Statistics for Managers Using Microsoft Excel
7th Edition

Chapter 5

Discrete Probability Distributions

Learning Objectives

In this chapter, you learn:

- The properties of a probability distribution
- To compute the expected value and variance of a probability distribution
- To calculate the covariance and understand its use in finance
- To compute probabilities from binomial, hypergeometric, and Poisson distributions
- How the binomial, hypergeometric, and Poisson distributions can be used to solve business problems

Definitions

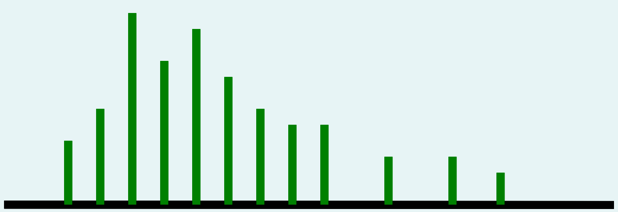
- **Discrete** variables produce outcomes that come from a counting process (e.g. number of classes you are taking).

- **Continuous** variables produce outcomes that come from a measurement (e.g. your annual salary, or your weight).

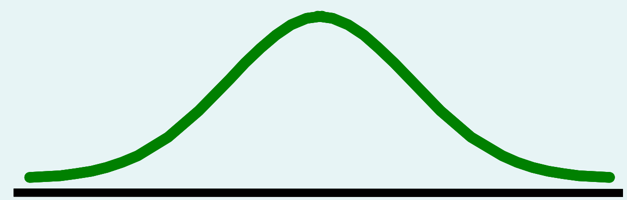
Types Of Variables

Types Of Variables

Ch. 5 — **Discrete Variable**



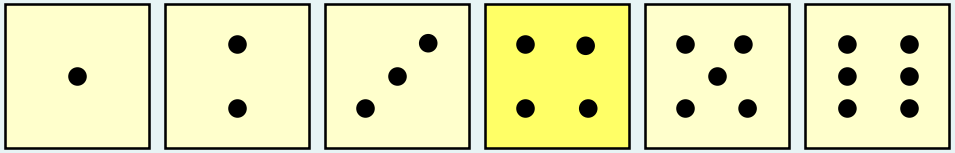
Continuous Variable — Ch. 6



Discrete Random Variables

- Can only assume a countable number of values

Examples:



- **Roll a die twice**
  Let X be the number of times 4 occurs
  (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)

Probability Distribution For A Discrete Random Variable

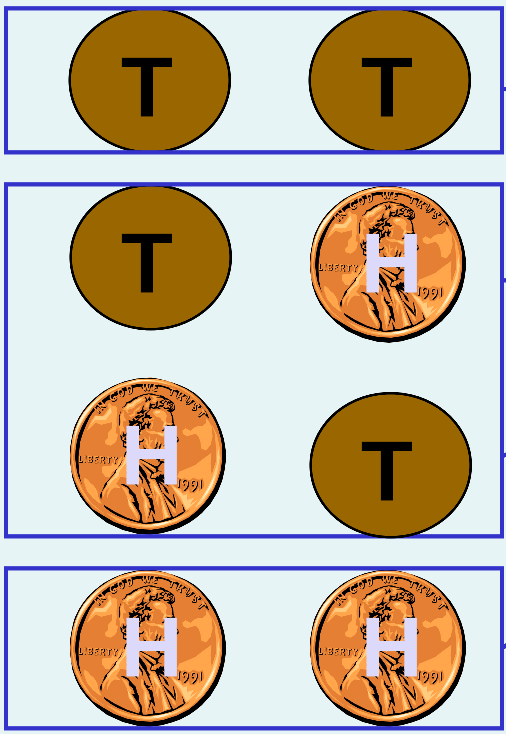
- A **probability distribution for a discrete random variable** is a mutually exclusive listing of all possible numerical outcomes for that variable and a probability of occurrence associated with each outcome.

Number of Classes Taken	Probability
2	0.20
3	0.40
4	0.24
5	0.16

Example of a Discrete Random Variable Probability Distribution

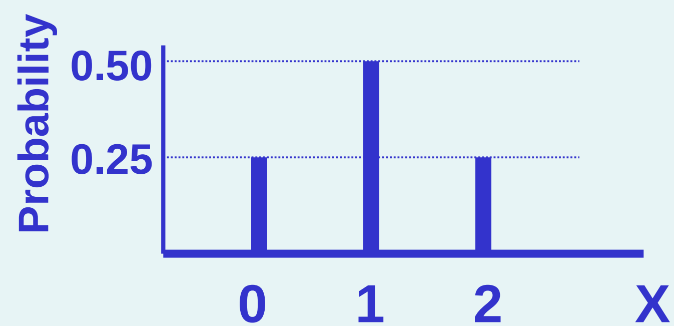
Experiment: Toss 2 Coins. Let X = # heads.

