The expected value does not have to be a value the random variable can assume.

Variance and Standard Deviation

- The variance summarizes the variability in the values of a random variable.

$$Var(x) = \sigma^2 = \sum (x - \mu)^2 f(x)$$

- The variance is a weighted average of the squared deviations of a random variable from its mean. The weights are the probabilities.
- The standard deviation, σ, is defined as the positive square root of the variance.

Discrete Probability Distributions (1 of 2)

Example: DiCarlo Motors

x	*f(x)*	*xf(x)*
0	.18	.00
1	.39	.39
2	.24	.48
3	.14	.42
4	.04	.16
5	.01	.05
	1.00	**1.50**

$E(x) = 1.50 =$ expected number of cars sold in a day

Discrete Probability Distributions (2 of 2)

Example: DiCarlo Motors

x	$x - \mu$	$(x-\mu)^2$	$f(x)$	$(x-\mu)^2 f(x)$
0	$0 - 1.5 = -1.5$	2.25	.18	$2.25(.18) = .4050$
1	$1 - 1.5 = -.5$	.25	.39	.0975
2	$2 - 1.5 = .5$	.25	.24	.0600
3	$3 - 1.5 = 1.5$	2.25	.14	.3150
4	$4 - 1.5 = 2.5$	6.25	.04	.2500
5	$5 - 1.5 = 3.5$	12.25	.01	.1225
			1.00	**1.2500**

Variance of daily sales $= \sigma^2 = 1.25$

Standard deviation of daily sales = 1.118 cars

Using Excel to Compute the Expected Value, Standard Deviation, and Variance

- Excel Formula and Value Worksheets

	A	B	C	D
1	Sales	Probability	Sq Dev from Mean	
2	0	0.18	=(A2-B9)^2	
3	1	0.39	=(A3-B9)^2	
4	2	0.24	=(A4-B9)^2	
5	3	0.14	=(A5-B9)^2	
6	4	0.04	=(A6-B9)^2	
7	5	0.01	=(A7-B9)^2	
8				
9	Mean	=SUMPRODUCT(A2:A7,B2:B7)		
10				
11	Variance	=SUMPRODUCT(C2:C7,B2:B7)		
12				
13	Std Deviation	=SQRT(B11)		
14				

	A	B	C	D
1	Sales	Probability	Sq Dev from Mean	
2	0	0.18	2.25	
3	1	0.39	0.25	
4	2	0.24	0.25	
5	3	0.14	2.25	
6	4	0.04	6.25	
7	5	0.01	12.25	
8				
9	Mean	1.5		
10				
11	Variance	1.25		
12				
13	Std Deviation	1.118		
14				

Bivariate Distributions

- A probability distribution involving two random variables is called a bivariate probability distribution.

- Each outcome of a bivariate experiment consists of two values, one for each random variable.

  Example: Rolling a pair of dice

- When dealing with bivariate probability distributions, we are often interested in the relationship between the random variables.

A Bivariate Discrete Probability Distribution (1 of 6)

Example: DiCarlo Motors

The crosstabulation of daily car sales for 300 days at DiCarlo's Saratoga and Geneva dealership is given below:

Geneva Dealership	Saratoga Dealership 0	1	2	3	4	5	Total
0	21	30	24	9	2	0	**86**
1	21	36	33	18	2	1	**111**
2	9	42	9	12	3	2	**77**
3	3	9	6	3	5	0	**26**
Total	**54**	**117**	**72**	**42**	**12**	**3**	**300**

A Bivariate Discrete Probability Distribution (2 of 6)

Example: DiCarlo Motors

Bivariate empirical discrete probability distribution for daily sales at DiCarlo dealerships in Saratoga and Geneva is shown below.

