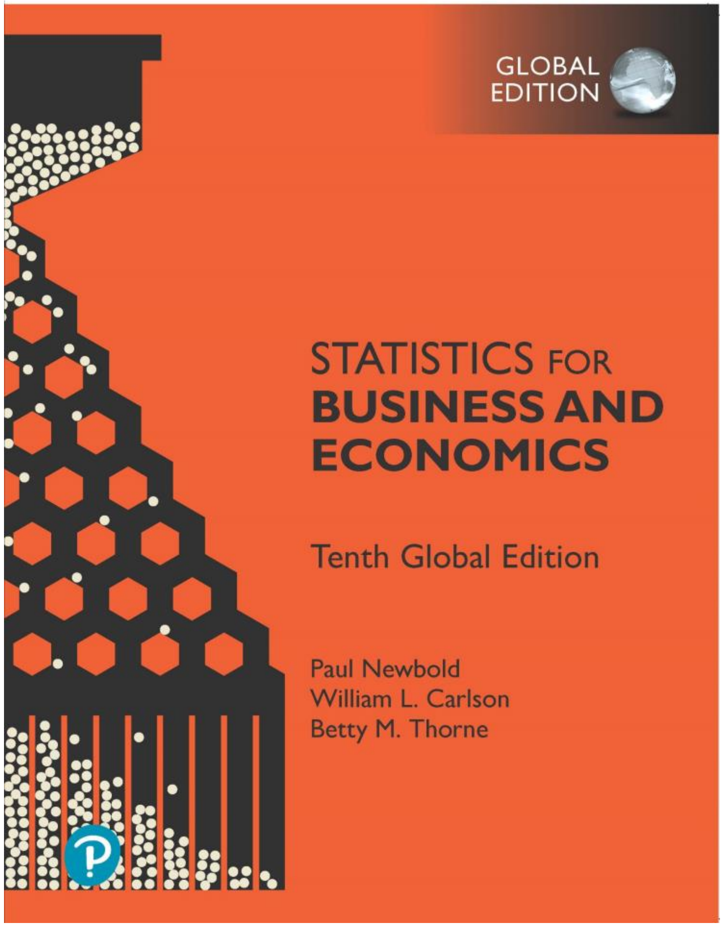
Statistics for Business and Economics

Tenth Edition, Global Edition



Chapter 4

Discrete Random Variables and Probability Distributions

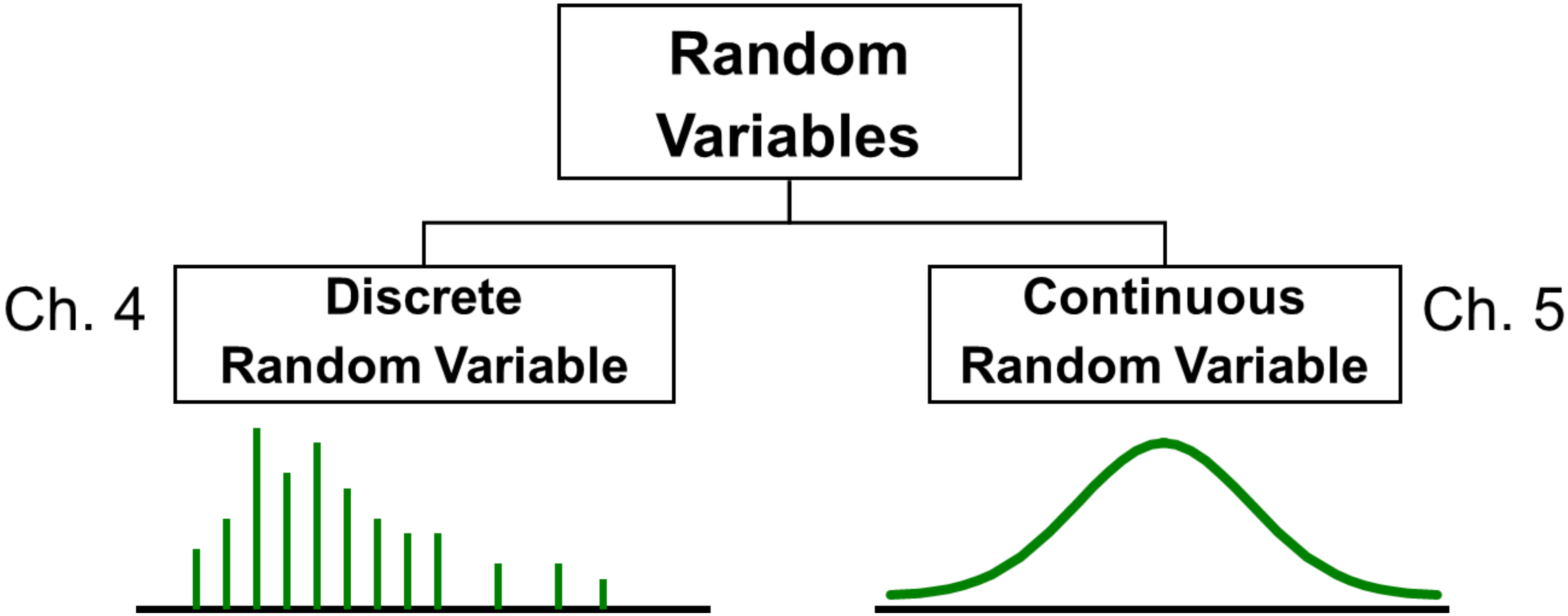
Chapter Goals

After completing this chapter, you should be able to:

- Interpret the mean and standard deviation for a discrete random variable
- Use the binomial probability distribution to find probabilities
- Describe when to apply the binomial distribution
- Use the hypergeometric and Poisson discrete probability distributions to find probabilities
- Explain covariance and correlation for jointly distributed discrete random variables
- Explain an application to portfolio investment

Section 4.1 Random Variables

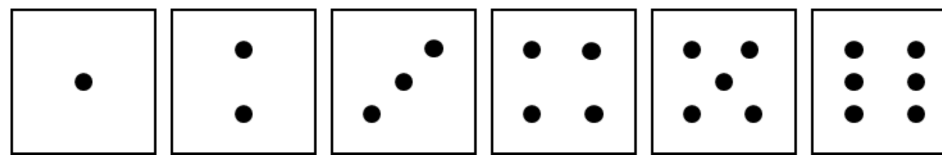
- **Random Variable**
  - Represents a possible numerical value from a random experiment