4 possible outcomes



Probability Distribution

X Value	Probability
0	1/4 = 0.25
1	2/4 = 0.50
2	1/4 = 0.25



Discrete Variables
Expected Value (Measuring Center)

- Expected Value (or mean) of a discrete variable (Weighted Average)

$$\mu = E(X) = \sum_{i=1}^{N} X_i P(X = X_i)$$

- Example: Toss 2 coins, X = # of heads, compute expected value of X:

E(X) = ((0)(0.25) + (1)(0.50) + (2)(0.25)) = 1.0

X	$P(X=X_i)$
0	0.25
1	0.50
2	0.25

Discrete Random Variables Measuring Dispersion

- Variance of a discrete random variable

$$\sigma^2 = \sum_{i=1}^{N} [X_i - E(X)]^2 P(X = X_i)$$

- Standard Deviation of a discrete random variable

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{N} [X_i - E(X)]^2 P(X = X_i)}$$

where:

E(X) = Expected value of the discrete random variable X
X_i = the i^{th} outcome of X
$P(X=X_i)$ = Probability of the i^{th} occurrence of X

Discrete Random Variables Measuring Dispersion

(continued)

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

$$\sigma = \sqrt{\sum [X_i - E(X)]^2 P(X_i)}$$

$$\sigma = \sqrt{(0-1)^2(0.25) + (1-1)^2(0.50) + (2-1)^2(0.25)} = \sqrt{0.50} = 0.707$$

Possible number of heads = 0, 1, or 2

Covariance

- The covariance measures the strength of the linear relationship between two discrete random variables X and Y.

- A positive covariance indicates a positive relationship.

- A negative covariance indicates a negative relationship.

The Covariance Formula

- The covariance formula:

$$\sigma_{XY} = \sum_{i=1}^{N} [X_i - E(X)][(Y_i - E(Y)] \, P(X = X_i, Y = Y_i)$$

where: X = discrete random variable X

X_i = the i^{th} outcome of X

Y = discrete random variable Y

Y_i = the i^{th} outcome of Y

$P(X=X_{i,Y=}Y_i)$ = probability of occurrence of the i^{th} outcome of X and the i^{th} outcome of Y

Investment Returns
The Mean

Consider the return per $1000 for two types of investments.

Prob.	Economic Condition	Investment	
		Passive Fund X	Aggressive Fund Y
0.2	Recession	- $25	- $200
0.5	Stable Economy	+ $50	+ $60
0.3	Expanding Economy	+ $100	+ $350

Investment Returns
The Mean

$$E(X) = \mu_X = (-25)(.2) +(50)(.5) + (100)(.3) = 50$$

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

Interpretation: Fund X is averaging a $50.00 return and fund Y is averaging a $95.00 return per $1000 invested.

Investment Returns
Standard Deviation

$$\sigma_X = \sqrt{(-25-50)^2(.2)+(50-50)^2(.5)+(100-50)^2(.3)}$$
$$= 43.30$$

$$\sigma_Y = \sqrt{(-200-95)^2(.2)+(60-95)^2(.5)+(350-95)^2(.3)}$$
$$= 193.71$$

Interpretation: Even though fund Y has a higher average return, it is subject to much more variability and the probability of loss is higher.

Investment Returns
Covariance

$$\sigma_{XY} = (-25-50)(-200-95)(.2)+(50-50)(60-95)(.5)$$
$$+(100-50)(350-95)(.3)$$
$$= 8{,}250$$

Interpretation: Since the covariance is large and positive, there is a positive relationship between the two investment funds, meaning that they will likely rise and fall together.

