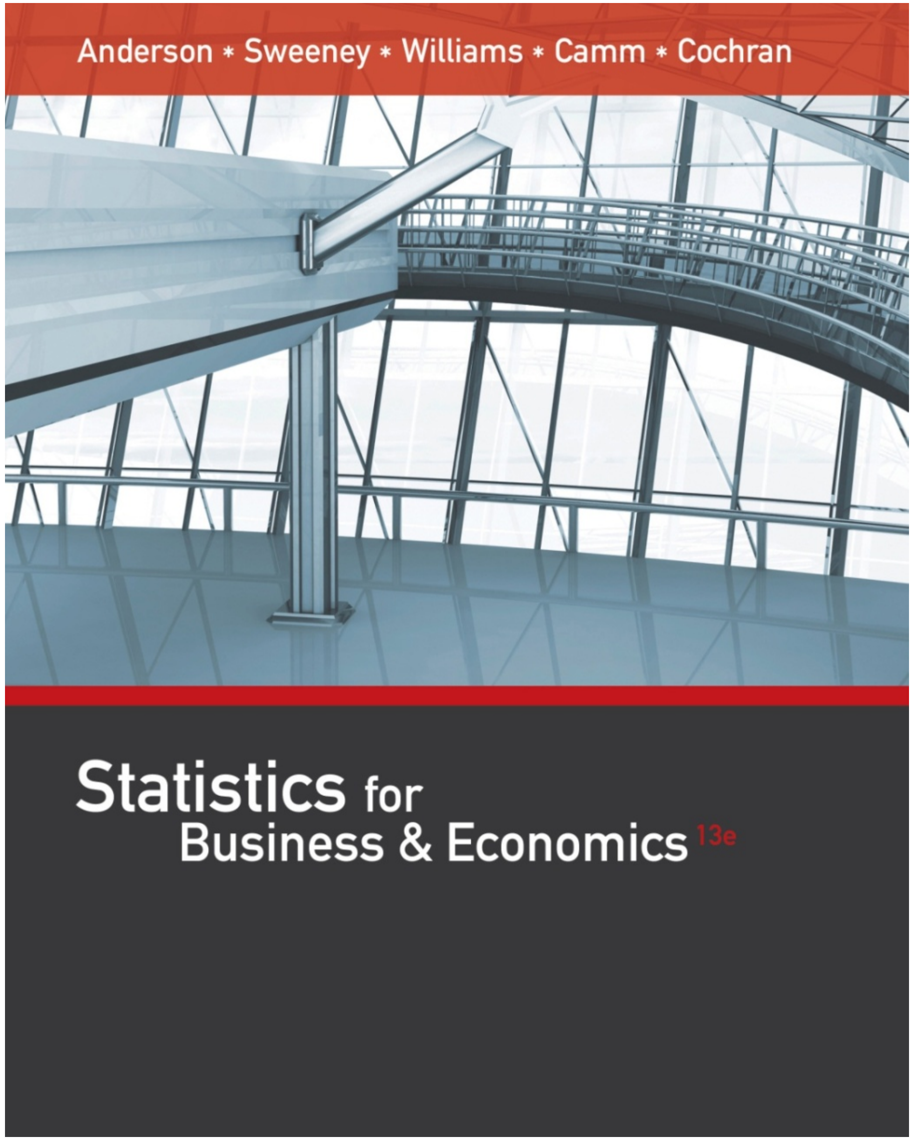
Statistics for Business and Economics (13e)