Geneva Dealership	Saratoga Dealership 0	1	2	3	4	5	Total
0	.0700	.1000	.0800	.0300	.0067	.0000	**.2867**
1	.0700	.1200	.1100	.0600	.0067	.0033	**.3700**
2	.0300	.1400	.0300	.0400	.0100	.0067	**.2567**
3	.0100	.0300	.0200	.0100	.0167	.0000	**.0867**
Total	**.18**	**.39**	**.24**	**.14**	**.04**	**.01**	**1.0000**

A Bivariate Discrete Probability Distribution (3 of 6)

- **Example:** DiCarlo Motors

Expected value and variance for daily car sales at Geneva dealership.

x	$f(x)$	$xf(x)$	$x-E(x)$	$(x-E(x))^2$	$(s-E(x))^2 f(x)$
0	.2867	.0000	−1.1435	1.3076	.3749
1	.3700	.3700	−.1435	0.0206	.0076
2	.2567	.5134	.8565	0.8565	.1883
3	.0867	.2601	1.8565	1.8565	.2988
		$E(x)=1.1435$			$Var(x)=.8696$

CENGAGE

A Bivariate Discrete Probability Distribution (4 of 6)

- **Example:** DiCarlo Motors

  Expected value and variance for total daily car sales data.

s	$f(s)$	$sf(s)$	$s-E(s)$	$(s-E(s))^2$	$(s-E(s))^2 f(s)$
0	.0700	.0000	−2.6433	6.9872	.4891
1	.1700	.1700	−1.6433	2.7005	.4591
2	.2300	.4600	−0.6433	0.4139	.0952
3	.2900	.8700	0.3567	0.1272	.0369
4	.1267	.5067	1.3567	1.8405	.2331
5	.0667	.3333	2.3567	5.5539	.3703
6	.0233	.1400	3.3567	11.2672	.2629
7	.0233	.1633	4.3567	18.9805	.4429
8	.0000	.0000	5.3567	28.6939	.0000
		$E(s)=2.6433$			$Var(s)=2.3895$

A Bivariate Discrete Probability Distribution (5 of 6)

Covariance for random variables x and y.

$$Var_{xy} = [Var(x+y) - Var(x) - Var(y)]/2$$

$$(2.3895 - .8696 - 1.25)/2$$

$$= .1350$$

A Bivariate Discrete Probability Distribution (6 of 6)

Correlation between random variables *x* and *y*

$$\rho_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

$$\sigma_x = \sqrt{.8696} = .9325$$

$$\sigma_y = \sqrt{1.25} = 1.1180$$

$$\rho_{xy} = \frac{.1350}{(.9325)(1.1180)} = .1295$$

Binomial Probability Distribution (1 of 9)

Four Properties of a Binomial Experiment

1. The experiment consists of a sequence of n identical trials.
2. Two outcomes, success and failure, are possible on each trial.
3. The probability of a success, denoted by p, and failure denoted by $1-p$ does not change from trial to trial. (This referred to as the stationarity assumption.)
4. The trials are independent.

Binomial Probability Distribution (2 of 9)

- Our interest is in the number of successes occurring in the *n* trials.
- We let *x* denote the number of successes occurring in the *n* trials.

Binomial Probability Distribution (3 of 9)

Binomial Probability Function

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)}$$

where:

$x =$ the number of successes

$p =$ the probability of a success on one trial

$n =$ the number of trials

$f(x) =$ the probability of x successes in n trials

$n! = n(n-1)(n-2)$ ….. $(2)(1)$

Binomial Probability Distribution (4 of 9)

- Binomial Probability Function

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)}$$

$\frac{n!}{x!(n-x)!}$: Number of experimental outcomes providing exactly x successes in n trials

$p^x (1-p)^{(n-x)}$: Probability of a particular sequence of trial outcomes with x successes in n trials

Binomial Probability Distribution (5 of 9)

Example: Martin Clothing store

The store manager wants to determine the purchase decisions of next three customers who enter the clothing store. On the basis of past experience, the store manager estimates the probability that any one customer will make a purchase is .30.

What is the probability that two of the next three customers will make a purchase?

Binomial Probability Distribution (6 of 9)

Example: Martin Clothing store

Using *S* to denote success (a purchase) and *F* to denote failure (no purchase), we are interested in experimental outcomes involving two successes in the three trials.

- The probability of the first two customers buying and the third customer not buying denoted by (*S*, *S*, *F*), is given by

$$(p)(p)(1-p)$$

- With a .30 probability of a customer buying on any one trial, the probability of the first two customers buying and the third customer not buying is $(0.3)(0.3)(1-0.3) = .063$.