Discrete Random Variable

- Takes on no more than a countable number of values

Examples:



- **Roll a die twice**

  Let *X* be the number of times 4 comes up (then *X* could be 0, 1, or 2 times)

- **Toss a coin 5 times.**

  Let *X* be the number of heads (then *X* = 0, 1, 2, 3, 4, or 5)

Continuous Random Variable

- Can take on any value in an interval
  - Possible values are measured on a continuum

Examples:

- **Weight of packages filled by a mechanical filling process**
- **Temperature of a cleaning solution**
- **Time between failures of an electrical component**

Section 4.2 Probability Distributions for Discrete Random Variables (1 of 2)

Let X be a discrete random variable and x be one of its possible values

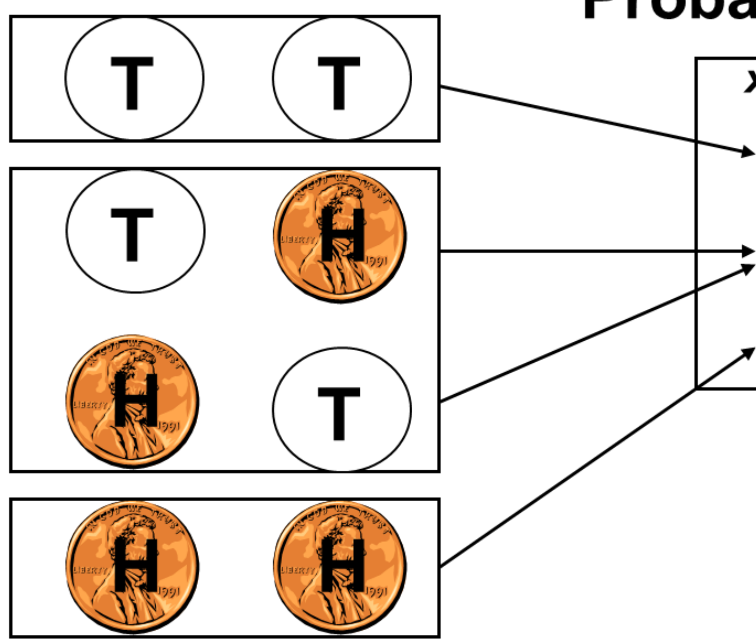
- The probability that random variable X takes specific value x is denoted $P(X = x)$
- The probability distribution function of a random variable is a representation of the probabilities for all the possible outcomes.
  - Can be shown algebraically, graphically, or with a table

Section 4.2 Probability Distributions for Discrete Random Variables (2 of 2)

Experiment: Toss 2 Coins. Let X = # heads.

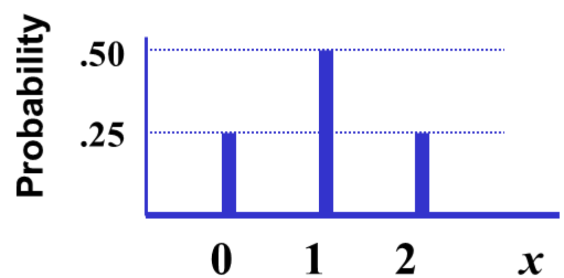
Show $P(x)$, i.e., $P(X = x)$, for all values of x:

4 possible outcomes



Probability Distribution

x Value	Probability
0	$\frac{1}{4} = .25$
1	$\frac{2}{4} = .50$
2	$\frac{1}{4} = .25$



Probability Distribution Required Properties

- $0 \le P(x) \le 1$ for any value of *x*

- The individual probabilities sum to 1;

$$\sum_{x} P(x) = 1$$

(The notation indicates summation over all possible *x* values)

Cumulative Probability Function (1 of 2)

- The cumulative probability function, denoted $F(x_0)$, shows the probability that *X* does not exceed the value x_0

$$F\left(x_0\right) = P\left(X \le x_0\right)$$

Where the function is evaluated at all values of x_0

Cumulative Probability Function (2 of 2)

Example: Toss 2 Coins. Let X = # heads.

x Value	$P(x)$	$F(x)$
0	0.25	0.25
1	0.50	0.75
2	0.25	1.00

Derived Relationship

The derived relationship between the probability distribution and the cumulative probability distribution

- Let *X* be a random variable with probability distribution $P(x)$ and cumulative probability distribution $F(x_0)$. Then

$$F(x_0) = \sum_{x \le x_0} P(x)$$

(the notation implies that summation is over all possible values of *x* that are less than or equal to x_0)

Derived Properties

Derived properties of cumulative probability distributions for discrete random variables

- Let X be a discrete random variable with cumulative probability distribution $F(x_0)$. Then

1. $0 \le F(x_0) \le 1$ for every number x_0
2. For $x_0 < x_1$, then $F(x_0) \le F(x_1)$

Section 4.3 Properties of Discrete Random Variables

- Expected Value (or mean) of a discrete random variable *X*:

$$E[X] = \mu = \sum_x xP(x)$$

- Example: Toss 2 coins,

  x = # of heads,

  compute expected value of *x*:

$$\mathbf{E(x) = (0 \times .25) + (1 \times .50) + (2 \times .25)}$$
$$\mathbf{= 1.0}$$

x	$P(x)$
0	.25
1	.50
2	.25

Variance and Standard Deviation

- Variance of a discrete random variable *X*

$$\sigma^2 = E\left[(X-\mu)^2\right] = \sum_x (x-\mu)^2 P(x)$$

Can also be expressed as

$$\sigma^2 = E\left[X^2\right] - \mu^2 = \sum_x x^2 P(x) - \mu^2$$

- Standard Deviation of a discrete random variable *X*

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_x (x-\mu)^2 P(x)}$$

Standard Deviation Example

- Example: Toss 2 coins, *X* = # heads, compute standard deviation (recall $E[X]=1$)

$$\sigma = \sqrt{\sum_x (x-\mu)^2 P(x)}$$

$$\sigma = \sqrt{(0-1)^2(.25)+(1-1)^2(.50)+(2-1)^2(.25)} = \sqrt{.50} = .707$$

Possible number of heads = 0, 1, or 2

Functions of Random Variables

- If $P(x)$ is the probability function of a discrete random variable *X*, and $g(X)$ is some function of *X*, then the expected value of function *g* is

$$E\left[g(X)\right] = \sum_{x} g(x)P(x)$$

Linear Functions of Random Variables

- Let random variable *X* have mean μ_x and variance $\sigma^2{}_x$
- Let *a* and *b* be any constants.
- Let $Y = a + bX$
- Then the mean and variance of *Y* are