Anderson, Sweeney, Williams, Camm, Cochran

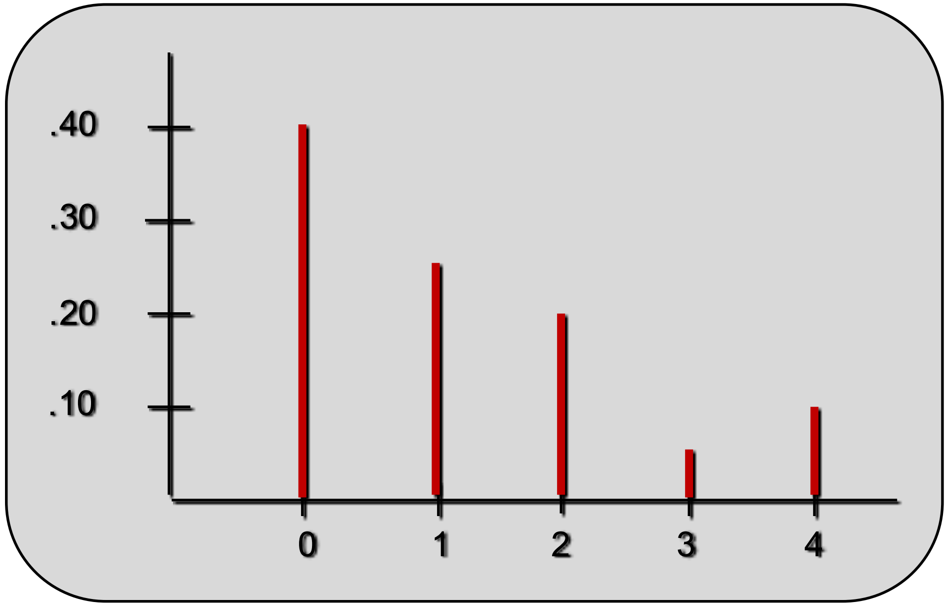
Slides by John Loucks

St. Edwards University



Chapter 5: Discrete Probability Distributions

- Random Variables
- Developing Discrete Probability Distributions
- Expected Value and Variance
- Binomial Probability Distribution
- Poisson Probability Distribution
- Hypergeometric Probability Distribution



Random Variables

- 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

- Example: JSL Appliances

Let x = number of TVs sold at the store in one day,

where x can take on 5 values (0, 1, 2, 3, 4)

We can count the TVs sold, and there is a finite upper limit on the number that might be sold (which is the number of TVs in stock).

Discrete Random Variable with a Finite Number of Values

- Example: JSL Appliances

Let x = number of customers 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

Illustration	Random Variable x	Type
Family size	x = Number of dependents reported on tax return	Discrete
Distance from home to stores on a highway	x = Distance in miles from home to the store site	Continuous
Own dog or cat	x = 1 if own no pet; = 2 if own dog(s) only; = 3 if own cat(s) only; = 4 if own dog(s) and cat(s)	Discrete

Discrete Probability Distributions

- 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

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

- 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:

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

Discrete Probability Distributions

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

Discrete Probability Distributions

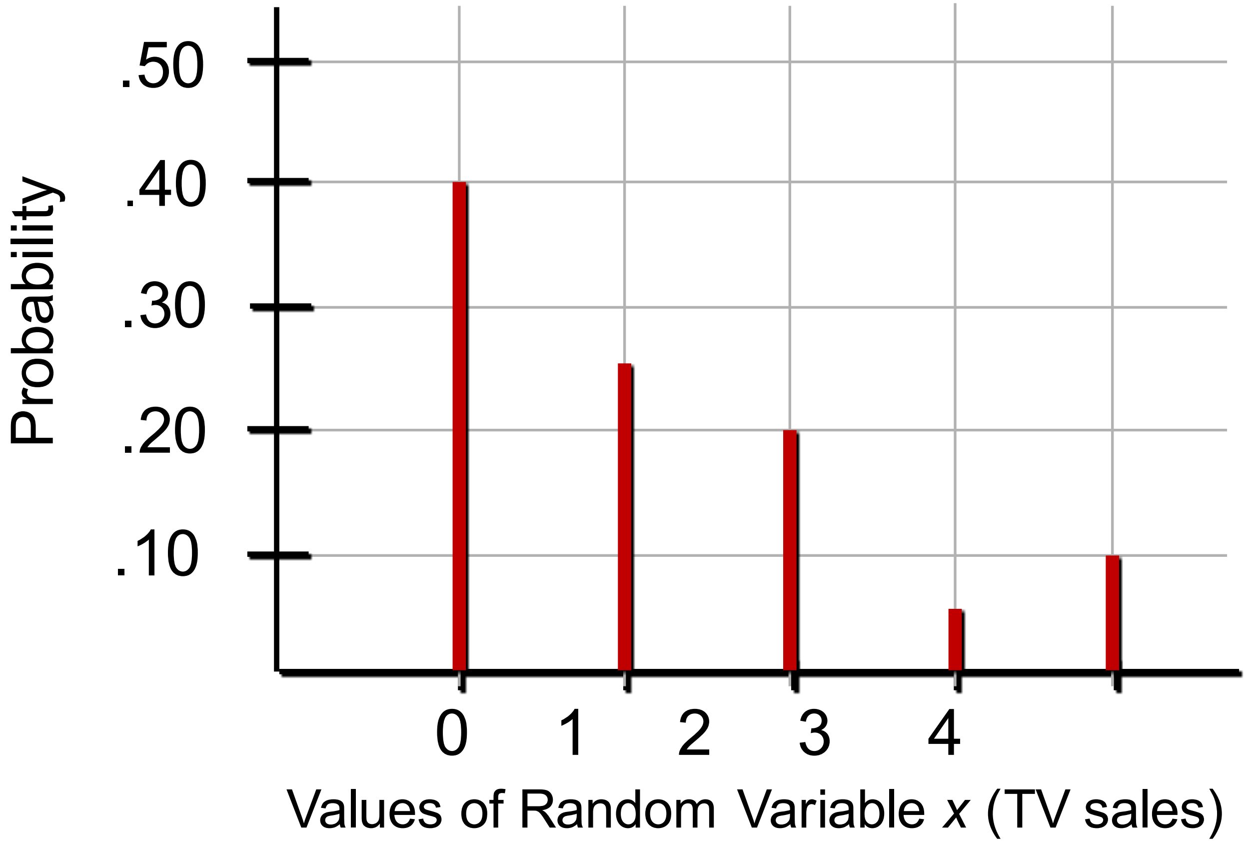
- Example: JSL Appliances

Using past data on TV sales, a tabular representation of the probability distribution for sales was developed.

Units Sold	Number of Days	x	$f(x)$
0	80	0	.40 = 80/200
1	50	1	.25
2	40	2	.20
3	10	3	.05
4	20	4	.10
	200		1.00

Discrete Probability Distributions

- Example: JSL Appliances



Graphical representation of probability distribution

Discrete Probability Distributions

- 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.
- Several discrete probability distributions specified by formulas are the discrete-uniform, binomial, Poisson, and hypergeometric distributions.

Discrete Probability Distributions

- 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.

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

- 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.

Expected Value

- Example: JSL Appliances

x	$f(x)$	$xf(x)$
0	.40	.00
1	.25	.25
2	.20	.40
3	.05	.15
4	.10	.40

$E(x) =$ 1.20 = expected number of TVs sold in a day

Variance

- Example: JSL Appliances

x	$x - \mu$	$(x - \mu)^2$	$f(x)$	$(x - \mu)^2 f(x)$
0	-1.2	1.44	.40	.576
1	-0.2	0.04	.25	.010
2	0.8	0.64	.20	.128
3	1.8	3.24	.05	.162
4	2.8	7.84	.10	.784

Variance of daily sales = σ^2 = 1.660

Standard deviation of daily sales = 1.2884 TVs

Binomial Probability Distribution

- 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, does not change from trial to trial. (This is referred to as the stationarity assumption.)
  4. The trials are independent.

Binomial Probability Distribution

- Our interest is in the number of successes occurring in the n trials.
- Let x denote the number of successes occurring in the n trials.

Binomial Probability Distribution

- 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

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

Binomial Probability Distribution

- Binomial Probability Function

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

Number of experimental outcomes providing exactly *x* successes in *n* trials

Probability of a particular sequence of trial outcomes with *x* successes in *n* trials

Binomial Probability Distribution

- Example: Evans Electronics

Evans Electronics is concerned about a low retention rate for its employees. In recent years, management has seen a turnover of 10% of the hourly employees annually.

Thus, for any hourly employee chosen at random, management estimates a probability of 0.1 that the person will not be with the company next year.

Choosing 3 hourly employees at random, what is the probability that 1 of them will leave the company this year?

Binomial Probability Distribution

- Example: Evans Electronics
  - The probability of the first employee leaving and the second and third employees staying, denoted (S, F, F), is given by

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

  - With a .10 probability of an employee leaving on any one trial, the probability of an employee leaving on the first trial and not on the second and third trials is given by

$$(.10)(.90)(.90) = (.10)(.90)^2 = .081$$

Binomial Probability Distribution

- Example: Evans Electronics
  - Two other experimental outcomes result in one success and two failures. The probabilities for all three experimental outcomes involving one success follow.

