Statistical Analysis in Fin Mkts

MSF 502

Li Cai

5 Discrete Probability Distributions

C H A P T E R

Chapter 5 Learning Objectives (LOs)

LO 5.1: Distinguish between discrete and continuous random variables.

LO 5.2: Describe the probability distribution of a discrete random variable.

LO 5.3: Calculate and interpret summary measures for a discrete random variable.

LO 5.4: Differentiate among risk neutral, risk averse, and risk loving consumers.

LO 5.5: Compute summary measures to evaluate portfolio returns.

Chapter 5 Learning Objectives (LOs)

LO 5.6: Describe the binomial distribution and compute relevant probabilities.

LO 5.7: Describe the Poisson distribution and compute relevant probabilities.

5.1 Random Variables and Discrete Probability Distributions

LO 5.1 Distinguish between discrete and continuous random variables.

- **Random variable**
  - **A function that assigns numerical values to the outcomes of a random experiment.**
  - **Denoted by uppercase letters (e.g., X).**

- **Values of the random variable are denoted by corresponding lowercase letters.**
  - **Corresponding values of the random variable: $x_1, x_2, x_3, \ldots$**

LO 5.1

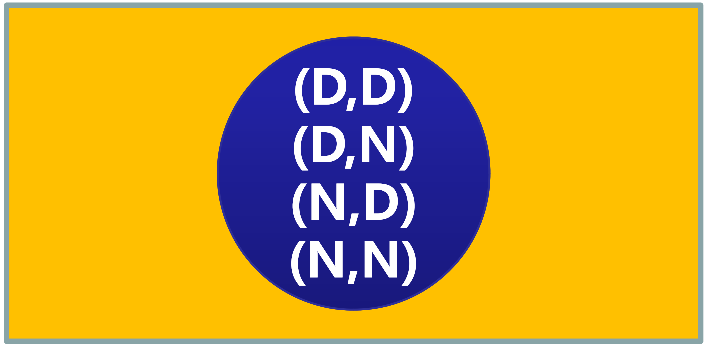
5.1 Random Variables and Discrete Probability Distributions

- **Random variables may be classified as:**
  - **Discrete**
    - **The random variable assumes a countable number of distinct values.**
  - **Continuous**
    - **The random variable is characterized by (infinitely) uncountable values within any interval.**

LO 5.1

5.1 Random Variables and Discrete Probability Distributions

- **Consider an experiment in which two shirts are selected from the production line and each can be defective (D) or non-defective (N).**
  - **Here is the sample space:**
  - **The random variable *X* is the number of defective shirts.**
  - **The possible number of defective shirts is the set {0, 1, 2}.**



- **Since these are the only possible outcomes, this is a *discrete* random variable.**

5.1 Random Variables and Discrete Probability Distributions

LO 5.2 Describe the probability distribution of a discrete random variable.

- **Every random variable is associated with a probability distribution that describes the variable completely.**
  - **A probability mass function is used to describe discrete random variables.**
  - **A probability density function is used to describe continuous random variables.**
  - **A cumulative distribution function may be used to describe both discrete and continuous random variables.**

LO 5.2

5.1 Random Variables and Discrete Probability Distributions

- **The probability mass function of a discrete random variable *X* is a list of the values of *X* with the associated probabilities, that is, the list of all possible pairs**

$$\left(x, P\left(X = x\right)\right)$$

- **The cumulative distribution function of *X* is defined as**

$$P\left(X \leq x\right)$$

LO 5.2

5.1 Random Variables and Discrete Probability Distributions

- **Two key properties of discrete probability distributions:**
  - **The probability of each value *x* is a value between 0 and 1, or equivalently**

$$0 \leq P(X = x) \leq 1$$

  - **The sum of the probabilities equals 1. In other words,**

$$\sum P(X = x_i) = 1$$

where the sum extends over all values *x* of *X*.