The Sum of Two Random Variables

- Expected Value of the sum of two random variables:

$$E(X+Y) = E(X) + E(Y)$$

- Variance of the sum of two random variables:

$$Var(X+Y) = \sigma^2_{X+Y} = \sigma^2_X + \sigma^2_Y + 2\sigma_{XY}$$

- Standard deviation of the sum of two random variables:

$$\sigma_{X+Y} = \sqrt{\sigma^2_{X+Y}}$$

Portfolio Expected Return and Expected Risk

- Investment portfolios usually contain several different funds (random variables)
- The expected return and standard deviation of two funds together can now be calculated.
- Investment Objective: Maximize return (mean) while minimizing risk (standard deviation).

Portfolio Expected Return and Portfolio Risk

- Portfolio expected return (weighted average return):

$$E(P) = w\,E(X) + (1 - w)\,E(Y)$$

- Portfolio risk (weighted variability)

$$\sigma_P = \sqrt{w^2\sigma_X^2 + (1 - w)^2\sigma_Y^2 + 2w(1 - w)\sigma_{XY}}$$

Where w = proportion of portfolio value in asset X

(1 - w) = proportion of portfolio value in asset Y

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 is in Investment X and 60% is in Investment Y:

$$E(P) = 0.4(50) + (0.6)(95) = 77$$

$$\sigma_P = \sqrt{(0.4)^2(43.30)^2 + (0.6)^2(193.71)^2 + 2(0.4)(0.6)(8{,}250)}$$
$$= 133.30$$

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

Probability Distributions

Probability Distributions

- Ch. 5 — **Discrete** Probability Distributions
  - Binomial
  - Poisson
  - Hypergeometric
- Ch. 6 — **Continuous** Probability Distributions
  - Normal
  - Uniform
  - Exponential

Binomial Probability Distribution

- A fixed number of observations, n
  - e.g., 15 tosses of a coin; ten light bulbs taken from a warehouse
- Each observation is categorized as to whether or not the "event of interest" occurred
  - e.g., head or tail in each toss of a coin; defective or not defective light bulb
  - Since these two categories are mutually exclusive and collectively exhaustive
    - When the probability of the event of interest is represented as π, then the probability of the event of interest not occurring is 1 - π
- Constant probability for the event of interest occurring (π) for each observation
  - Probability of getting a tail is the same each time we toss the coin

Binomial Probability Distribution

(continued)

- Observations are independent
  - The outcome of one observation does not affect the outcome of the other
  - Two sampling methods deliver independence
    - Infinite population without replacement
    - Finite population with replacement

Possible Applications for the Binomial Distribution

- 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 Counting Techniques

- Suppose the event of interest is obtaining heads on the toss of a fair coin. You are to toss the coin three times. In how many ways can you get two heads?
- Possible ways: HHT, HTH, THH, so there are three ways you can getting two heads.
- This situation is fairly simple. We need to be able to count the number of ways for more complicated situations.

Counting Techniques
Rule of Combinations

- The number of combinations of selecting X objects out of n objects is

$$ {}_{n}C_{x} = \frac{n!}{X!(n-X)!} $$

where:

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

$X! = (X)(X-1)(X-2) \cdots (2)(1)$

$0! = 1$ (by definition)

Counting Techniques Rule of Combinations

- How many possible 3 scoop combinations could you create at an ice cream parlor if you have 31 flavors to select from?
- The total choices is n = 31, and we select X = 3.

$$_{31}C_3 = \frac{31!}{3!(31-3)!} = \frac{31!}{3!28!} = \frac{31 \bullet 30 \bullet 29 \bullet 28!}{3 \bullet 2 \bullet 1 \bullet 28!} = 31 \bullet 5 \bullet 29 = 4{,}495$$

Binomial Distribution Formula

$$P(X=x\,|n,\pi) = \frac{n!}{x!\,(n-x)!}\,\pi^{x}\,(1-\pi)^{n-x}$$

P(X=x|n,π) = probability of **x** events of interest in **n** trials, with the probability of an "event of interest" being **π** for each trial

x = number of "events of interest" in sample, (x = 0, 1, 2, ..., n)

n = sample size (number of trials or observations)

π = probability of "event of interest"

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

n = 4

π = 0.5

1 - π = (1 - 0.5) = 0.5

X = 0, 1, 2, 3, 4

Example: Calculating a Binomial Probability

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

$x = 1$, $n = 5$, and $\pi = 0.1$

$$P(X = 1 | 5, 0.1) = \frac{n!}{x!(n-x)!}\pi^{x}(1-\pi)^{n-x}$$

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

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

$$= 0.32805$$

The Binomial Distribution Example

Suppose the probability of purchasing a defective computer is 0.02. What is the probability of purchasing 2 defective computers in a group of 10?