$$\mu_Y = E(a + bX) = a + b\mu_x$$

$$\sigma^2{}_Y = Var(a + bX) = b^2\sigma^2 x$$

- so that the standard deviation of *Y* is

$$\sigma_Y = |b|\sigma_x$$

Properties of Linear Functions of Random Variables

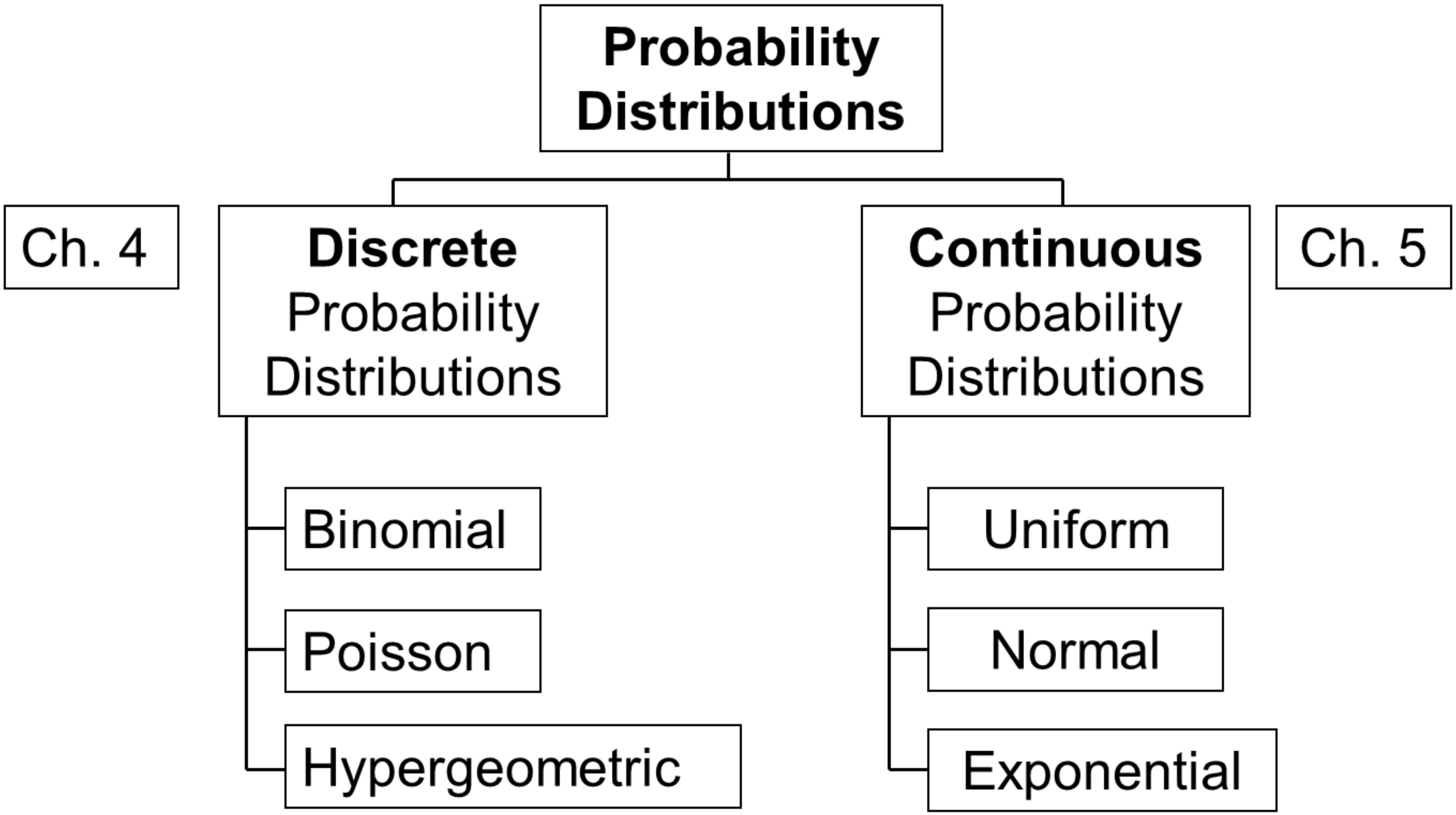
- Let *a* and *b* be any constants.
- a) $E(a) = a$ and $Var(a) = 0$

  i.e., if a random variable always takes the value *a*, it will have mean *a* and variance 0

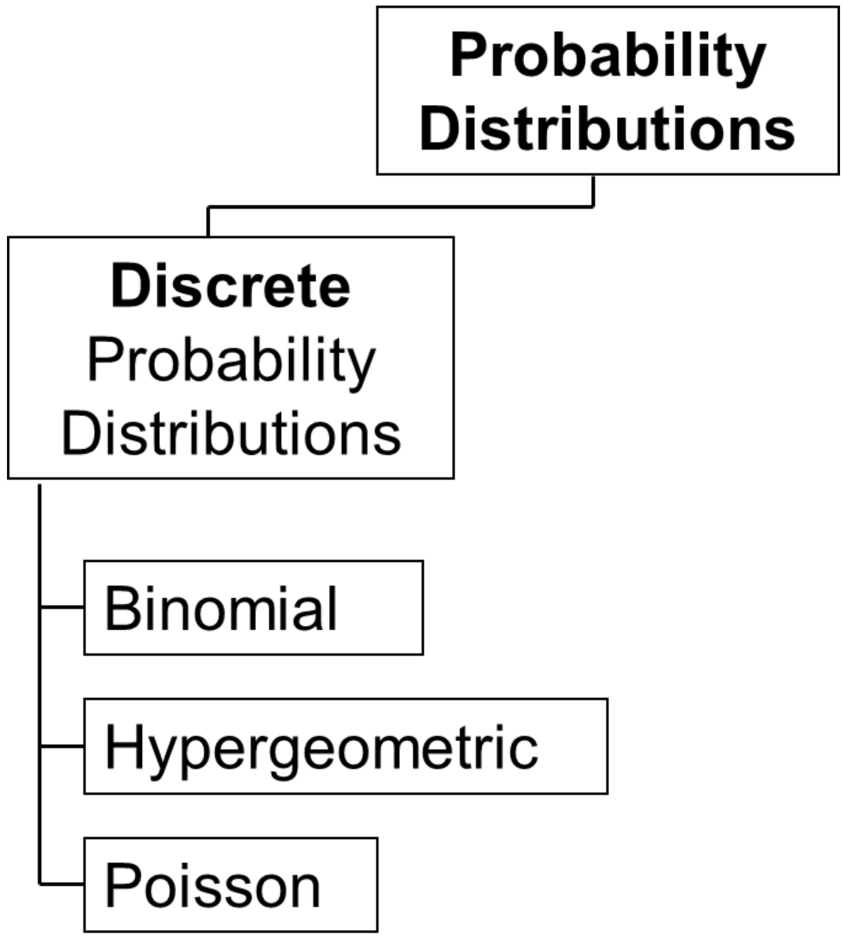
- b) $E(bX) = b\mu_x$ and $Var(bX) = b^2\sigma^2{}_x$

