IE 440

PROCESS IMPROVEMENT THROUGH PLANNED EXPERIMENTATION

Some Important Discrete Probability Distributions

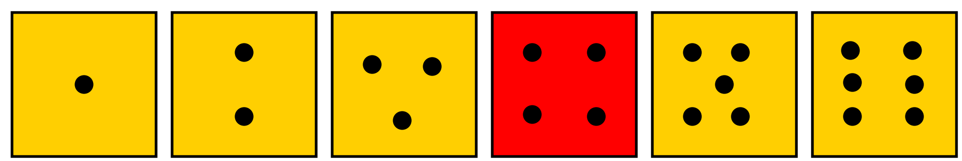
Dr. Xueping Li
University of Tennessee

Chapter Topics

- The Probability of a Discrete Random Variable
- Covariance and Its Applications in Finance
- Binomial Distribution
- Poisson Distribution
- Hypergeometric Distribution

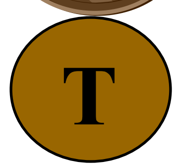
Random Variable

- Random Variable
  - Outcomes of an experiment expressed numerically
  - E.g., Toss a die twice; count the number of times the number 4 appears (0, 1 or 2 times)



  - E.g., Toss a coin; assign $10 to head and -$30 to a tail

= $10



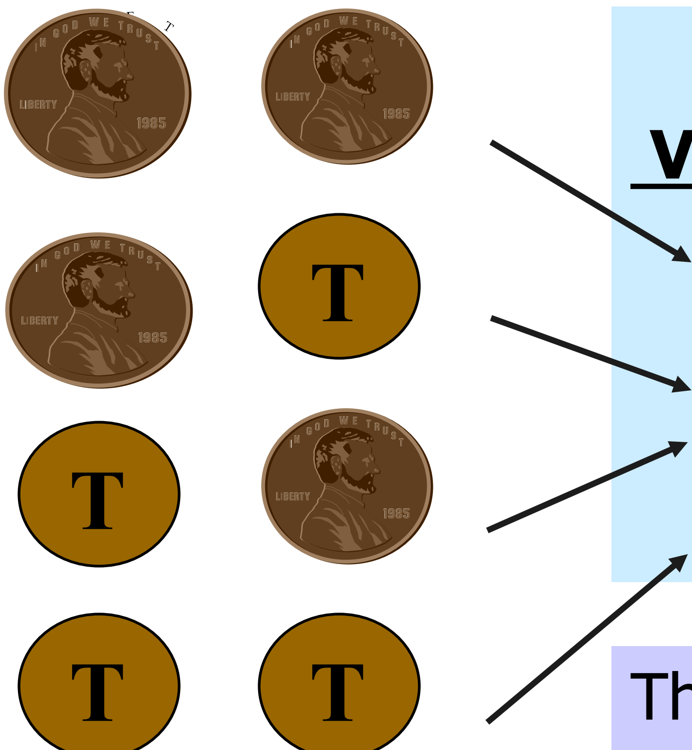
= -$30

Discrete Random Variable

- Discrete Random Variable
  - Obtained by counting (0, 1, 2, 3, etc.)
  - Usually a finite number of different values
  - E.g., Toss a coin 5 times; count the number of tails (0, 1, 2, 3, 4, or 5 times)

Discrete Probability Distribution Example

Event: Toss 2 Coins   Count # Tails



Probability Distribution

Values	Probability
0	1/4 = .25
1	2/4 = .50
2	1/4 = .25

This is using the *A Priori Classical Probability* approach.