LO 5.2

5.1 Random Variables and Discrete Probability Distributions

- **A discrete probability distribution may be viewed as a table, algebraically, or graphically.**
- **For example, consider the experiment of rolling a six-sided die. A tabular presentation is:**

x	1	2	3	4	5	6
$P(X = x)$	1/6	1/6	1/6	1/6	1/6	1/6

- **Each outcome has an associated probability of 1/6. Thus, the pairs of values and their probabilities form the probability mass function for *X*.**

LO 5.2

5.1 Random Variables and Discrete Probability Distributions

- **Another tabular view of a probability distribution is based on the cumulative probability distribution.**
  - **For example, consider the experiment of rolling a six-sided die. The cumulative probability distribution is**

x	1	2	3	4	5	6
$P(X \leq x)$	1/6	2/6	3/6	4/6	5/6	6/6

- **The cumulative probability distribution gives the probability of *X* being less than or equal to *x*. For example,** $P(X \leq 4) = 4/6$

LO 5.2

5.1 Random Variables and Discrete Probability Distributions

- **A probability distribution may be expressed algebraically.**
- **For example, for the six-sided die experiment, the probability distribution of the random variable *X* is:**

$$P(X = x) = \begin{cases} 1/6 & \text{if } x = 1,2,3,4,5,6 \\ 0 & \text{otherwise} \end{cases}$$

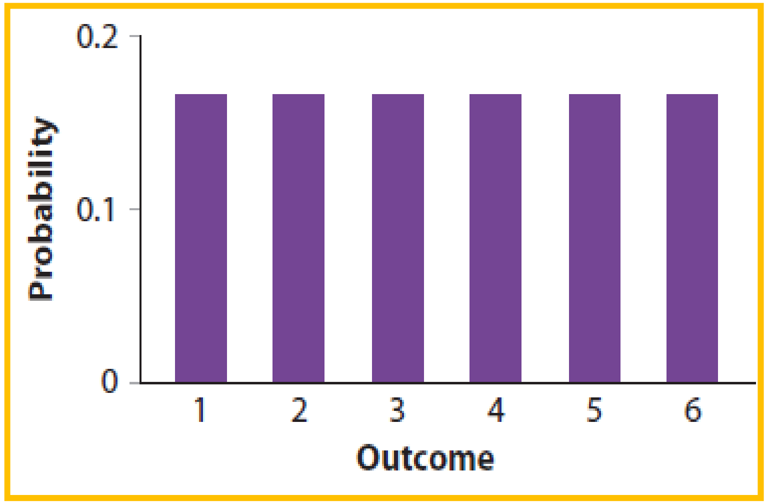
- **Using this formula we can find**

$$P(X = 5) = 1/6 \qquad P(X = 7) = 0$$

LO 5.2

5.1 Random Variables and Discrete Probability Distributions

- **A probability distribution may be expressed graphically.**
  - **The values *x* of *X* are placed on the horizontal axis and the associated probabilities on the vertical axis.**
  - **A line is drawn such that its height is associated with the probability of *x*.**
  - **For example, here is the graph representing the six-sided die experiment:**
  - **This is a uniform distribution since the bar heights are all the same.**



LO 5.2

5.1 Random Variables and Discrete Probability Distributions

- **Example: Consider the probability distribution which reflects the number of credit cards that Bankrate.com's readers carry:**

Number of Credit Cards	Percentage
0	2.5%
1	9.8
2	16.6
3	16.5
4*	54.6

*denotes 4 or more credit cards.

SOURCE: www.bankrate.com, Financial Literacy Series, 2007.