  i.e., the expected value of $b \cdot X$ is $b \cdot E(x)$

Probability Distributions



Section 4.4 The Binomial Distribution



Bernoulli Distribution

- Consider only two outcomes: “success” or “failure”
- Let P denote the probability of success
- Let $1 - P$ be the probability of failure
- Define random variable X:

  $x = 1$ if success, $x = 0$ if failure

- Then the Bernoulli probability distribution is

$$P(0) = (1 - P) \quad \text{and} \quad P(1) = P$$

Mean and Variance of a Bernoulli Random Variable

- The mean is $\mu_x = P$

$$\mu_x = E[X] = \sum_x xP(x) = (0)(1-P) + (1)P = P$$

- The variance is $\sigma_x^2 = P(1-P)$

$$\sigma_x^2 = E\left[(X-\mu_x)^2\right] = \sum_x (x-\mu_x)^2 P(x)$$

$$= (0-P)^2(1-P) + (1-P)^2 P = P(1-P)$$

Developing the Binomial Distribution

- The number of sequences with x successes in n independent trials is:

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

Where $n! = n \cdot (n-1) \cdot (n-2) \cdot \ldots \cdot 1$ and $0! = 1$

- These sequences are mutually exclusive, since no two can occur at the same time

Binomial Probability Distribution

- A fixed number of observations, *n*
  - e.g., 15 tosses of a coin; ten light bulbs taken from a warehouse
- Two mutually exclusive and collectively exhaustive categories
  - e.g., head or tail in each toss of a coin; defective or not defective light bulb
  - Generally called “success” and “failure”
  - Probability of success is P, probability of failure is $1 - P$
- Constant probability for each observation
  - e.g., Probability of getting a tail is the same each time we toss the coin
- Observations are independent
  - The outcome of one observation does not affect the outcome of the other

Possible Binomial Distribution Settings

- A manufacturing plant labels items as either defective or acceptable
- A firm bidding for contracts will either get a contract or not
- A marketing research firm receives survey responses of “yes I will buy” or “no I will not”
- New job applicants either accept the offer or reject it

The Binomial Distribution

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

$P(x) =$ probability of x successes in n trials, with probability of success P on each trial

x = number of 'successes' in sample, $(x = 0,\ 1,\ 2,\ \ldots,\ n)$

n = sample size (number of independent trials or observations)

P = probability of "success"

Example: Flip a coin four times, let x = # heads:

$$n = 4$$

$$P = 0.5$$

$$1 - P = (1 - 0.5) = 0.5$$

$$x = 0,\ 1,\ 2,\ 3,\ 4$$

Example 1: Calculating a Binomial Probability

What is the probability of one success in five observations if the probability of success is 0.1?