Experimental Outcome	Probability of Experimental Outcome
(S, F, F)	$p(1 - p)(1 - p) = (.1)(.9)(.9) = .081$
(F, S, F)	$(1 - p)p(1 - p) = (.9)(.1)(.9) = .081$
(F, F, S)	$(1 - p)(1 - p)p = (.9)(.9)(.1) = \underline{.081}$
	Total = .243

Binomial Probability Distribution

- Example: Evans Electronics

Using the probability function:

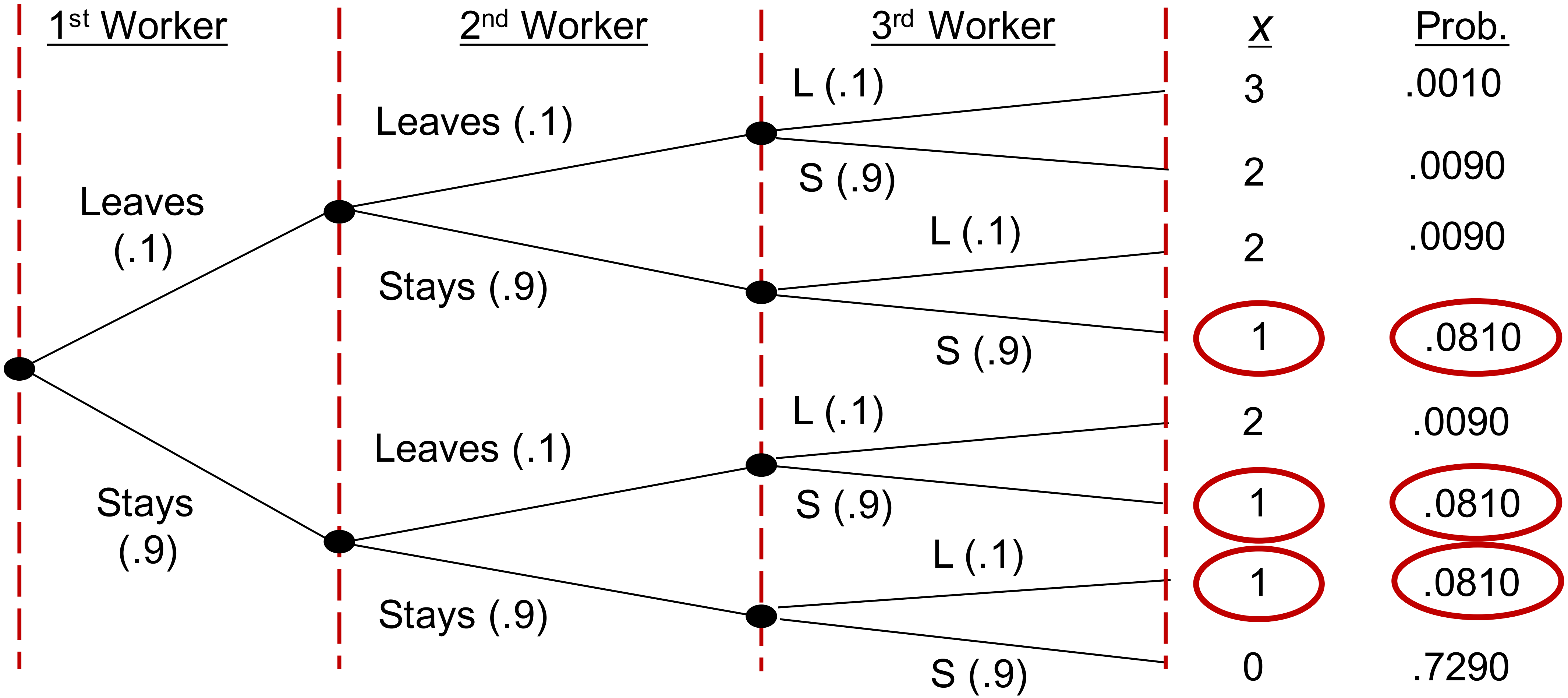
Let: $p = .10$, $n = 3$, $x = 1$

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

$$f(1) = \frac{3!}{1!(3-1)!}(0.1)^1(0.9)^2 = .243$$

Binomial Probability Distribution

- Example: Evans Electronics



Binomial Probabilities and Cumulative Probabilities

- Statisticians have developed tables that give probabilities and cumulative probabilities for a binomial experiment 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.