  - **Is this a valid probability distribution?**
  - **What is the probability that a reader carries no credit cards?**
  - **What is the probability that a reader carries less than two?**
  - **What is the probability that a reader carries at least two credit cards?**

LO 5.2

5.1 Random Variables and Discrete Probability Distributions

- **Consider the probability distribution which reflects the number of credit cards that Bankrate.com's readers carry:**
  - **Yes, because $0 \leq P(X = x) \leq 1$ and $\Sigma P(X = x) = 1$.**
  - **$P(X = 0) = 0.025$**
  - **$P(X < 2) = P(X = 0) + P(X = 1) = 0.025 + 0.098 = 0.123$.**
  - **$P(X \geq 2) = P(X = 2) + P(X = 3) + P(P = 4^*) = 0.166 + 0.165 + 0.546 = 0.877$. Alternatively, $P(X \geq 2) = 1 – P(X < 2) = 1 – 0.123 = 0.877$.**

Number of Credit Cards	Percentage
0	2.5%
1	9.8
2	16.6
3	16.5
4*	54.6

*denotes 4 or more credit cards.

SOURCE: www.bankrate.com, Financial Literacy Series, 2007.

5.2 Expected Value, Variance, and Standard Deviation

LO 5.3 Calculate and interpret summary measures for a discrete random variable.

- **Summary measures for a random variable include the**
  - **Mean (Expected Value)**
  - **Variance**
  - **Standard Deviation**

5.2 Expected Value, Variance, and Standard Deviation

LO 5.3

- **Expected Value** ⟺ **Population Mean**
  $E(X)$ ⟺ **μ**
  - **$E(X)$ is the long-run average value of the random variable over infinitely many independent repetitions of an experiment.**
  - **For a discrete random variable X with values $x_1, x_2, x_3, \ldots$ that occur with probabilities $P(X = x_i)$, the expected value of X is**

$$E(X) = \mu = \sum x_i P(X = x_i)$$

LO 5.3

5.2 Expected Value, Variance, and Standard Deviation

- **Variance and Standard Deviation**
  - **For a discrete random variable *X* with values $x_1, x_2, x_3, \ldots$ that occur with probabilities $P(X = x_i)$,**

$$Var(X) = \sigma^2 = \sum (x_i - \mu)^2 P(X = x_i)$$
$$= \sum x_i^2 P(X = x_i) - \mu^2$$

  - **The standard deviation is the square root of the variance.** $SD(X) = \sigma = \sqrt{\sigma^2}$

LO 5.3

5.2 Expected Value, Variance, and Standard Deviation

- **Example: Brad Williams, owner of a car dealership in Chicago, decides to construct an incentive compensation program based on performance.**

Bonus (in $1,000s)	Performance Type	Probability
$10	Superior	0.15
$6	Good	0.25
$3	Fair	0.40
$0	Poor	0.20

- **Calculate the expected value of the annual bonus amount.**
- **Calculate the variance and standard deviation of the annual bonus amount.**

LO 5.3

5.2 Expected Value, Variance, and Standard Deviation

- **Solution: Let the random variable *X* denote the bonus amount (in $1,000s) for an employee.**

Value, x_i	Probability, $P(X = x_i)$	Weighted Value, $x_i P(X = x_i)$	Weighted Squared Deviation, $(x_i - \mu)^2 P(X = x_i)$
10	0.15	$10 \times 0.15 = 1.5$	$(10 - 4.2)^2 \times 0.15 = 5.05$
6	0.25	$6 \times 0.25 = 1.5$	$(6 - 4.2)^2 \times 0.25 = 0.81$
3	0.40	$3 \times 0.40 = 1.2$	$(3 - 4.2)^2 \times 0.40 = 0.58$
0	0.20	$0 \times 0.20 = 0$	$(0 - 4.2)^2 \times 0.20 = 3.53$
		Total = 4.2	Total = 9.97

- $E(X) = \mu = \Sigma x_i P(X = x_i) = 4.2$ **or $4,200**
- $Var(X) = \sigma^2 = \Sigma(x_i - \mu)^2 P(X = x_i) = 9.97$ **(in $1,000s)²**.
- $SD(X) = \sqrt{\sigma^2} = \sqrt{9.97} = 3.158$ or $3,158.

5.2 Expected Value, Variance, and Standard Deviation

LO 5.4 Differentiate among risk neutral, risk averse, and risk loving consumers.