Binomial Probability Distribution (7 of 9)

Example: Martin Clothing store

Two other experimental outcomes result in two successes and one failure. The probabilities for all three experimental outcomes involving two successes follow:

Experimental outcome	Probability
(*S*, *S*, *F*)	.063
(*S*, *F*, *S*)	.063
(*F*, *S*, *S*)	.063

CENGAGE

Binomial Probability Distribution (8 of 9)

<u>Example:</u> Martin Clothing store

Using the probability function:

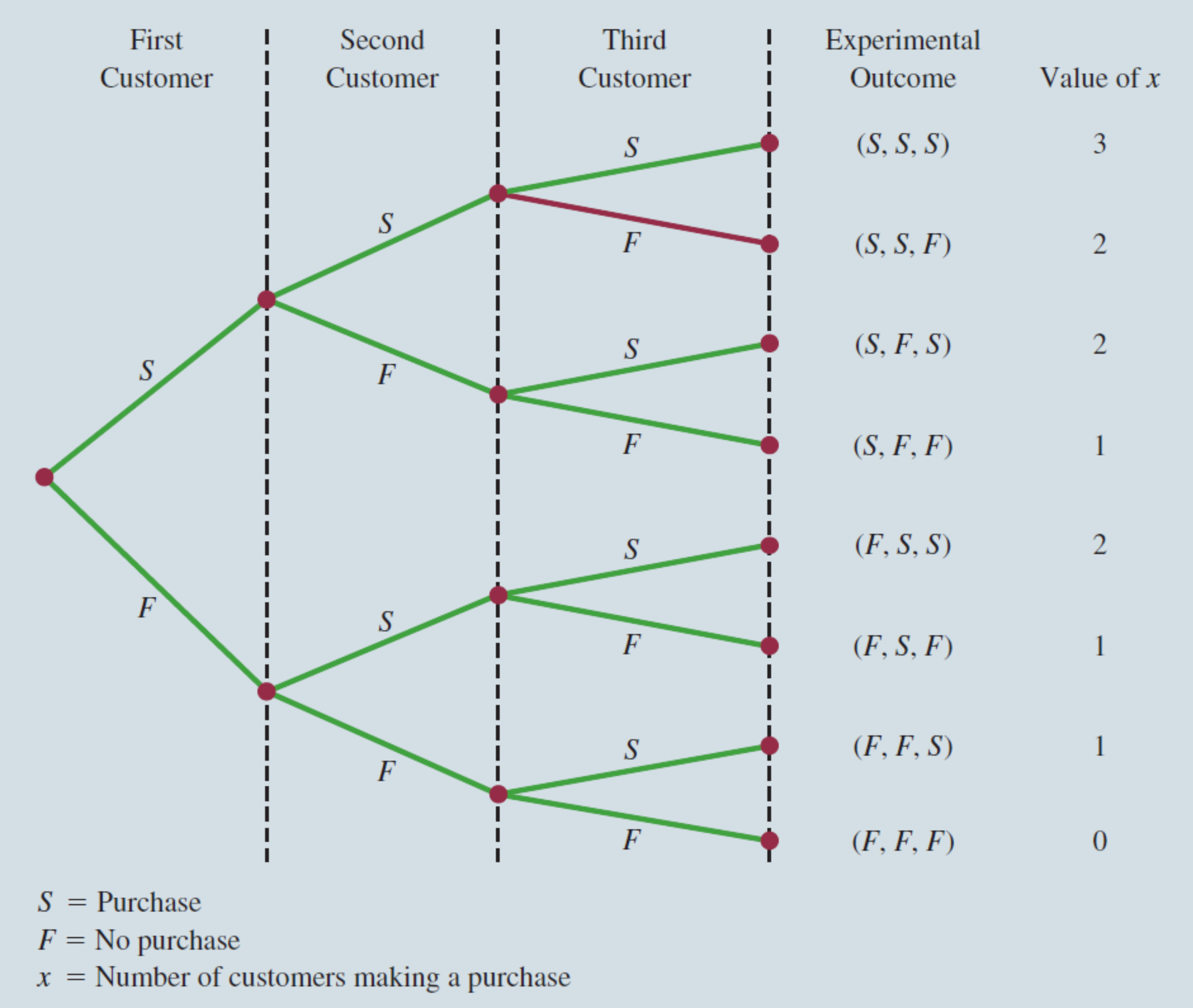
Let: $p = .30$, $n = 3$, $x = 2$

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)}$$

$$f(1) = \frac{3!}{2!(3-2)!} (0.3)^2 (0.7)^1 = .189$$

Binomial Probability Distribution (9 of 9)