Binomial Probability Distribution

- Using Tables of Binomial Probabilities

		p									
n	***x***	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
3	0	.8574	.7290	.6141	.5120	.4219	.3430	.2746	.2160	.1664	.1250
	1	.1354	.2430	.3251	.3840	.4219	.4410	.4436	.4320	.4084	.3750
	2	.0071	.0270	.0574	.0960	.1406	.1890	.2389	.2880	.3341	.3750
	3	.0001	.0010	.0034	.0080	.0156	.0270	.0429	.0640	.0911	.1250

Binomial Probability Distribution

- Expected Value

$$E(x) = \mu = np$$

- Variance

$$Var(x) = \sigma^2 = np(1 - p)$$

- Standard Deviation

$$\sigma = \sqrt{np(1-p)}$$

Binomial Probability Distribution

- Example: Evans Electronics
  - Expected Value

$$E(x) = np = 3(.1) = .3 \text{ employees out of } 3$$

  - Variance

$$Var(x) = np(1 - p) = 3(.1)(.9) = .27$$

  - Standard Deviation

$$\sigma = \sqrt{3(.1).9)} = .52 \text{ employees}$$

Poisson Probability Distribution

- 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 ($x = 0, 1, 2, \ldots$).

Poisson Probability Distribution

- Examples of Poisson distributed random variables:
  - number of knotholes in 14 linear feet of pine board
  - number of vehicles arriving at a toll booth in one hour

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

Poisson Probability Distribution

- Two 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

- 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

μ = mean number of occurrences in an interval

e = 2.71828

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

Poisson Probability Distribution

- Poisson Probability Function–
  - Since there is no stated upper limit for the number of occurrences, the probability function $f(x)$ is applicable for values $x = 0, 1, 2, \dots$ 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

- Example: Mercy Hospital

Patients arrive at the emergency room of Mercy Hospital at the average rate of 6 per hour on weekend evenings.

What is the probability of 4 arrivals in 30 minutes on a weekend evening?

Poisson Probability Distribution

- Example: Mercy Hospital

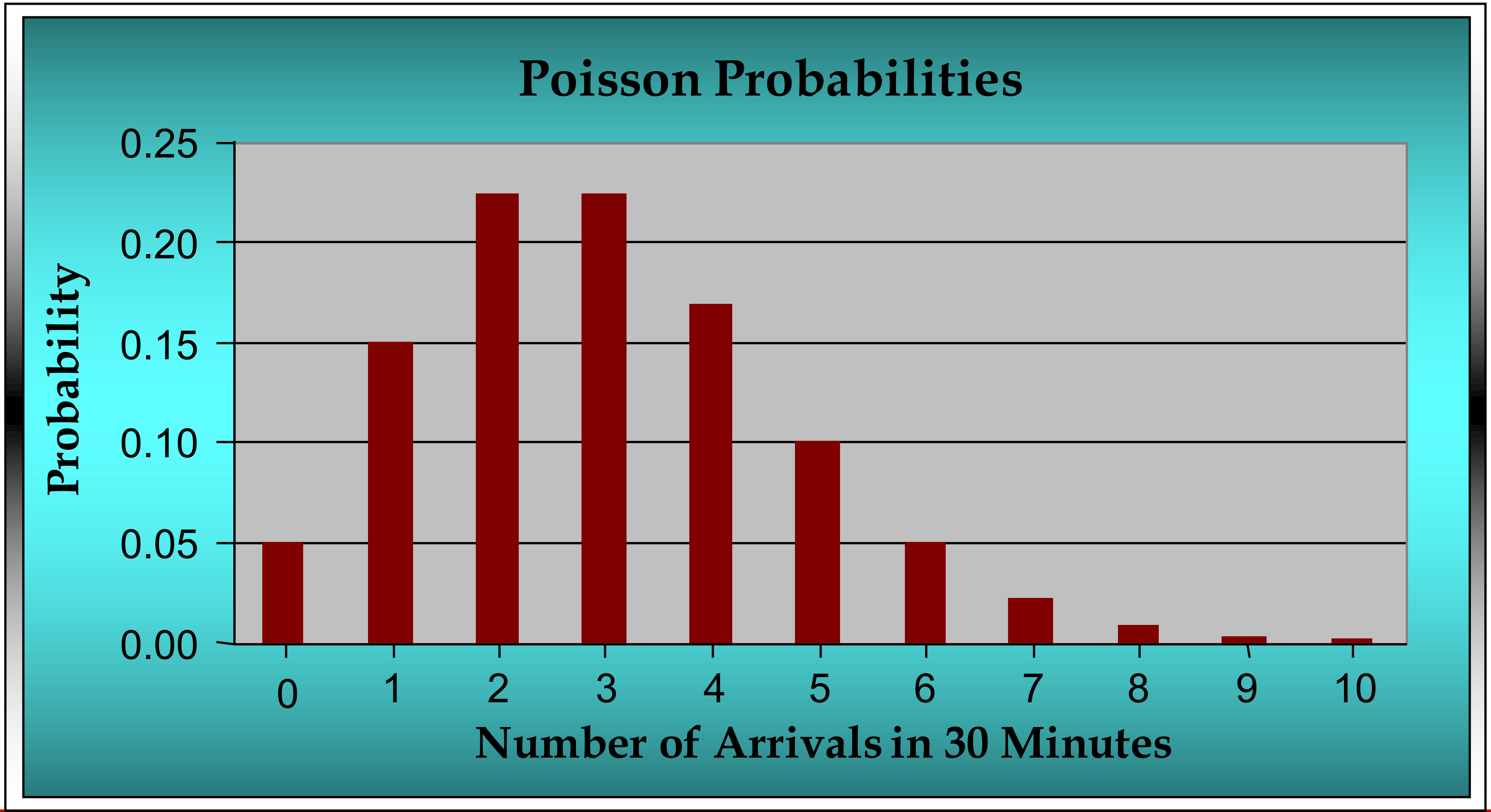
Using the probability function:

μ = 6/hour = 3/half-hour, x = 4

$$f(4) = \frac{3^4 (2.71828)^{-3}}{4!} = .1680$$

Poisson Probability Distribution

- Example: Mercy Hospital



Poisson Probability Distribution

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

$$\mu = \sigma^2$$

- Example: Mercy Hospital

Variance for Number of Arrivals during 30-Minute periods

$$\mu = \sigma^2 = 3$$

Hypergeometric Probability Distribution

- 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

- 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 success

Hypergeometric Probability Distribution

- Hypergeometric Probability Function

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

number of ways x successes can be selected from a total of r successes in the population

number of ways $n - x$ failures can be selected from a total of $N - r$ failures in the population

number of ways n elements can be selected from a population of size N

Hypergeometric Probability Distribution

- Hypergeometric Probability Function
  - The probability function $f(x)$ on the is usually applicable for values of $x = 0, 1, 2, \ldots n$.
  - However, only values of x where: 1) $x \leq r$ and 2) $n - x \leq N - r$ are valid.
  - If these two conditions do not hold for a value of x, the corresponding $f(x)$ equals 0.

Hypergeometric Probability Distribution

- Example: Neveready's Batteries

Bob Neveready has removed two dead batteries from a flashlight and inadvertently mingled them with the two good batteries he intended as replacements. The four batteries look identical.

Bob now randomly selects two of the four batteries. What is the probability he selects the two good batteries?

Hypergeometric Probability Distribution

- Example: Neveready's Batteries

Using the probability function:

$$f(x) = \frac{\binom{r}{x}\binom{N-r}{n-x}}{\binom{N}{n}} = \frac{\binom{2}{2}\binom{2}{0}}{\binom{4}{2}} = \frac{\left(\frac{2!}{2!0!}\right)\left(\frac{2!}{0!2!}\right)}{\left(\frac{4!}{2!2!}\right)} = \frac{1}{6} = .167$$

where: *x* = 2 = number of <u>good</u> batteries selected
n = 2 = number of batteries selected
N = 4 = number of batteries in total
r = 2 = number of <u>good</u> batteries in total

Hypergeometric Probability Distribution

- 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

- Example: Neveready's Batteries
  - Mean

$$\mu = n\left(\frac{r}{N}\right) = 2\left(\frac{2}{4}\right) = 1$$

  - Variance

$$\sigma^2 = 2\left(\frac{2}{4}\right)\left(1-\frac{2}{4}\right)\left(\frac{4-2}{4-1}\right) = \frac{1}{3} = .333$$

Hypergeometric Probability Distribution

- 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 $E(x) = np$ and $Var(x) = np(1 - p)$.
- Note that these are the expressions for the expected value and variance of a binomial distribution.

continued →

Hypergeometric Probability Distribution

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

End of Chapter 5