Random Variables and Their Distributions

Learning Objectives

This week, we learn:

- The notion of random variables
- The properties of a probability distribution
- To compute the expected value and variance of a probability distribution
- To compute probabilities from binomial and Poisson distributions
- How the binomial and Poisson distributions can be used to solve business problems
- Applications to Finance

Definitions

- **Random variable** is a quantity that depends on chance.

Example: gender of 3 children (tossing 3 coins).

Let X be the number of girls. X is unknown but we know this:

X can be 0, 1, 2, or 3.

{X=0}, {X=1}, {X=2}, {X=3} are events, they have probabilities.

Random variables, example

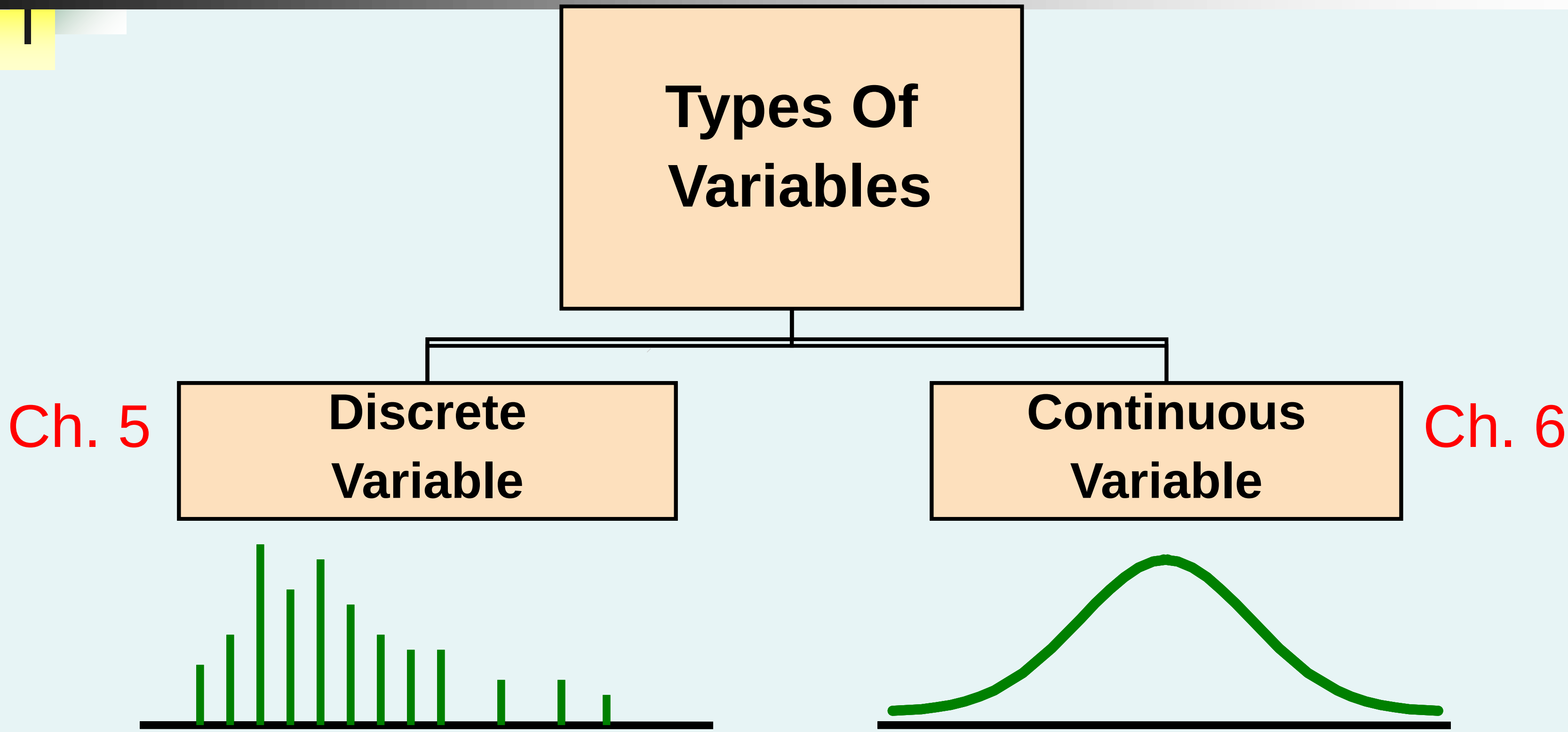
- X is the number of girls in a family of 3 children.