- **Risk Neutrality and Risk Aversion**
  - **Risk averse consumers:**
    - **Expect a reward for taking a risk.**
    - **May decline a risky prospect even if it offers a positive expected gain.**
  - **Risk neutral consumers:**
    - **Completely ignore risk.**
    - **Always accept a prospect that offers a positive expected gain.**

LO 5.4

5.2 Expected Value, Variance, and Standard Deviation

- **Risk Neutrality and Risk Aversion**
  - **Risk loving consumers:**
    - **May accept a risky prospect even if the expected gain is negative.**
- **Application of Expected Value to Risk**
  - **Suppose you have a choice of receiving $1,000 in cash or receiving a beautiful painting from your grandmother.**
  - **The actual value of the painting is uncertain. Here is a probability distribution of the possible worth of the painting. What should you do?**

x	$P(X = x)$
$2,000	0.20
$1,000	0.50
$500	0.30

LO 5.4

5.2 Expected Value, Variance, and Standard Deviation

- **Application of Expected Value to Risk**
  - **First calculate the expected value:**

x	$P(X = x)$
\$2,000	0.20
\$1,000	0.50
\$500	0.30

$$E(X) = \sum x_i P(X = x_i)$$
$$= \$2{,}000 \times 0.20 + \$1{,}000 \times 0.50 + \$500 \times 0.30$$
$$= \$1{,}050$$

  - **Since the expected value is more than \$1,000 it may seem logical to choose the painting over cash.**
  - **However, we have not taken into account risk.**

5.3 Portfolio Returns

LO 5.5 Compute summary measures to evaluate a portfolio returns.

- **Investment opportunities often use both:**
  - **Expected return as a measure of reward.**
  - **Variance or standard deviation of return as a measure of risk.**
- **Portfolio is defined as a collection of assets such as stocks and bonds.**
  - **Let *X* and *Y* represent two random variables of interest, denoting, say, the returns of two assets.**
  - **Since an investor may have invested in both assets, we would like to evaluate the portfolio return formed by a linear combination of *X* and *Y* .**

LO 5.5

5.3 Portfolio Returns

- **Properties of random variables useful in evaluating portfolio returns.**
  - **Given two random variables *X* and *Y*,**
    - **The expected value of *X* and *Y* is**

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

    - **The variance of *X* and *Y* is**

$$Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$$

where *Cov*(*X*,*Y*) is the covariance between *X* and *Y*.

    - **For constants *a*, *b*, the formulas extend to**

$$E(aX+bY)=aE(X)+bE(Y)$$

$$Var(aX+bY)=a^2Var(X)+b^2Var(Y)+2abCov(X,Y)$$

LO 5.5

5.3 Portfolio Returns

- **Expected return, variance, and standard deviation of portfolio returns.**
  - **Given a portfolio with two assets, Asset *A* and Asset *B*, the expected return of the portfolio $E(R_p)$ is computed as:**

$$E\left(R_p\right) = w_A E\left(R_A\right) + w_B E\left(R_B\right)$$

- **where**
  w_A and w_B are the portfolio weights
  $w_A + w_B = 1$
  $E(R_A)$ and $E(R_B)$ are the expected returns on assets *A* and *B*, respectively.

LO 5.5

5.3 Portfolio Returns

- **Expected return, variance, and standard deviation of portfolio returns.**
  - **Using the covariance or the correlation coefficient of the two returns, the portfolio variance of return is:**

$$Var\left(R_p\right) = w_A^{\,2}\sigma_A^{\,2} + w_B^{\,2}\sigma_B^{\,2} + 2w_A w_B \rho_{AB}\sigma_A\sigma_B$$

where σ^2_A and σ^2_B are the variances of the returns for Asset *A* and Asset *B*, respectively,
σ_{AB} is the covariance between the returns for Assets *A* and *B*
ρ_{AB} is the correlation coefficient between the returns for Asset *A* and Asset *B*.