Discrete Probability Distribution

- List of All Possible [X_j , $P(X_j)$] Pairs
  - X_j = Value of random variable
  - $P(X_j)$ = Probability associated with value
- Mutually Exclusive (Nothing in Common)
- Collective Exhaustive (Nothing Left Out)

$$0 \le P\left(X_j\right) \le 1 \qquad \sum P\left(X_j\right) = 1$$

Summary Measures

- Expected Value (The Mean)
  - Weighted average of the probability distribution
  - $$\mu = E(X) = \sum_j X_j P(X_j)$$
  - E.g., Toss 2 coins, count the number of tails, compute expected value:

$$\mu = \sum_j X_j P(X_j)$$

$$= (0)(.25) + (1)(.5) + (2)(.25) = 1$$

Summary Measures

(continued)

- Variance
  - Weighted average squared deviation about the mean
  - $$\sigma^2 = E\left[(X-\mu)^2\right] = \sum (X_j - \mu)^2 P(X_j)$$
  - E.g., Toss 2 coins, count number of tails, compute variance:

$$\sigma^2 = \sum (X_j - \mu)^2 P(X_j)$$
$$= (0-1)^2(.25) + (1-1)^2(.5) + (2-1)^2(.25)$$
$$= .5$$

Covariance and Its Application

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

X : discrete random variable

X_i : i^{th} outcome of X

Y : discrete random variable

Y_i : i^{th} outcome of Y

$P(X_i Y_i)$: probability of occurrence of the i^{th} outcome of X and the i^{th} outcome of Y

Computing the Mean for Investment Returns

Return per $1,000 for two types of investments

			Investment	
P(X_i)	**P(Y_i)**	**Economic Condition**	**Dow Jones Fund *X***	**Growth Stock *Y***
.2	.2	Recession	-$100	-$200
.5	.5	Stable Economy	+ 100	+ 50
.3	.3	Expanding Economy	+ 250	+ 350

$$E(X) = \mu_X = (-100)(.2) + (100)(.5) + (250)(.3) = \$105$$

$$E(Y) = \mu_Y = (-200)(.2) + (50)(.5) + (350)(.3) = \$90$$

Computing the Variance for Investment Returns

			Investment	
P(X_i)	***P*(Y_i)**	**Economic Condition**	**Dow Jones Fund *X***	**Growth Stock *Y***
.2	.2	Recession	-$100	-$200
.5	.5	Stable Economy	+ 100	+ 50
.3	.3	Expanding Economy	+ 250	+ 350

$$\sigma_X^2 = (.2)(-100-105)^2 + (.5)(100-105)^2 + (.3)(250-105)^2$$
$$= 14,725 \qquad \sigma_X = 121.35$$

$$\sigma_Y^2 = (.2)(-200-90)^2 + (.5)(50-90)^2 + (.3)(350-90)^2$$
$$= 37,900 \qquad \sigma_Y = 194.68$$

Computing the Covariance for Investment Returns

		Investment	
$P(X_iY_i)$	**Economic Condition**	**Dow Jones Fund X**	**Growth Stock Y**
.2	Recession	-$100	-$200
.5	Stable Economy	+ 100	+ 50
.3	Expanding Economy	+ 250	+ 350

$$\sigma_{XY} = (-100-105)(-200-90)(.2) + (100-105)(50-90)(.5) + (250-105)(350-90)(.3) = 23{,}300$$

The covariance of 23,000 indicates that the two investments are positively related and will vary together in the same direction.

Computing the Coefficient of Variation for Investment Returns

- $$CV(X) = \frac{\sigma_X}{\mu_X} = \frac{121.35}{105} = 1.16 = 116\%$$
- $$CV(Y) = \frac{\sigma_Y}{\mu_Y} = \frac{194.68}{90} = 2.16 = 216\%$$
- Investment *X* appears to have a lower risk (variation) per unit of average payoff (return) than investment *Y*
- Investment *X* appears to have a higher average payoff (return) per unit of variation (risk) than investment *Y*

Sum of Two Random Variables

- The expected value of the sum is equal to the sum of the expected values

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

- The variance of the sum is equal to the sum of the variances plus twice the covariance

$$Var(X+Y)=\sigma_{X+Y}^2=\sigma_X^2+\sigma_Y^2+2\sigma_{XY}$$

- The standard deviation is the square root of the variance

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

Portfolio Expected Return and Risk

- The portfolio expected return for a two-asset investment is equal to the weighted average of the two assets