x = 1, n = 5, and P = 0.1

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

$$= \frac{5!}{1!(5-1)!} (0.1)^1 (1-0.1)^{5-1}$$

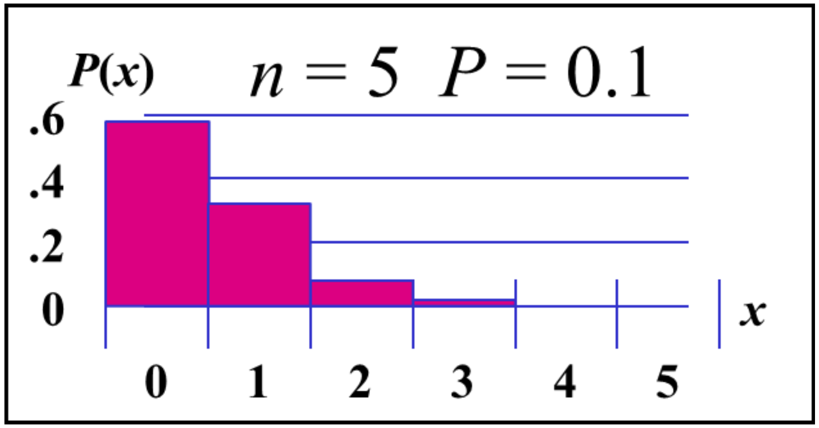
$$= (5)(0.1)(0.9)^4$$

$$= .32805$$

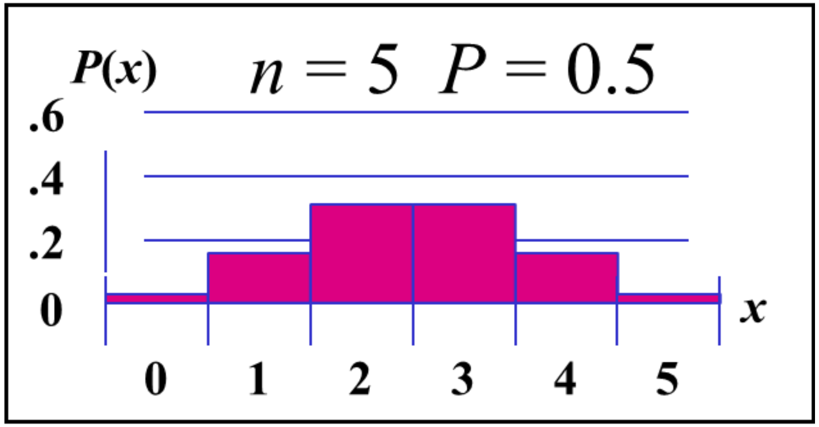
Shape of Binomial Distribution

- The shape of the binomial distribution depends on the values of *P* and *n*

- Here, *n* = 5 and *P* = 0.1



- Here, *n* = 5 and *P* = 0.5



Mean and Variance of a Binomial Distribution

- Mean

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

- Variance and Standard Deviation

$$\sigma^2 = nP(1-P)$$

$$\sigma = \sqrt{nP(1-P)}$$

Where n = sample size

P = probability of success

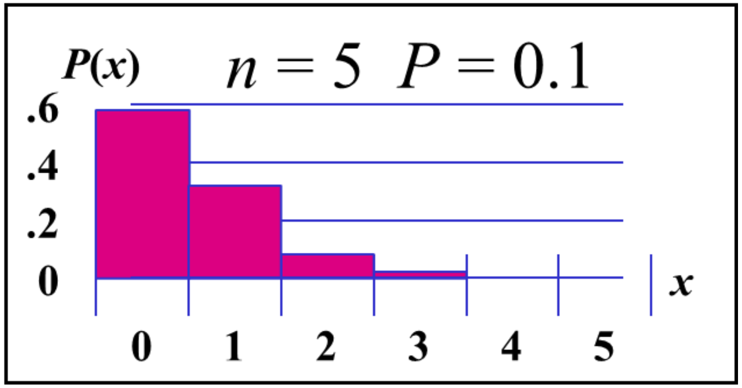
$(1 - P)$ = probability of failure

Binomial Characteristics

Examples

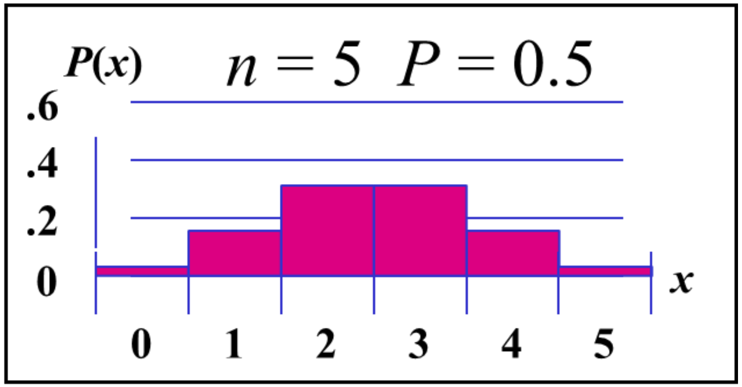
$$\mu = nP = (5)(0.1) = 0.5$$

$$\sigma = \sqrt{nP(1-P)} = \sqrt{(5)(0.1)(1-0.1)} = 0.6708$$



$$\mu = nP = (5)(0.5) = 2.5$$

$$\sigma = \sqrt{nP(1-P)} = \sqrt{(5)(0.5)(1-0.5)} = 1.118$$



Using Binomial Tables

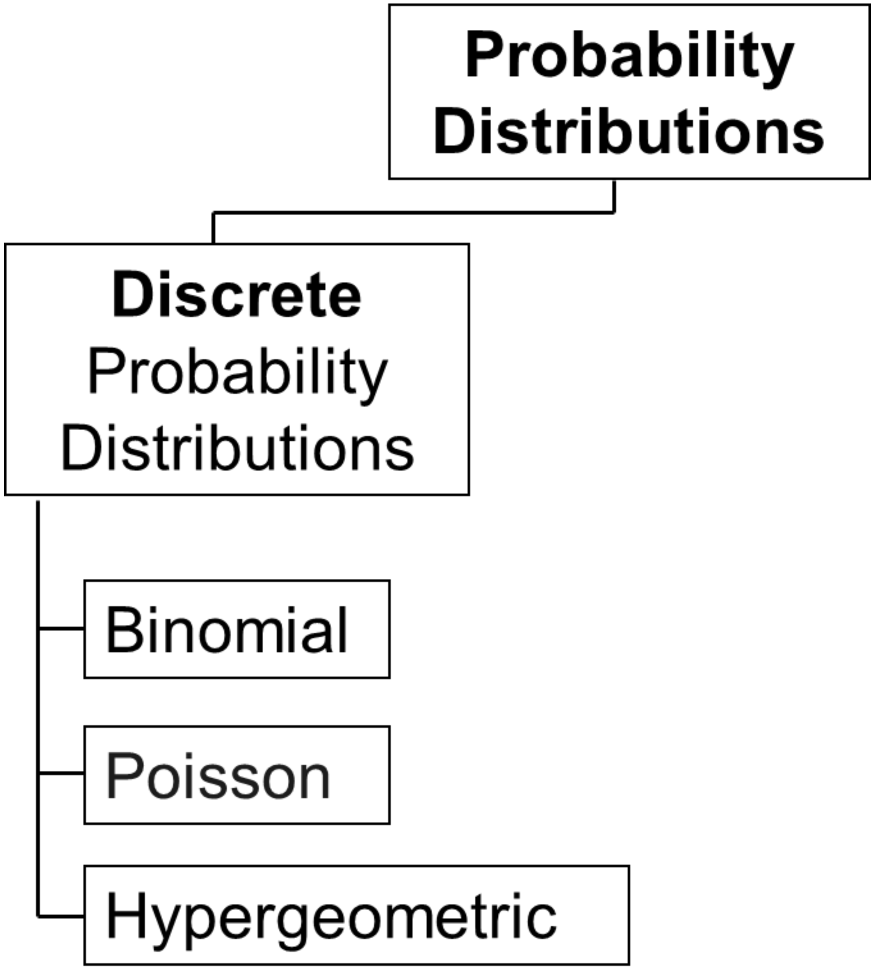
N	x	…	p = .20	p = .25	p = .30	p = .35	p = .40	p = .45	p = .50
10	0	…	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010
	1	…	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098
	2	…	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439
	3	…	0.2013	0.2503	0.2668	**0.2522**	0.2150	0.1665	0.1172
	4	…	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051
	5	…	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461
	6	…	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051
	7	…	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172
	8	…	0.0001	0.0004	0.0014	0.0043	0.0106	**0.0229**	0.0439
	9	…	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098
	10	…	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010

Examples:

$$n = 10,\ x = 3,\ P = 0.35: \qquad P\left(x = 3 \mid n = 10,\ p = 0.35\right) = .2522$$

$$n = 10,\ x = 8,\ P = 0.45: \qquad P\left(x = 8 \mid n = 10,\ p = 0.45\right) = .0229$$

Section 4.5 The Poisson Distribution (1 of 3)



Section 4.5 The Poisson Distribution (2 of 3)

- The Poisson distribution is used to determine the probability of a random variable which characterizes the number of occurrences or successes of a certain event in a given continuous interval (such as time, surface area, or length).

Section 4.5 The Poisson Distribution (3 of 3)

- Assume an interval is divided into a very large number of equal subintervals where the probability of the occurrence of an event in any subinterval is very small.