Value of X	Probability	Equally likely outcomes
0	1/8	BBB
1	3/8	GBB, BGB, BBG
2	3/8	GGB, GBG, BGG
3	1/8	GGG

- G=girl, B=boy

A coin is tossed 3 times. X is the number of heads.
Same values, same probabilities. **Same distribution.**

Types Of Variables

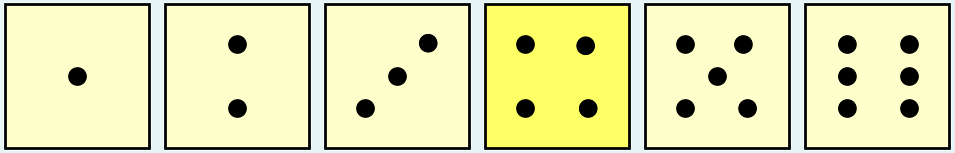


Discrete variables take only separate, isolated values.
Continuous variables can take any value in an interval.

Discrete Random Variables

- Can only assume a countable number of values

Examples:



- **Roll a die**

Let X be the number of dots.

- **Number of shares bought**

- **Number of satisfied customers**

- **Proportion of satisfied customers, in a sample of 100**

Probability Distribution For A Discrete Random Variable

- *A **probability distribution for a discrete random variable** is a list of its all possible values along with the corresponding probabilities.*

- Example:

x **values of X**	**P(x)** **probabilities**
0	1/8
1	3/8
2	3/8
3	1/8

Expected Value

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

$$\mu = \mathrm{E(X)} = \sum_x xP(x)$$

- 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)
0	0.25
1	0.50
2	0.25

Variance and Standard Deviation

- Variance of a discrete random variable

$$\sigma^2 = Var(X) = E(X-\mu)^2 = \sum_x [x - \mathrm{E}(X)]^2 \mathrm{P}(x)$$

- Standard Deviation of a discrete random variable

$$\sigma = \sqrt{\sigma^2}$$

Covariance and Applications to Finance

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:

$$\sigma_{XY} = E(X-\mu_X)(Y-\mu_Y) = \sum_x \sum_y (x-EX)(y-EY)P(x,y)$$

where: X = discrete random variable X
x = value of X
Y = discrete random variable Y
y = value of Y
P(x,y) = P(X=x and Y=y)
= probability of a joint event {X=x and Y=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.

These are expected returns.

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.

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:

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

The Weighted Sum Two Random Variables

- Expected Value of a weighted sum:

$$\boxed{\mathrm{E}(\mathrm{aX}+b\mathrm{Y})=aE(X)+bE(Y)}$$

- Variance of a weighted sum:

$$\boxed{\mathrm{Var}(\mathrm{aX}+b\mathrm{Y})=a^2\sigma_{\mathrm{X}}^2+b^2\sigma_{\mathrm{Y}}^2+2ab\sigma_{XY}}$$

- Standard deviation of a weighted sum:

$$\sigma_{aX+bY}=\sqrt{\sigma_{aX+BY}^2}$$

Portfolio Expected Return and Expected Risk

- Investment portfolios usually contain several different funds (random variables)

- *Investment Objective: Maximize return (mean) while minimizing risk (standard deviation).*

Portfolio Expected Return and Portfolio Risk

Consider a portfolio with:

w = proportion of portfolio value in asset X

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

- Portfolio expected return (weighted average return):

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

- Portfolio risk (weighted variability)

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

Portfolio Example 1

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:

$$\boxed{E(P) = 0.4(50) + (0.6)(95) = 77}$$

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

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

Portfolio Example 2

We would like to invest $10,000 into companies A and B. Shares of A cost $20 per share. The market analysis shows that their expected return is $1 per share with a standard deviation of $0.5. Shares of B cost $50 per share, with an expected return of $2.50 and a standard deviation of $1 per share, and returns from the two companies are independent.