LO 5.5

5.3 Portfolio Returns

- **Example: Consider an investment portfolio of \$40,000 in Stock A and \$60,000 in Stock B.**
  - **Given the following information, calculate the expected return of this portfolio.**

Stock A	Stock B
$E(R_A) = \mu_A = 9.5\%$	$E(R_B) = \mu_B = 7.6\%$
$SD(R_A) = \sigma_A = 12.93\%$	$SD(R_B) = \sigma_B = 8.20\%$
$Cov(R_A, R_B) = \sigma_{AB} = 18.60\%$	

  - **SOLUTION:** First we compute the portfolio weights. Since \$40,000 is invested in Stock A and \$60,000 in Stock B, we compute

$$w_A = \frac{40,000}{100,000} = 0.40 \quad \text{and} \quad w_B = \frac{60,000}{100,000} = 0.60.$$

Thus, using the formula for portfolio expected return, we solve:

$$E(R_p) = (0.40 \times 9.5\%) + (0.60 \times 7.6\%) = 3.80\% + 4.56\% = 8.36\%.$$

LO 5.5

5.3 Portfolio Returns

- **Example: Consider an investment portfolio of \$40,000 in Stock A and \$60,000 in Stock B.**

Stock A	Stock B
$E(R_A) = \mu_A = 9.5\%$	$E(R_B) = \mu_B = 7.6\%$
$SD(R_A) = \sigma_A = 12.93\%$	$SD(R_B) = \sigma_B = 8.20\%$
$Cov(R_A, R_B) = \sigma_{AB} = 18.60\%$	

- **Calculate the correlation coefficient between the returns on Stocks A and B.**

- **Solution:** $\rho_{AB} = \dfrac{\sigma_{AB}}{\sigma_A \sigma_B} = \dfrac{18.60}{12.93 \times 8.20} = 0.1754$

LO 5.5

5.3 Portfolio Returns

- **Example: Consider an investment portfolio of \$40,000 in Stock A and \$60,000 in Stock B.**

Stock A	Stock B
$E(R_A) = \mu_A = 9.5\%$	$E(R_B) = \mu_B = 7.6\%$
$SD(R_A) = \sigma_A = 12.93\%$	$SD(R_B) = \sigma_B = 8.20\%$
$Cov(R_A, R_B) = \sigma_{AB} = 18.60\%$	

- **Calculate the portfolio variance.**
- **Solution:**

$$
\begin{aligned}
Var(R_p) &= w_A^2\sigma_A^2 + w_B^2\sigma_B^2 + 2w_A w_B \sigma_{AB} \\
&= (0.40)^2(12.93)^2 + (0.60)^2(8.20)^2 + 2(0.40)(0.60)(18.60) \\
&= 26.75 + 24.21 + 8.93 \\
&= 59.89
\end{aligned}
$$

LO 5.5

5.3 Portfolio Returns

- **Example: Consider an investment portfolio of $40,000 in Stock A and $60,000 in Stock B.**

Stock A	Stock B
$E(R_A) = \mu_A = 9.5\%$	$E(R_B) = \mu_B = 7.6\%$
$SD(R_A) = \sigma_A = 12.93\%$	$SD(R_B) = \sigma_B = 8.20\%$
$Cov(R_A, R_B) = \sigma_{AB} = 18.60\%$	

  - **Calculate the portfolio standard deviation.**
  - **Solution:**

$$SD(R_p) = \sqrt{59.89} = 7.74, \text{ or } 7.74\%$$

5.4 The Binomial Probability Distribution

LO 5.6 Describe the binomial distribution and compute relevant probabilities.