$$x = 2,\ n = 10,\ \text{and}\ \pi = 0.02$$

$$P(X = 2 | 10, 0.02) = \frac{n!}{x!(n-x)!}\pi^{x}(1-\pi)^{n-x}$$

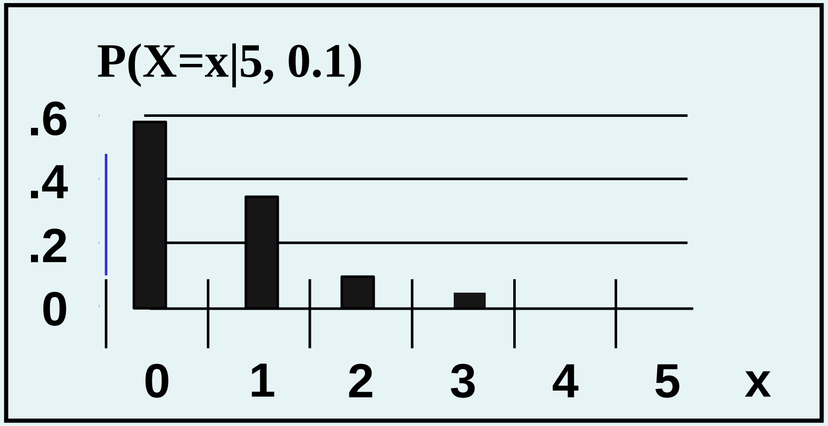
$$= \frac{10!}{2!(10-2)!}(.02)^{2}(1-.02)^{10-2}$$

$$= (45)(.0004)(.8508)$$

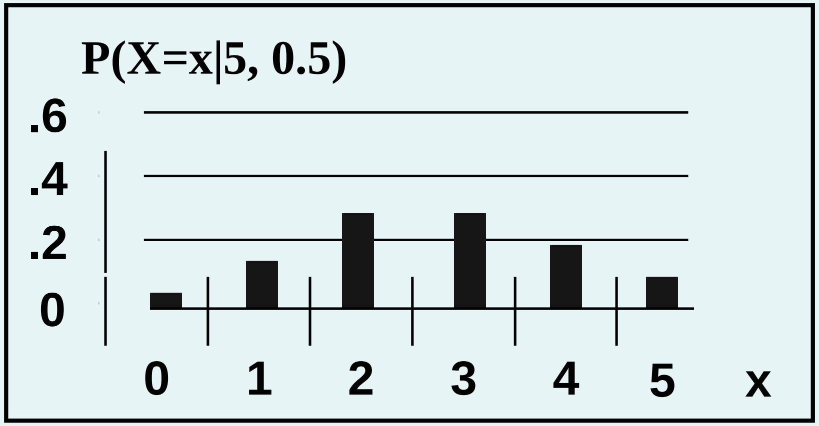
$$= .01531$$

The Binomial Distribution Shape

- The shape of the binomial distribution depends on the values of π and n
- Here, $n = 5$ and $\pi = .1$



- Here, $n = 5$ and $\pi = .5$



The Binomial Distribution Using Binomial Tables (Available On Line)

n = 10									
x	…	π=.20	π=.25	π=.30	**π=.35**	π=.40	π=.45	π=.50	
0	…	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010	10
1	…	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098	9
2	…	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439	8
3	…	0.2013	0.2503	0.2668	**0.2522**	0.2150	0.1665	0.1172	7
4	…	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051	6
5	…	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461	5
6	…	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051	4
7	…	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172	3
8	…	0.0001	**0.0004**	0.0014	0.0043	0.0106	0.0229	0.0439	**2**
9	…	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098	1
10	…	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0
	…	π=.80	**π=.75**	π=.70	π=.65	π=.60	π=.55	π=.50	**x**