What is the most optimal portfolio consisting of shares of A and B, in terms of the maximum expected return at the minimum risk?

Probability Distributions Commonly Used in Practice

Probability Distributions

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

Bernoulli Distribution

- The simplest non-trivial distribution:
- X takes two values, X=0 and X=1

(Bernoulli trials have two outcomes - pass and fail, good and defective, girl and boy, “success” and “failure”).

- Distribution of X:

$P(X=1) = P(1) = \pi$ = probability of a success

$P(X=0) = P(0) = 1-\pi$ = probability of a failure

- $E(X) = \pi, \quad Var(X) = \pi(1- \pi)$

Binomial Distribution

- X = number of "successes" in n independent Bernoulli trials

- Examples:

  number of A grades

  number of defective products

  number of votes for a particular candidate

  number of games won

  number of Stanley Cups won

  etc.

Business Applications of 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

Binomial Distribution Formula

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

$P(X=x|n,\pi)$ = 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

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)

Here, we have X successes in n Bernoulli trials

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.

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

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?

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

$$\begin{aligned} P(X = 2 \mid 10, 0.02) &= \frac{n!}{x!(n-x)!}\pi^{x}(1-\pi)^{n-x} \\ &= \frac{10!}{2!(10-2)!}(.02)^{2}(1-.02)^{10-2} \\ &= (45)(.0004)(.8508) \\ &= .01531 \end{aligned}$$

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

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)

Poisson Distribution

- **Poisson distribution** describes a number of unexpected events that occur independently, one by one, at random times
  - 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 number of thunderstorms in a given month
  - The number of traffic accidents in a given day
  - The number of calls to a customer service

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

$$\boxed{\mu = \lambda}$$

- Variance and Standard Deviation

$$\boxed{\sigma^2 = \lambda}$$

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

where λ = expected number of events

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, E4, FALSE)**		
9	1	0.1494	**=POISSON.DIST(A9, E4, FALSE)**		
10	2	0.2240	**=POISSON.DIST(A10, E4, FALSE)**		
11	3	0.2240	**=POISSON.DIST(A11, E4, FALSE)**		
12	4	0.1680	**=POISSON.DIST(A12, E4, FALSE)**		
13	5	0.1008	**=POISSON.DIST(A13, E4, FALSE)**		
14	6	0.0504	**=POISSON.DIST(A14, E4, FALSE)**		
15	7	0.0216	**=POISSON.DIST(A15, E4, FALSE)**		
16	8	0.0081	**=POISSON.DIST(A16, E4, FALSE)**		
17	9	0.0027	**=POISSON.DIST(A17, E4, FALSE)**		
18	10	0.0008	**=POISSON.DIST(A18, E4, FALSE)**		
19	11	0.0002	**=POISSON.DIST(A19, E4, FALSE)**		
20	12	0.0001	**=POISSON.DIST(A20, E4, FALSE)**		
21	13	0.0000	**=POISSON.DIST(A21, E4, FALSE)**		
22	14	0.0000	**=POISSON.DIST(A22, E4, FALSE)**		
23	15	0.0000	**=POISSON.DIST(A23, E4, FALSE)**		

Deciding between Binomial and Poisson distribution:

- 1. The meaning of X:

Binomial variable = number of “successes” in n trials

Poisson variable = number of events that occur unexpectedly at random times

Deciding between Binomial and Poisson distribution:

2. Possible values of X.

Binomial X = 0 through n

Poisson X = from 0 to infinity

3. Parameters of the distribution:

Binomial: n = number of trials, pi = probability of success;

Poisson: lambda = frequency of events

Chapter Summary

In this chapter we discussed

- The probability distribution of a discrete random variable
- Expectation, variance, and standard deviation of a random variable
- Evaluating the risk of a portfolio
- The Binomial distribution
- The Poisson distribution