Example: Martin Clothing store



Using Excel to Compute Binomial Probabilities

- Excel Formula and Value Worksheets

	A	B	C	D	E
1			**Number of Trials (*n*)**	3	
2			**Probability of Success (*p*)**	0.3	
3					
4		***x***	***f(x)***		
5		0	=BINOM.DIST(B5,D1,D2,FALSE)		
6		1	=BINOM.DIST(B6,D1,D2,FALSE)		
7		2	=BINOM.DIST(B7,D1,D2,FALSE)		
8		3	=BINOM.DIST(B8,D1,D2,FALSE)		
9					

	A	B	C	D	E
1		**Number of Trials (*n*)**		3	
2		**Probability of Success (*p*)**		0.3	
3					
4		***x***	***f(x)***		
5		0	0.343		
6		1	0.441		
7		2	0.189		
8		3	0.027		
9					

CENGAGE

Using Excel to Compute Cumulative Binomial Probabilities

- Excel Formula and Values Worksheets

  For number of purchases with 10 customers:

	A	B	C	D	E
1			Number of Trials (n)	10	
2			Probability of Success (p)	0.3	
3					
4		x	$f(x)$	Cum Prob	
5		0	=BINOM.DIST(B5,D1,D2,FALSE)	=BINOM.DIST(B5,D1,D2,TRUE)	
6		1	=BINOM.DIST(B6,D1,D2,FALSE)	=BINOM.DIST(B6,D1,D2,TRUE)	
7		2	=BINOM.DIST(B7,D1,D2,FALSE)	=BINOM.DIST(B7,D1,D2,TRUE)	
8		3	=BINOM.DIST(B8,D1,D2,FALSE)	=BINOM.DIST(B8,D1,D2,TRUE)	
9		4	=BINOM.DIST(B9,D1,D2,FALSE)	=BINOM.DIST(B9,D1,D2,TRUE)	
10		5	=BINOM.DIST(B10,D1,D2,FALSE)	=BINOM.DIST(B10,D1,D2,TRUE)	
11		6	=BINOM.DIST(B11,D1,D2,FALSE)	=BINOM.DIST(B11,D1,D2,TRUE)	
12		7	=BINOM.DIST(B12,D1,D2,FALSE)	=BINOM.DIST(B12,D1,D2,TRUE)	
13		8	=BINOM.DIST(B13,D1,D2,FALSE)	=BINOM.DIST(B13,D1,D2,TRUE)	
14		9	=BINOM.DIST(B14,D1,D2,FALSE)	=BINOM.DIST(B14,D1,D2,TRUE)	
15		10	=BINOM.DIST(B15,D1,D2,FALSE)	=BINOM.DIST(B15,D1,D2,TRUE)	
16					

	A	B	C	D	E
1		Number of Trials (n)		10	
2		Probability of Success (p)		0.3	
3					
4		x	$f(x)$	Cum Prob	
5		0	0.0282	0.0282	
6		1	0.1211	0.1493	
7		2	0.2335	0.3828	
8		3	0.2668	0.6496	
9		4	0.2001	0.8497	
10		5	0.1029	0.9527	
11		6	0.0368	0.9894	
12		7	0.0090	0.9984	
13		8	0.0014	0.9999	
14		9	0.0001	1.0000	
15		10	0.0000	1.0000	
16					

Binomial Probabilities and Cumulative Probabilities

- Statisticians have developed tables that give probabilities and cumulative probabilities for a binomial random variable.
- These tables can be found in some statistics textbooks.
- With modern calculators and the capability of statistical software packages, such tables are almost unnecessary.