Examples:

n = 10, π = 0.35, x = 3: P(X = 3|10, 0.35) = 0.2522

n = 10, π = 0.75, x = 8: P(X = 8|10, 0.75) = 0.0004

Binomial Distribution Characteristics

- Mean

$$\mu = E(X) = n\pi$$

- Variance and Standard Deviation

$$\sigma^2 = n\pi(1 - \pi)$$

$$\sigma = \sqrt{n\pi(1 - \pi)}$$

Where n = sample size

π = probability of the event of interest for any trial

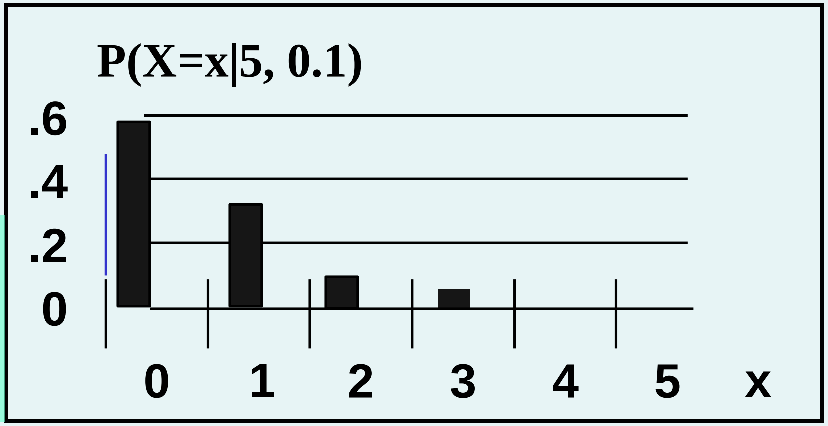
(1 – π) = probability of no event of interest for any trial

The Binomial Distribution Characteristics

Examples

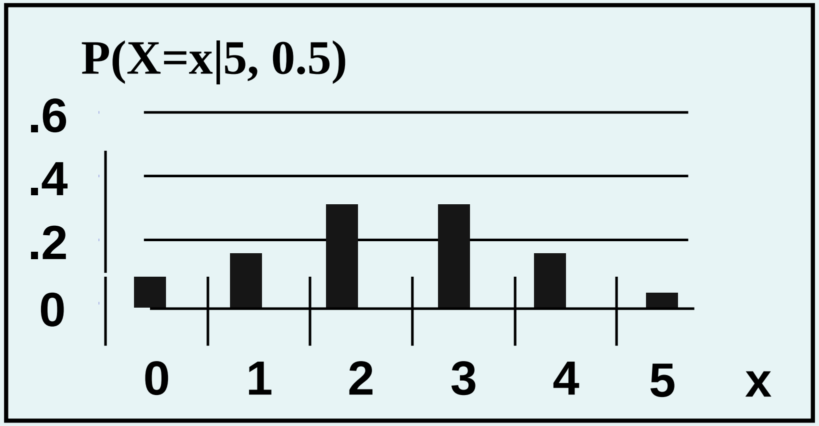
$$\mu = n\pi = (5)(.1) = 0.5$$

$$\sigma = \sqrt{n\pi(1-\pi)} = \sqrt{(5)(.1)(1-.1)} = 0.6708$$



$$\mu = n\pi = (5)(.5) = 2.5$$

$$\sigma = \sqrt{n\pi(1-\pi)} = \sqrt{(5)(.5)(1-.5)} = 1.118$$



Using Excel For The Binomial Distribution

	A	B	
1	**Binomial Probabilities**		
2			
3	**Data**		
4	Sample size	4	
5	Probability of an event of interest	0.1	
6			
7	**Statistics**		
8	Mean	0.4	**=B4 * B5**
9	Variance	0.36	**=B8 * (1 - B5)**
10	Standard deviation	0.6	**=SQRT(B9)**
11			
12	**Binomial Probabilities Table**		
13	***X***	***P(X)***	
14	0	0.6561	**=BINOM.DIST(A14, B4, B5, FALSE)**
15	1	0.2916	**=BINOM.DIST(A15, B4, B5, FALSE)**
16	2	0.0486	**=BINOM.DIST(A16, B4, B5, FALSE)**
17	3	0.0036	**=BINOM.DIST(A17, B4, B5, FALSE)**
18	4	0.0001	**=BINOM.DIST(A18, B4, B5, FALSE)**

The Poisson Distribution Definitions

- You use the **Poisson distribution** when you are interested in the number of times an event occurs in a given **area of opportunity**.
- An **area of opportunity** is a continuous unit or interval of time, volume, or such area in which more than one occurrence of an event can occur.
  - The number of scratches in a car's paint
  - The number of mosquito bites on a person
  - The number of computer crashes in a day

The Poisson Distribution

- Apply the Poisson Distribution when:
  - You wish to count the number of times an event occurs in a given area of opportunity
  - The probability that an event occurs in one area of opportunity is the same for all areas of opportunity
  - The number of events that occur in one area of opportunity is independent of the number of events that occur in the other areas of opportunity
  - The probability that two or more events occur in an area of opportunity approaches zero as the area of opportunity becomes smaller
  - The average number of events per unit is λ (lambda)

Poisson Distribution Formula

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

where:

x = number of events in an area of opportunity

λ = expected number of events

e = base of the natural logarithm system (2.71828...)