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

where

$w =$ portion of the portfolio value assigned to asset X

- Portfolio risk

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

Computing the Expected Return and Risk of the Portfolio Investment

		Investment	
$P(X_i Y_i)$	**Economic Condition**	**Dow Jones Fund X**	**Growth Stock Y**
.2	Recession	-\$100	-\$200
.5	Stable Economy	+ 100	+ 50
.3	Expanding Economy	+ 250	+ 350

Suppose a portfolio consists of an equal investment in each of X and Y:

$$E(P) = 0.5(105) + 0.5(90) = 97.5$$

$$\sigma_P = \sqrt{(0.5)^2(14725) + (0.5)^2(37900) + 2(0.5)(0.5)(23300)} = 157.5$$

Doing It in PHStat

- PHStat | Decision Making | Covariance and Portfolio Analysis
  - Fill in the "Number of Outcomes:"
  - Check the "Portfolio Management Analysis" box
  - Fill in the probabilities and outcomes for investment *X* and *Y*
  - Manually compute the CV using the formula in the previous slide
- Here is the Excel spreadsheet that contains the results of the previous investment example:

Microsoft Excel Worksheet

Important Discrete Probability Distributions

- **Discrete Probability Distributions**
  - **Binomial**
  - **Hypergeometric**
  - **Poisson**

Binomial Probability Distribution

- 'n' Identical Trials
  - E.g., 15 tosses of a coin; 10 light bulbs taken from a warehouse
- 2 Mutually Exclusive Outcomes on Each Trial
  - E.g., Heads or tails in each toss of a coin; defective or not defective light bulb
- Trials are Independent
  - The outcome of one trial does not affect the outcome of the other

Binomial Probability Distribution

(continued)

- Constant Probability for Each Trial
  - E.g., Probability of getting a tail is the same each time we toss the coin
- 2 Sampling Methods
  - Infinite population without replacement
  - Finite population with replacement

Binomial Probability Distribution Function

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

$P(X)$: probability of X successes given n and p

X: number of "successes" in sample $(X=0,1,\cdots,n)$

p: the probability of each "success"

n: sample size

Tails in 2 Tosses of Coin

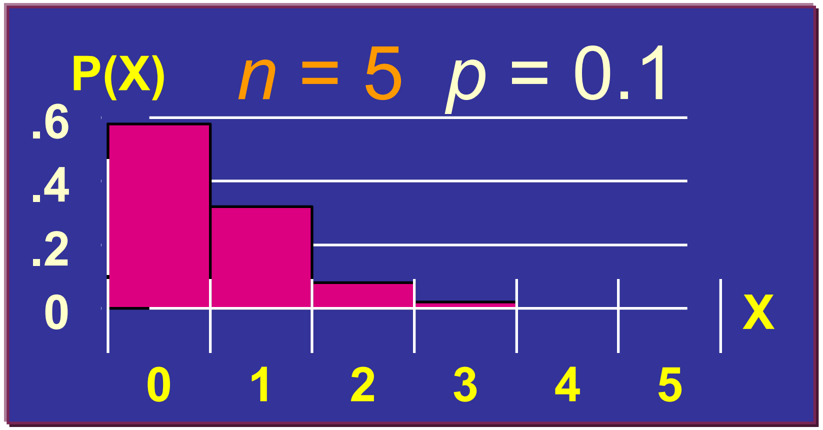
X	P(X)
0	1/4 = .25
1	2/4 = .50
2	1/4 = .25

Binomial Distribution Characteristics

- Mean
  - $\mu = E(X) = np$
  - E.g., $\mu = np = 5(.1) = .5$
- Variance and Standard Deviation
  - $\sigma^2 = np(1-p)$

    $\sigma = \sqrt{np(1-p)}$
  - E.g.,

$$\sigma = \sqrt{np(1-p)} = \sqrt{5(.1)(1-.1)} = .6708$$



Binomial Distribution in PHStat

- PHStat | Probability & Prob. Distributions | Binomial
- Example in Excel Spreadsheet

Microsoft Excel
Worksheet

Example: Binomial Distribution

A mid-term exam has 30 multiple choice questions, each with 5 possible answers. What is the probability of randomly guessing the answer for each question and passing the exam (i.e., having guessed at least 18 questions correctly)?