Poisson distribution assumptions

1. The probability of the occurrence of an event is constant for all subintervals.
2. There can be no more than one occurrence in each subinterval.
3. Occurrences are independent; that is, an occurrence in one interval does not influence the probability of an occurrence in another interval.

Poisson Distribution Function

The expected number of events per unit is the parameter λ (lambda), which is a constant that specifies the average number of occurrences (successes) for a particular time and/or space

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$

where:

$P(x) =$ the probability of *x* successes over a given time or space, given λ

$\lambda =$ the expected number of successes per time or space unit, $\lambda > 0$

e = base of the natural logarithm system $(2.71828\ldots)$

Poisson Distribution Characteristics

Mean and variance of the Poisson distribution

- Mean

$$\mu_x = E[X] = \lambda$$

- Variance and Standard Deviation

$$\sigma_x^2 = E\left[(X - \mu_x)^2\right] = \lambda$$

$$\sigma = \sqrt{\lambda}$$

where $\lambda =$ expected number of successes per time or space unit

Using Poisson Tables

	λ								
X	0.10	0.20	0.30	0.40	**0.50**	0.60	0.70	0.80	0.90
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659
2	0.0045	0.0164	0.0333	0.0536	**0.0758**	0.0988	0.1217	0.1438	0.1647
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
4	0.0000	0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
5	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Example: Find $P(X=2)$ if $\lambda = .50$

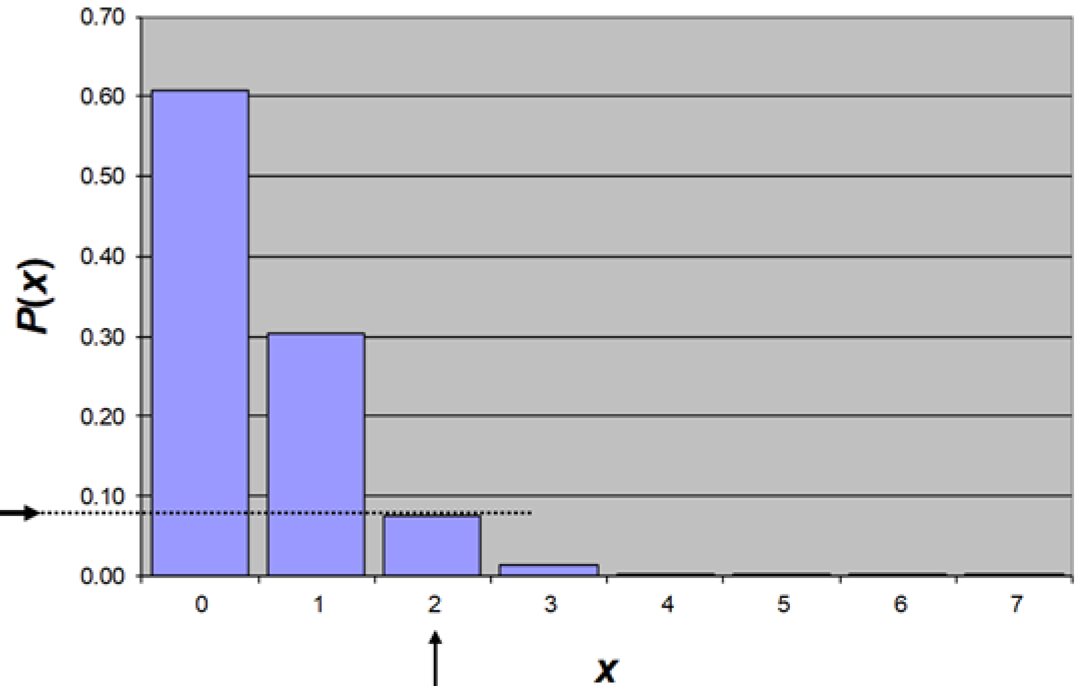
$$P(X=2) = \frac{e^{-\lambda}\lambda^{X}}{X!} = \frac{e^{-0.50}(0.50)^{2}}{2!} = .0758$$

Graph of Poisson Probabilities

Graphically:

$\lambda = .50$

X	$\lambda = 0.50$
0	0.6065
1	0.3033
2	0.0758
3	0.0126
4	0.0016
5	0.0002
6	0.0000
7	0.0000

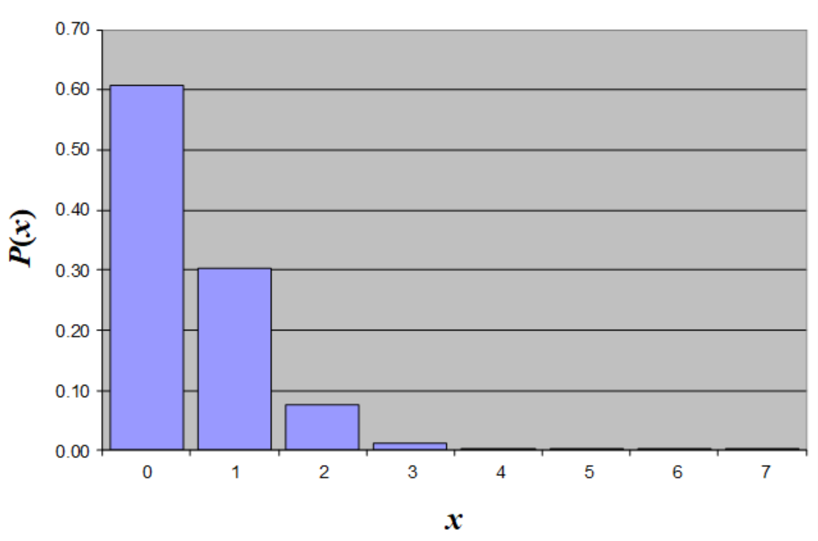


$P(X = 2) = .0758$

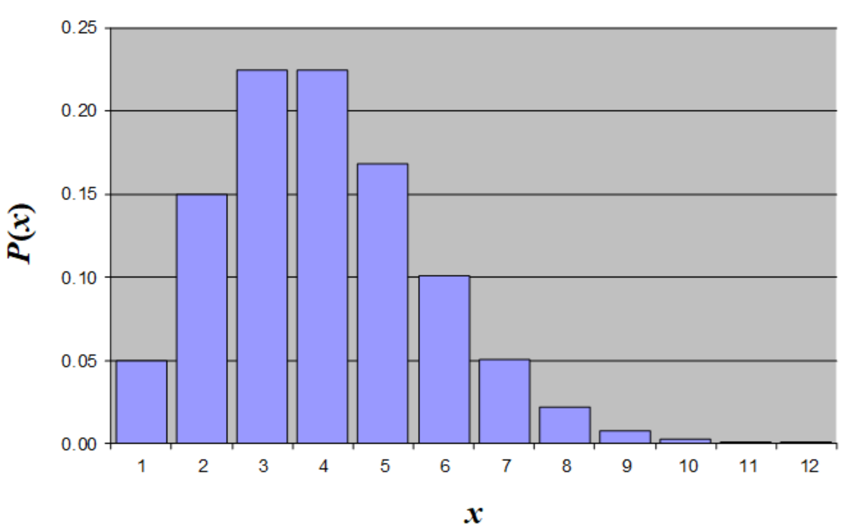
Poisson Distribution Shape

- The shape of the Poisson Distribution depends on the parameter λ :