Poisson Distribution Characteristics

- Mean

$$\mu = \lambda$$

- Variance and Standard Deviation

$$\sigma^2 = \lambda$$

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

where λ = expected number of events

Using Poisson Tables (Available On Line)

X	λ								
	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 $P(X = 2 \mid \lambda = 0.50)$

$$P(X = 2 \mid 0.50) = \frac{e^{-\lambda}\lambda^{X}}{X!} = \frac{e^{-0.50}(0.50)^2}{2!} = 0.0758$$

Using Excel For The Poisson Distribution

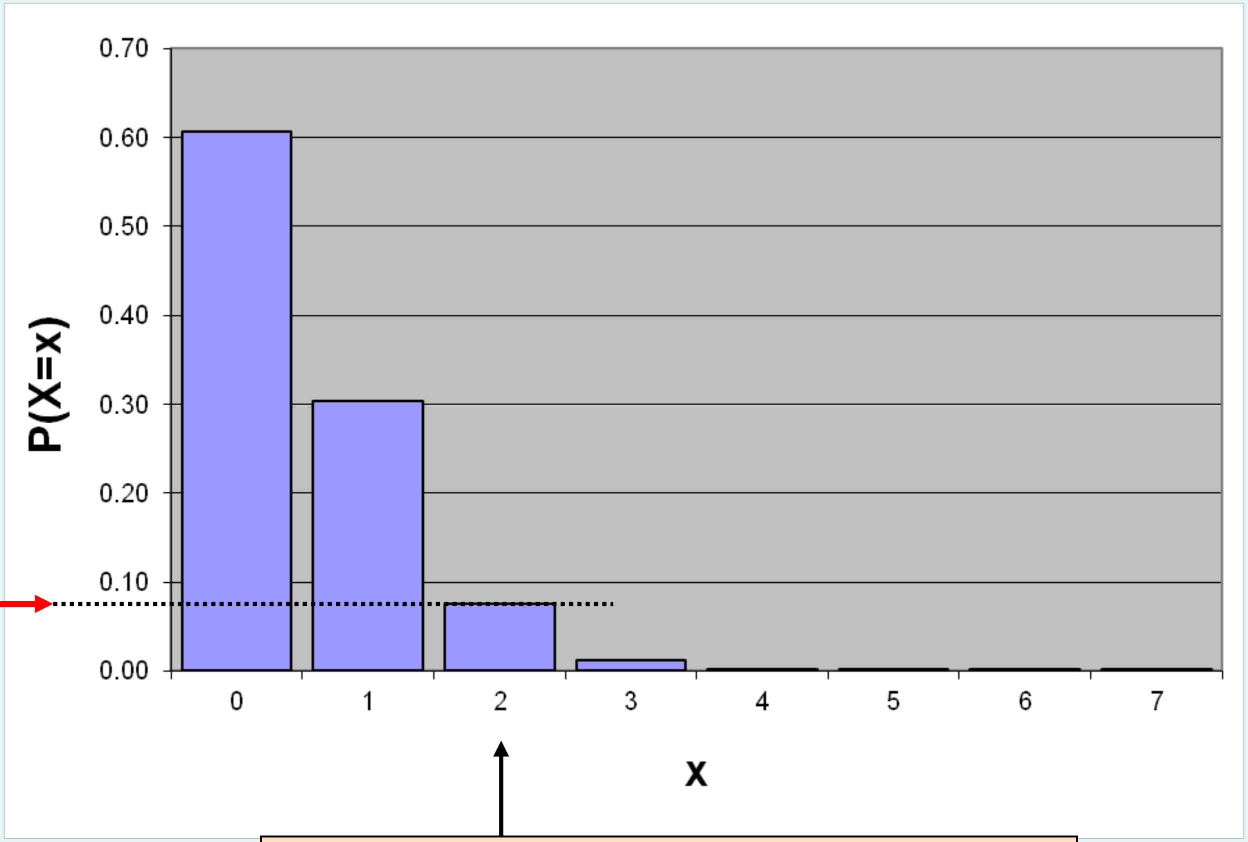
	A	B	C	D	E
1	**Poisson Probabilities**				
2					
3	**Data**				
4	Mean/Expected number of events of interest:				3
5					
6	**Poisson Probabilities Table**				
7	***X***	***P(X)***			
8	0	0.0498	=POISSON.DIST(A8, \$E\$4, FALSE)		
9	1	0.1494	=POISSON.DIST(A9, \$E\$4, FALSE)		
10	2	0.2240	=POISSON.DIST(A10, \$E\$4, FALSE)		
11	3	0.2240	=POISSON.DIST(A11, \$E\$4, FALSE)		
12	4	0.1680	=POISSON.DIST(A12, \$E\$4, FALSE)		
13	5	0.1008	=POISSON.DIST(A13, \$E\$4, FALSE)		
14	6	0.0504	=POISSON.DIST(A14, \$E\$4, FALSE)		
15	7	0.0216	=POISSON.DIST(A15, \$E\$4, FALSE)		
16	8	0.0081	=POISSON.DIST(A16, \$E\$4, FALSE)		
17	9	0.0027	=POISSON.DIST(A17, \$E\$4, FALSE)		
18	10	0.0008	=POISSON.DIST(A18, \$E\$4, FALSE)		
19	11	0.0002	=POISSON.DIST(A19, \$E\$4, FALSE)		
20	12	0.0001	=POISSON.DIST(A20, \$E\$4, FALSE)		
21	13	0.0000	=POISSON.DIST(A21, \$E\$4, FALSE)		
22	14	0.0000	=POISSON.DIST(A22, \$E\$4, FALSE)		
23	15	0.0000	=POISSON.DIST(A23, \$E\$4, FALSE)		