Are the assumptions for the binomial distribution met?

Microsoft Excel Worksheet

Yes, the assumptions are met. Using results from PHStat:

$$n = 30 \qquad p = 0.2$$

$$P(X \geq 18) = 1.84245(10)^{-6}$$

Poisson Distribution

Siméon Poisson

Poisson Distribution

- Discrete events ("successes") occurring in a given area of opportunity ("interval")
  - "Interval" can be time, length, surface area, etc.
- The probability of a "success" in a given "interval" is the same for all the "intervals"
- The number of "successes" in one "interval" is independent of the number of "successes" in other "intervals"
- The probability of two or more "successes" occurring in an "interval" approaches zero as the "interval" becomes smaller
  - E.g., # customers arriving in 15 minutes
  - E.g., # defects per case of light bulbs

Poisson Probability Distribution Function

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

$P(X)$: probability of X "successes" given λ

X: number of "successes" per unit

λ: expected (average) number of "successes"

e: 2.71828 (base of natural logs)

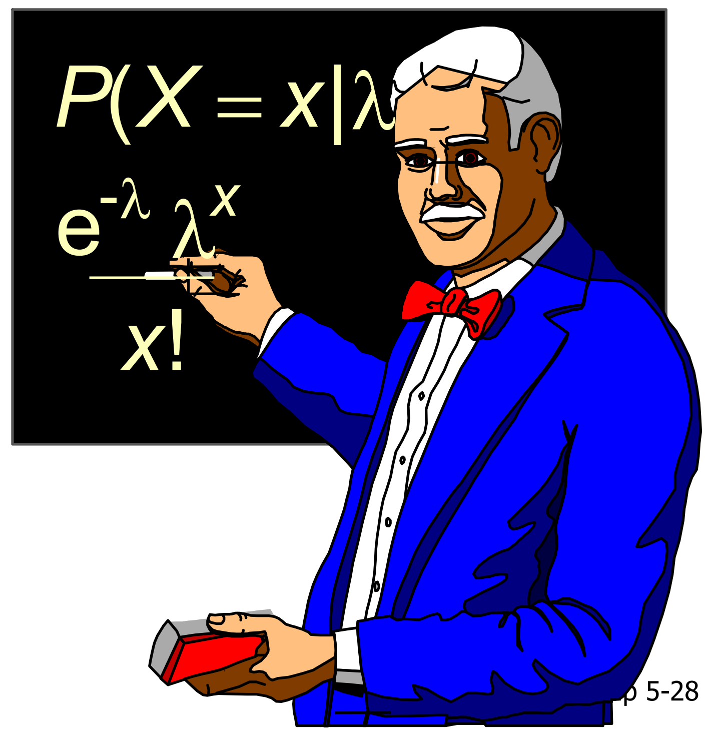
E.g., Find the probability of 4 customers arriving in 3 minutes when the mean is 3.6.

$$P(X) = \frac{e^{-3.6}3.6^{4}}{4!} = .1912$$

Poisson Distribution in PHStat

- PHStat | Probability & Prob. Distributions | Poisson
- Example in Excel Spreadsheet