$\lambda = 0.50$



$\lambda = 3.00$



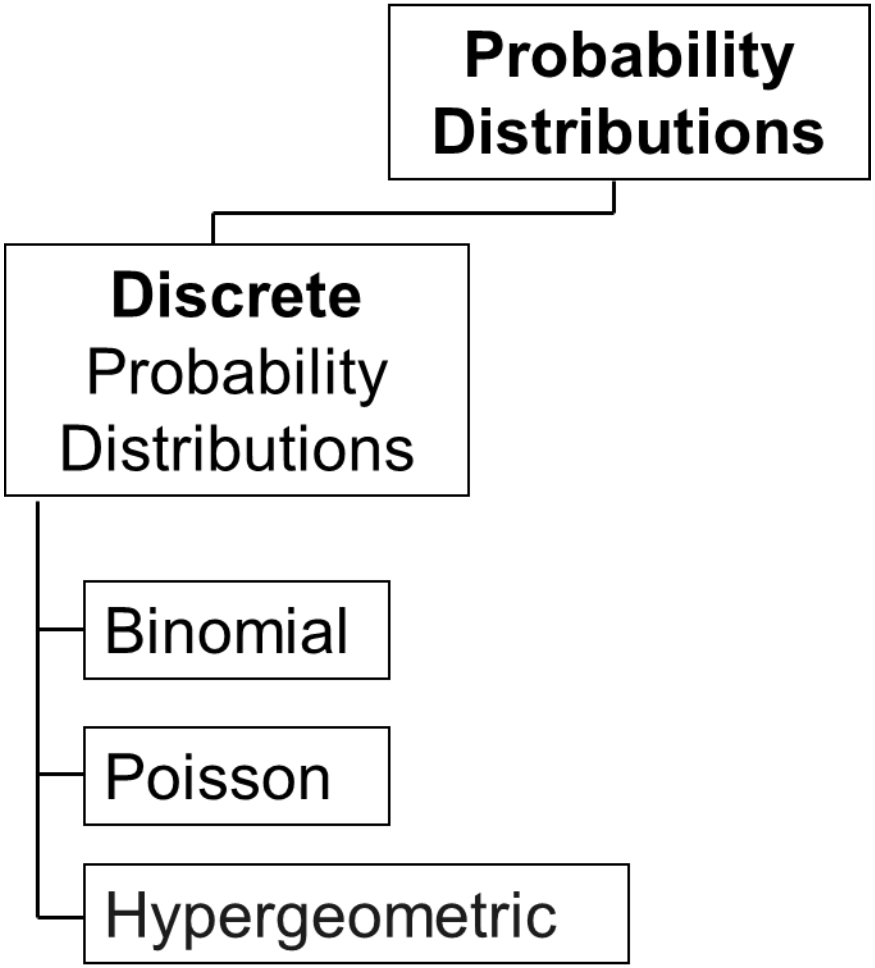
Poisson Approximation to the Binomial Distribution

Let *X* be the number of successes from *n* independent trials, each with probability of success *P*. The distribution of the number of successes, *X* , is binomial, with mean nP.

If the number of trials, *n*, is large and nP is of only moderate size (preferably $nP \le 7$), this distribution can be approximated by the Poisson distribution with $\lambda = nP$. The probability distribution of the approximating distribution is

$$P(x) = \frac{e^{-nP}(nP)^x}{x!} \quad \text{for} \quad x = 0, 1, 2, \ldots$$

Section 4.6 The Hypergeometric Distribution (1 of 2)



Section 4.6 The Hypergeometric Distribution (2 of 2)

- “n” trials in a sample taken from a finite population of size N
- Sample taken without replacement
- Outcomes of trials are dependent
- Concerned with finding the probability of “X” successes in the sample where there are “S” successes in the population

Hypergeometric Probability Distribution

$$P(x) = \frac{C_x^S C_{n-x}^{N-S}}{C_n^N} = \frac{\dfrac{S!}{x!(S-x)!} \times \dfrac{(N-S)!}{(n-x)!(N-S-n+x)!}}{\dfrac{N!}{n!(N-n)!}}$$

Where N = population size

S = number of successes in the population

$N - S$ = number of failures in the population

n = sample size

x = number of successes in the sample

$n - x$ = number of failures in the sample

Using the Hypergeometric Distribution

- Example: 3 different computers are checked from 10 in the department. 4 of the 10 computers have illegal software loaded. What is the probability that 2 of the 3 selected computers have illegal software loaded?

$$N = 10 \qquad n = 3$$

$$S = 4 \qquad x = 2$$

$$P(x=2) = \frac{C_x^S \, C_{n-x}^{N-S}}{C_n^N} = \frac{C_2^4 \, C_1^6}{C_3^{10}} = \frac{(6)(6)}{120} = 0.3$$

The probability that 2 of the 3 selected computers have illegal software loaded is 0.30, or 30%.