Graph of Poisson Probabilities

Graphically:

$\lambda = 0.50$

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

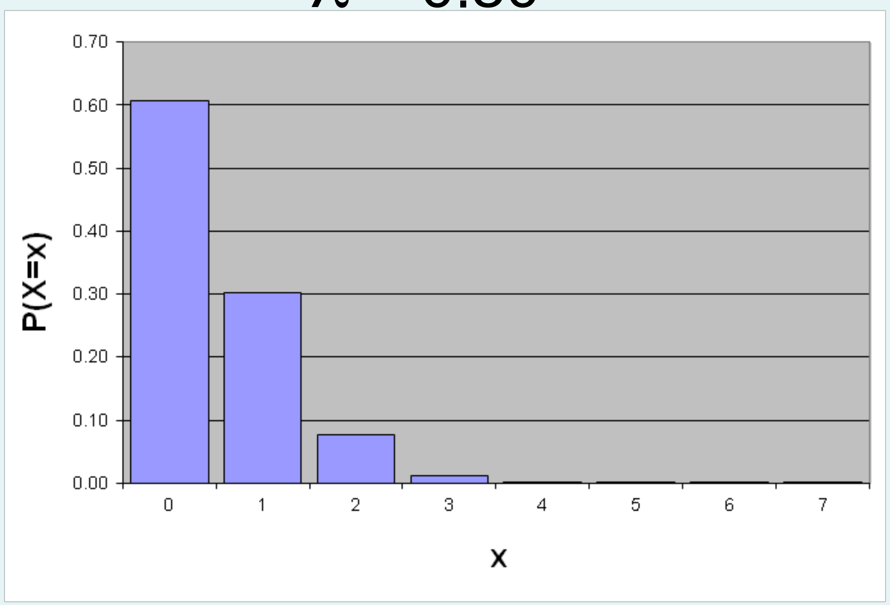


$P(X = 2 \mid \lambda=0.50) = 0.0758$

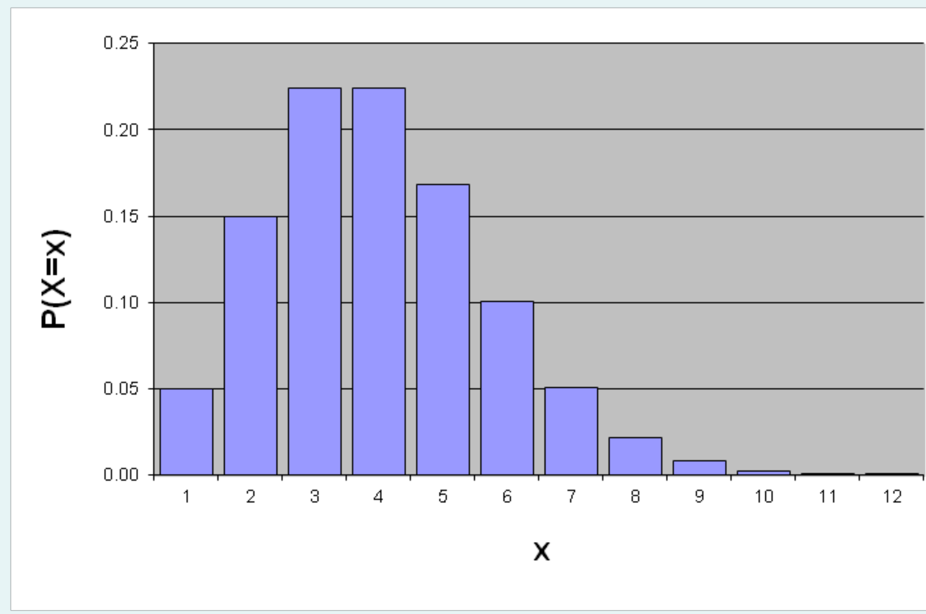
Poisson Distribution Shape

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

$\lambda = 0.50$



$\lambda = 3.00$



The Hypergeometric Distribution

- The **binomial distribution** is applicable when selecting from a finite population with replacement or from an infinite population without replacement.

- The **hypergeometric distribution** is applicable when selecting from a finite population without replacement.

The Hypergeometric Distribution

- "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" items of interest in the sample where there are "A" items of interest in the population

Hypergeometric Distribution Formula

$$P(X = x \mid n, N, A) = \frac{[{}_{A}C_{X}][{}_{N-A}C_{n-X}]}{{}_{N}C_{n}} = \frac{\binom{A}{X}\binom{N-A}{n-X}}{\binom{N}{n}}$$

Where

N = population size

A = number of items of interest in the population

N – A = number of events not of interest in the population

n = sample size

x = number of items of interest in the sample

n – x = number of events not of interest in the sample

Properties of the Hypergeometric Distribution

- The mean of the hypergeometric distribution is

$$\mu = E(X) = \frac{nA}{N}$$

- The standard deviation is

$$\sigma = \sqrt{\frac{nA(N-A)}{N^2}} \cdot \sqrt{\frac{N-n}{N-1}}$$