Microsoft Excel Worksheet

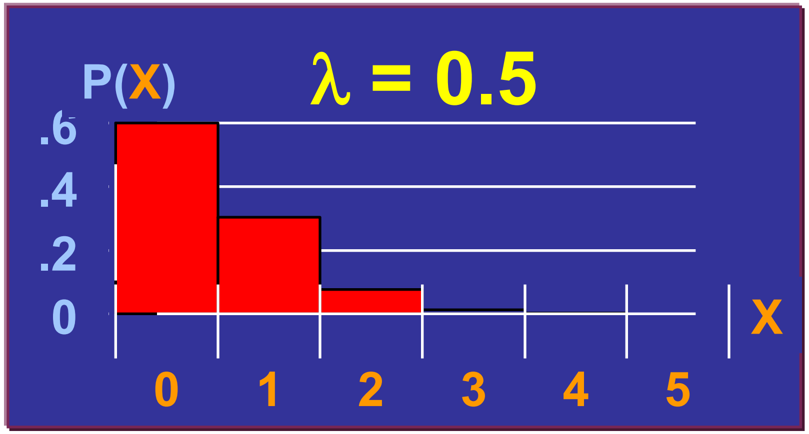


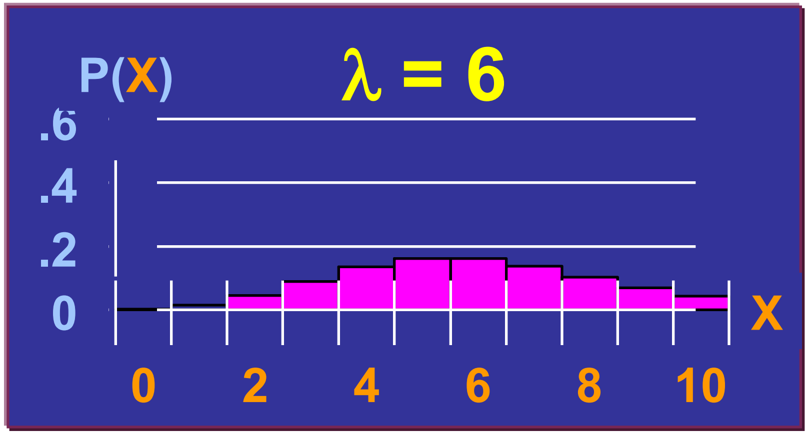
Poisson Distribution Characteristics

- Mean
  - $\mu = E(X) = \lambda$

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

- Standard Deviation and Variance
  - $\sigma^2 = \lambda \qquad \sigma = \sqrt{\lambda}$





Hypergeometric Distribution

- "n" Trials in a Sample Taken from a Finite Population of Size N
- Sample Taken Without Replacement
- Trials are Dependent
- Concerned with Finding the Probability of "X" Successes in the Sample Where There are "A" Successes in the Population

Hypergeometric Distribution Function

$$P(X) = \frac{\binom{A}{X}\binom{N-A}{n-X}}{\binom{N}{n}}$$

E.g., 3 Light bulbs were selected from 10. Of the 10, there were 4 defective. What is the probability that 2 of the 3 selected are defective?

$P(X)$: probability that X successes given $n, N,$ and A

n: sample size

N: population size

A: number of "successes" in population

X: number of "successes" in sample

$(X = 0, 1, 2, \cdots, n)$

$$P(2) = \frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}} = .30$$

Hypergeometric Distribution Characteristics

- Mean
  - $$\mu = E(X) = n\frac{A}{N}$$
- Variance and Standard Deviation
  - $$\sigma^2 = \frac{nA(N-A)}{N^2}\frac{N-n}{N-1}$$
    $$\sigma = \sqrt{\frac{nA(N-A)}{N^2}}\sqrt{\frac{N-n}{N-1}}$$

Finite Population Correction Factor

Hypergeometric Distribution in PHStat

- PHStat | Probability & Prob. Distributions | Hypergeometric …
- Example in Excel Spreadsheet

Microsoft Excel Worksheet

Chapter Summary

- Addressed the Probability of a Discrete Random Variable
- Defined Covariance and Discussed Its Application in Finance
- Discussed Binomial Distribution
- Addressed Poisson Distribution
- Discussed Hypergeometric Distribution