Expected Value and Variance for Binomial Distribution (1 of 2)

- Expected Value $E(x) = \mu = np$
- Variance $Var(x) = \sigma^2 = np(1 - p)$
- Standard Deviation $\sigma = \sqrt{np(1 - p)}$

Expected Value and Variance for Binomial Distribution (2 of 2)

Example: Martin Clothing store

Expected Value $E(x) = np = 3(.3) = .9$

$Var(x) = np(1 - p) = 3(.3)(1 - .3) = .63$

Standard Deviation $= \sigma = \sqrt{np(1-p)} = \sigma = \sqrt{.63} = .79$

Poisson Probability Distribution (1 of 7)

- A Poisson distributed random variable is often useful in estimating the number of occurrences over a specified interval of time or space.
- It is a discrete random variable that may assume an infinite sequence of values (x = 0, 1, 2, . . .).

Poisson Probability Distribution (2 of 7)

Examples

- Number of knotholes in 14 linear feet of pine board
- Number of vehicles arriving at a toll booth in one hour
- Number of leaks in 100 miles of pipeline

Bell Labs used the Poisson distribution to model the arrival of phone calls.

Poisson Probability Distribution (3 of 7)

Properties of a Poisson Experiment

1. The probability of an occurrence is the same for any two intervals of equal length.
2. The occurrence or nonoccurrence in any interval is independent of the occurrence or nonoccurrence in any other interval.

Poisson Probability Distribution (4 of 7)

Poisson Probability Function

$$f(x) = \frac{\mu^x e^{-\mu}}{x!}$$

where:

$x =$ the number of occurrences in an interval

$f(x) =$ the probability of x occurrences in an interval

$\mu =$ mean number of occurrences in an interval

$e = 2.71828$

$x! = x(x-1)(x-2)\ldots(2)(1)$

CENGAGE

Poisson Probability Distribution (5 of 7)

Poisson Probability Function

- Since there is no stated upper limit for the number of occurrences, the probability function $f(x)$ is applicable for values $x = 0, 1, 2, \ldots$ without limit.
- In practical applications, x will eventually become large enough so that $f(x)$ is approximately zero and the probability of any larger values of x becomes negligible.

Poisson Probability Distribution (6 of 7)

Example: Arrivals at the emergency room

The average number of patients arriving at the emergency room at a large hospital in a 15-minute period of time on weekday mornings is 10.

What is the probability of 5 arrivals in a 15-minute period of time on a weekday morning?

Poisson Probability Distribution (7 of 7)

Example: Arrivals at the emergency room

Using the probability function:

$$\mu = 10;\ x = 5$$

$$f(5) = \frac{10^5 (2.71828)^{-10}}{5!}$$

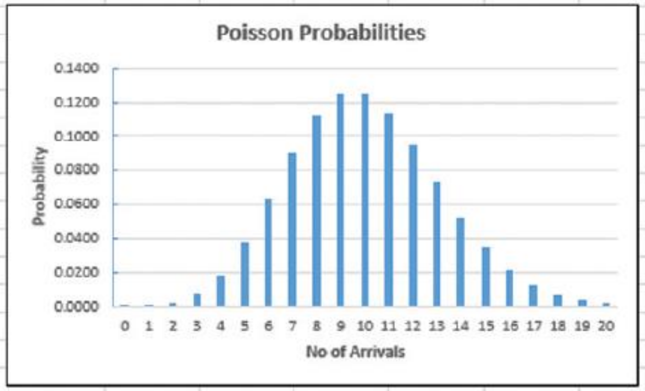
$$= .0378$$

Using Excel to Compute Poisson Probabilities

- Excel Formula and Values Worksheets

	A	B	C	D	E
1			Mean No. of Occurrences	10	
2					
3	No. of Arrivals (*x*)	Probability *f(x)*			
4	0	=POISSON.DIST(A4,D1,FALSE)			
5	1	=POISSON.DIST(A5,D1,FALSE)			
6	2	=POISSON.DIST(A6,D1,FALSE)			
7	3	=POISSON.DIST(A7,D1,FALSE)			
8	4	=POISSON.DIST(A8,D1,FALSE)			
9	5	=POISSON.DIST(A9,D1,FALSE)			
10	6	=POISSON.DIST(A10,D1,FALSE)			
11	7	=POISSON.DIST(A11,D1,FALSE)			
12	8	=POISSON.DIST(A12,D1,FALSE)			
13	9	=POISSON.DIST(A13,D1,FALSE)			
14	10	=POISSON.DIST(A14,D1,FALSE)			
15	11	=POISSON.DIST(A15,D1,FALSE)			
16	12	=POISSON.DIST(A16,D1,FALSE)			
17	13	=POISSON.DIST(A17,D1,FALSE)			
18	14	=POISSON.DIST(A18,D1,FALSE)			
19	15	=POISSON.DIST(A19,D1,FALSE)			
20	16	=POISSON.DIST(A20,D1,FALSE)			
21	17	=POISSON.DIST(A21,D1,FALSE)			
22	18	=POISSON.DIST(A22,D1,FALSE)			
23	19	=POISSON.DIST(A23,D1,FALSE)			
24	20	=POISSON.DIST(A24,D1,FALSE)			
25					

	A	B	C	D
1		Mean No. of Occurences		10
2				
3	No. of Arrivals (x)	Probability *f(x)*		
4	0	0.0000		
5	1	0.0005		
6	2	0.0023		
7	3	0.0076		
8	4	0.0189		
9	5	0.0378		
10	6	0.0631		
11	7	0.0901		
12	8	0.1126		
13	9	0.1251		
14	10	0.1251		
15	11	0.1137		
16	12	0.0948		
17	13	0.0729		
18	14	0.0521		
19	15	0.0347		
20	16	0.0217		
21	17	0.0128		
22	18	0.0071		
23	19	0.0037		
24	20	0.0019		
25				



Using Excel to Compute Cumulative Poisson Probabilities

- Excel Formula and Values Worksheets

	A	B	C	D	E
1			Mean No. of Occurrences	10	
2					
3	No. of Arrivals (x)	Probability $f(x)$			
4	0	=POISSON.DIST(A4,D1,TRUE)			
5	1	=POISSON.DIST(A5,D1,TRUE)			
6	2	=POISSON.DIST(A6,D1,TRUE)			
7	3	=POISSON.DIST(A7,D1,TRUE)			
8	4	=POISSON.DIST(A8,D1,TRUE)			
9	5	=POISSON.DIST(A9,D1,TRUE)			
10	6	=POISSON.DIST(A10,D1,TRUE)			
11	7	=POISSON.DIST(A11,D1,TRUE)			
12	8	=POISSON.DIST(A12,D1,TRUE)			
13	9	=POISSON.DIST(A13,D1,TRUE)			
14	10	=POISSON.DIST(A14,D1,TRUE)			
15	11	=POISSON.DIST(A15,D1,TRUE)			
16	12	=POISSON.DIST(A16,D1,TRUE)			
17	13	=POISSON.DIST(A17,D1,TRUE)			
18	14	=POISSON.DIST(A18,D1,TRUE)			
19	15	=POISSON.DIST(A19,D1,TRUE)			
20	16	=POISSON.DIST(A20,D1,TRUE)			
21	17	=POISSON.DIST(A21,D1,TRUE)			
22	18	=POISSON.DIST(A22,D1,TRUE)			
23	19	=POISSON.DIST(A23,D1,TRUE)			
24	20	=POISSON.DIST(A24,D1,TRUE)			
25					

	A	B	C	D	E
1		Mean No. of Occurrences		10	
2					
3	No. of Arrivals (x)	Probability $f(x)$			
4	0	0.0000			
5	1	0.0005			
6	2	0.0028			
7	3	0.0103			
8	4	0.0293			
9	5	0.0671			
10	6	0.1301			
11	7	0.2202			
12	8	0.3328			
13	9	0.4579			
14	10	0.5830			
15	11	0.6968			
16	12	0.7916			
17	13	0.8645			
18	14	0.9165			
19	15	0.9513			
20	16	0.9730			
21	17	0.9857			
22	18	0.9928			
23	19	0.9965			
24	20	0.9984			
25					

CENGAGE

Poisson Probability Distribution (1 of 2)

A property of the Poisson distribution is that the **mean and variance** are **equal**.

$$\mu = \sigma^2$$

Poisson Probability Distribution (2 of 2)

Example: Arrivals at the emergency room

Variance for the number of patients arriving at the emergency room at a large hospital in a 15-minute period of time on weekday mornings is

$$\mu = \sigma^2 = 10$$

Hypergeometric Probability Distribution (1 of 10)

The hypergeometric distribution is closely related to the binomial distribution.