Section 4.7 Jointly Distributed Discrete Random Variables

- A joint probability distribution is used to express the probability that simultaneously *X* takes the specific value *x* and *Y* takes the value *y*, as a function of *x* and *y*

$$P(x, y) = P(X = x \cap Y = y)$$

- The marginal probability distributions are

$$P(x) = \sum_{y} P(x, y) \qquad P(y) = \sum_{x} P(x, y)$$

Properties of Joint Probability Distributions

Properties of Joint Probability Distributions of Discrete Random Variables

- Let *X* and *Y* be discrete random variables with joint probability distribution $P(x, y)$

1. $0 \leq P(x, y) \leq 1$ for any pair of values *x* and *y*
2. the sum of the joint probabilities $P(x, y)$ over all possible pairs of values must be 1

Conditional Probability Distribution

- The conditional probability distribution of the random variable *Y* expresses the probability that *Y* takes the value *y* when the value *x* is specified for *X*.

$$P(y \mid x) = \frac{P(x, y)}{P(x)}$$

- Similarly, the conditional probability function of *X*, given *Y* = *y* is:

$$P(x \mid y) = \frac{P(x, y)}{P(y)}$$

Independence

- The jointly distributed random variables *X* and *Y* are said to be independent if and only if their joint probability distribution is the product of their marginal probability functions:

$$P(x, y) = P(x)P(y)$$

for all possible pairs of values *x* and *y*

- A set of *k* random variables are independent if and only if

$$P(x_1, x_2, \cdots, x_k) = P(x_1)P(x_2)\cdots P(x_k)$$

Conditional Mean and Variance

- The conditional mean is

$$\mu_{Y|X} = E\left[Y \mid X\right] = \sum_{y} \left(y \mid x\right) P\left(y \mid x\right)$$

- The conditional variance is

$$\sigma^2_{Y|X} = E\left[\left(Y - \mu_{Y|X}\right)^2 \mid X\right] = \sum_{y} \left[\left(y - \mu_{Y|X}\right)^2 \mid x\right] P\left(y \mid x\right)$$

Covariance

- Let *X* and *Y* be discrete random variables with means μ_X and μ_Y
- The expected value of $(X-\mu_X)(Y-\mu_Y)$ is called the covariance between *X* and *Y*
- For discrete random variables

$$Cov(X,Y)=E\left[(X-\mu_X)(Y-\mu_Y)\right]=\sum_x\sum_y(x-\mu_x)(y-\mu_y)P(x,y)$$

- An equivalent expression is

$$Cov(X,Y)=E(XY)-\mu_x\mu_y=\sum_x\sum_y xyP(x,y)-\mu_x\mu_y$$

Correlation

- The correlation between *X* and *Y* is:

$$\rho = Corr(X,Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

- $-1 \le \rho \le 1$
- $\rho = 0 \Rightarrow$ no linear relationship between *X* and *Y*
- $\rho > 0 \Rightarrow$ positive linear relationship between *X* and *Y*
  - when *X* is high (low) then *Y* is likely to be high (low)
  - $\rho = +1 \Rightarrow$ perfect positive linear dependency
- $\rho < 0 \Rightarrow$ negative linear relationship between *X* and *Y*
  - when *X* is high (low) then *Y* is likely to be low (high)
  - $\rho = -1 \Rightarrow$ perfect negative linear dependency

Covariance and Independence

- The covariance measures the strength of the linear relationship between two variables
- If two random variables are statistically independent, the covariance between them is 0
  - The converse is not necessarily true

Portfolio Analysis (1 of 2)

- Let random variable *X* be the price for stock *A*
- Let random variable *Y* be the price for stock *B*
- The market value, *W*, for the portfolio is given by the linear function

$$W = aX + bY$$

(*a* is the number of shares of stock *A*,
b is the number of shares of stock *B*)

Portfolio Analysis (2 of 2)

- The mean value for *W* is

$$\mu_W = E[W] = E[aX + bY]$$
$$= a\mu_X + b\mu_Y$$

- The variance for *W* is

$$\sigma_W^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2abCov(X,Y)$$

or using the correlation formula

$$\sigma_W^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2abCorr(X,Y)\sigma_X\sigma_Y$$

Example 2: Investment Returns

Return per $1,000 for two types of investments

		Investment	
***P*(*x*$_i$ *y*$_i$)**	**Economic condition**	**Passive Fund *X***	**Aggressive Fund *Y***
.2	Recession	−$25	− $200
.5	Stable Economy	+ 50	+ 60
.3	Expanding Economy	+100	+ 350

$$E(x) = \mu_x = (-25)(.2) + (50)(.5) + (100)(.3) = 50$$

$$E(y) = \mu_y = (-200)(.2) + (60)(.5) + (350)(.3) = 95$$

Computing the Standard Deviation for Investment Returns

		Investment	
***P*(*xi* *yi*)**	**Economic condition**	**Passive Fund *X***	**Aggressive Fund *Y***
0.2	Recession	−$25	−$200
0.5	Stable Economy	+ 50	+ 60
0.3	Expanding Economy	+ 100	+ 350

$$\sigma_x = \sqrt{(-25-50)^2(0.2)+(50-50)^2(0.5)+(100-50)^2(0.3)}$$
$$= 43.30$$

$$\sigma_y = \sqrt{(-200-95)^2(0.2)+(60-95)^2(0.5)+(350-95)^2(0.3)}$$
$$= 193.71$$

Covariance for Investment Returns

		Investment	
$P(x_i y_i)$	**Economic condition**	**Passive Fund *X***	**Aggressive Fund *Y***
.2	Recession	−$25	− $200
.5	Stable Economy	+ 50	+ 60
.3	Expanding Economy	+ 100	+ 350

$$
\begin{aligned}
Cov(X,Y) &= (-25-50)(-200-95)(.2)+(50-50)(60-95)(.5) \\
&\quad +(100-50)(350-95)(.3) \\
&= 8250
\end{aligned}
$$

Portfolio Example

Investment X: $\mu_x = 50$ $\sigma_x = 43.30$

Investment Y: $\mu_y = 95$ $\sigma_y = 193.21$

$$\sigma_{xy} = 8250$$

Suppose 40% of the portfolio (P) is in Investment X and 60% is in Investment Y:

$$E(P) = .4(50) + (.6)(95) = 77$$

$$\sigma_P = \sqrt{(.4)^2 (43.30)^2 + (.6)^2 (193.21)^2 + 2(.4)(.6)(8250)}$$

$$= 133.04$$

The portfolio return and portfolio variability are between the values for investments X and Y considered individually

Interpreting the Results for Investment Returns

- The aggressive fund has a higher expected return, but much more risk

$$\mu_y = 95 > \mu_x = 50$$

but

$$\sigma_y = 193.21 > \sigma_x = 43.30$$

- The Covariance of 8250 indicates that the two investments are positively related and will vary in the same direction

Chapter Summary

- Defined discrete random variables and probability distributions
- Discussed the Binomial distribution
- Reviewed the Poisson distribution
- Discussed the Hypergeometric distribution
- Defined covariance and the correlation between two random variables
- Examined application to portfolio investment