Where $\sqrt{\frac{N-n}{N-1}}$ is called the "Finite Population Correction Factor" from sampling without replacement from a finite population

Using the Hypergeometric Distribution

- Example: 3 different computers are checked out 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	n = 3
A = 4	x = 2

$$P(X = 2 \mid 3,10,4) = \frac{\binom{A}{X}\binom{N-A}{n-X}}{\binom{N}{n}} = \frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}} = \frac{(6)(6)}{120} = 0.3$$

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

Using Excel for the Hypergeometric Distribution

	A	B	
1	**Hypergeometric Probabilities**		
2			
3	**Data**		
4	Sample size	8	
5	No. of events of interest in population	10	
6	Population size	30	
7			
8	**Hypergeometric Probabilities Table**		
9	***X***	***P(X)***	
10	0	0.0215	**=HYPGEOM.DIST(A10, B4, B5, B6, FALSE)**
11	1	0.1324	**=HYPGEOM.DIST(A11, B4, B5, B6, FALSE)**
12	2	0.2980	**=HYPGEOM.DIST(A12, B4, B5, B6, FALSE)**
13	3	0.3179	**=HYPGEOM.DIST(A13, B4, B5, B6, FALSE)**
14	4	0.1738	**=HYPGEOM.DIST(A14, B4, B5, B6, FALSE)**
15	5	0.0491	**=HYPGEOM.DIST(A15, B4, B5, B6, FALSE)**
16	6	0.0068	**=HYPGEOM.DIST(A16, B4, B5, B6, FALSE)**
17	7	0.0004	**=HYPGEOM.DIST(A17, B4, B5, B6, FALSE)**
18	8	0.0000	**=HYPGEOM.DIST(A18, B4, B5, B6, FALSE)**

Chapter Summary

In this chapter we discussed

- The probability distribution of a discrete random variable
- The covariance and its application in finance
- The Binomial distribution
- The Poisson distribution
- The Hypergeometric distribution