However, for the hypergeometric distribution

- the trials are not independent, and
- the probability of success changes from trial to trial.

Hypergeometric Probability Distribution (2 of 10)

Hypergeometric Probability Function

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}}$$

where:

x = number of successes
n = number of trials
$f(x)$ = probability of x successes in n trials
N = number of elements in the population
r = number of elements in the population labeled for success

Hypergeometric Probability Distribution (3 of 10)

- Hypergeometric Probability Function

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} \quad \text{for } 0 \le x \le r$$

$\binom{r}{x}$: number of ways x successes can be selected from a total of r successes in the population

$\binom{N}{n}$: number of ways n elements can be selected from a population of size N

$\binom{N-r}{n-x}$: number of ways $n - x$ failures can be selected from a total of $N - r$ failures in the population

Hypergeometric Probability Distribution (4 of 10)

Hypergeometric Probability Function

- The probability function $f(x)$ on the previous slide is usually applicable for values of $x = 0, 1, 2, \dots n$.
- However, only following values of x are valid:

  1) $x \le r$ and

  2) $n - x \le N - r$

- If these two conditions do not hold for a value of x, the corresponding $f(x)$ equals 0.

Hypergeometric Probability Distribution (5 of 10)

Example: Ontario Electric

Electric fuses produced by Ontario Electric are packaged in boxes of 12 each. Suppose an inspector randomly selects 3 of the 12 fuses in a box for testing. If the box contains 5 defective fuses, what is the probability that the inspector will find exactly one of the three fuses defective?

Hypergeometric Probability Distribution (6 of 10)

<u>Example:</u> Ontario Electric

Using the probability function:

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{\binom{5}{1}\binom{7}{2}}{\binom{12}{3}} = \frac{\left(\frac{5!}{1!4!}\right)\left(\frac{7!}{2!5!}\right)}{\left(\frac{12!}{3!9!}\right)} = \frac{(5)(21)}{220} = .4773$$

where:

x = 1 = number of <u>defective</u> fuse selected
n = 3 = number of fuses selected
N = 12 = number of fuses in total
r = 5 = number of <u>defective</u> fuses in total

Hypergeometric Probability Distribution (7 of 10)

Mean

$$E(x) = \mu = n\left(\frac{r}{N}\right)$$

Variance

$$Var(x) = \sigma^2 = n\left(\frac{r}{N}\right)\left(1 - \frac{r}{N}\right)\left(\frac{N-n}{N-1}\right)$$

Hypergeometric Probability Distribution (8 of 10)

Example: Ontario Electric

Mean

$$\mu = n\left(\frac{r}{N}\right) = 3\left(\frac{5}{12}\right) = 1.25$$

Variance

$$\sigma^2 = 3\left(\frac{5}{12}\right)\left(1 - \frac{5}{12}\right)\left(\frac{12-3}{12-1}\right) = .60$$

Standard deviation

$$\sigma = .77$$

Hypergeometric Probability Distribution (9 of 10)

- Consider a hypergeometric distribution with *n* trials and let $p = (r/n)$ denote the probability of a success on the first trial.
- If the population size is large, the term $(N - n)/(N - 1)$ approaches 1.
- The expected value and variance can be written as
  - $E(x) = np$
  - $Var(x) = np(1 - p)$
- Note that these are the expressions for the expected value and variance of a binomial distribution.

Hypergeometric Probability Distribution (10 of 10)

- When the population size is large, a hypergeometric distribution can be approximated by a binomial distribution with n trials and a probability of success $p = (r/N)$.