- **A binomial random variable is defined as the number of successes achieved in the *n* trials of a Bernoulli process.**
  - **A Bernoulli process consists of a series of *n* independent and identical trials of an experiment such that on each trial:**
    - **There are only two possible outcomes:**
      p = probability of a success
      $1 - p = q$ = probability of a failure
    - **Each time the trial is repeated, the probabilities of success and failure remain the same.**

LO 5.6

5.4 The Binomial Probability Distribution

- **A binomial random variable *X* is defined as the number of successes achieved in the *n* trials of a Bernoulli process.**
- **A binomial probability distribution shows the probabilities associated with the possible values of the binomial random variable (that is, 0, 1, . . . , *n*).**
  - **For a binomial random variable *X* , the probability of *x* successes in *n* Bernoulli trials is**

$$P(X=x)=\binom{n}{x}p^x q^{n-x}=\frac{n!}{x!(n-x)!}p^x q^{n-x}$$

for $x=0,1,2,\ldots,n$. By definition, $0!=1$.

LO 5.6

5.4 The Binomial Probability Distribution

- **For a binomial distribution:**
  - **The expected value (*E*(*X*)) is:** $E(X) = \mu = np$
  - **The variance (*Var*(*X*)) is:** $Var(X) = \sigma^2 = npq$
  - **The standard deviation (*SD*(*X*)) is:** $SD(X) = \sigma = \sqrt{npq}$

LO 5.6

5.4 The Binomial Probability Distribution

- **Example: Approximately 20% of U.S. workers are afraid that they will never be able to retire. Suppose 10 workers are randomly selected.**
  - **What is the probability that none of the workers is afraid that they will never be able to retire?**
    - **Solution: Let *X* = 10, then**

$$P(X=0)=\frac{10!}{0!(10-0)!}\times(0.20)^0\times(0.80)^{10}$$

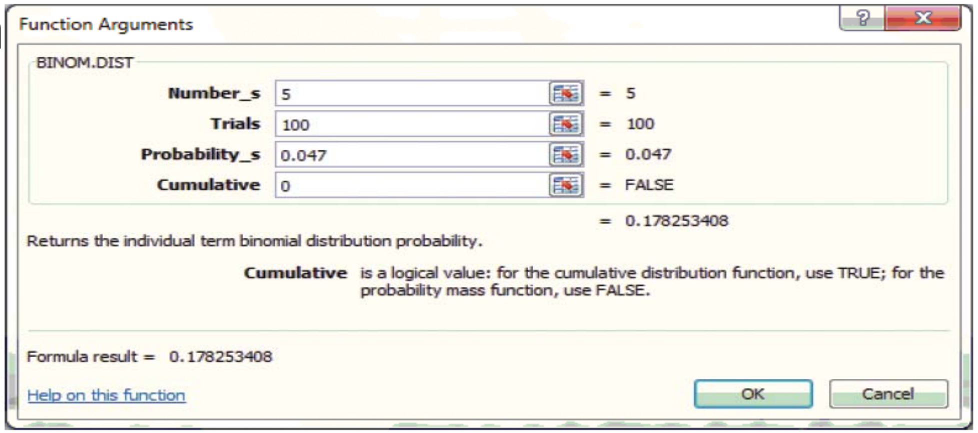
$$=\frac{10\times 9\times\cdots\times 1}{(1)\times(10\times 9\times\cdots\times 1)}\times 1\times(0.80)^{10}=1\times 1\times 0.1074$$

$$=0.1074.$$

LO 5.6

5.4 The Binomial Probability Distribution

- **Computing binomial probabilities with Excel:**
  - **In 2007 approximately 4.7% of the households in the Detroit metropolitan area were in some stage of foreclosure. What is the probability that exactly 5 of these 100 mortgage-holding households in Detroit are in some stage of foreclosure?**
    - **Solution: Using the binomial function on Excel, enter the four arguments shown**
    - **Excel returns the formula result as 0.1783; thus, $P(X = 5) = 0.1783$.**



5.5 The Poisson Probability Distribution

LO 5.7 Describe the Poisson distribution and compute relevant probabilities.

- **A binomial random variable counts the number of successes in a fixed number of Bernoulli trials.**
- **In contrast, a Poisson random variable counts the number of successes over a given interval of time or space.**
- **Examples of a Poisson random variable include**
  - ***With respect to time*—the number of cars that cross the Brooklyn Bridge between 9:00 am and 10:00 am on a Monday morning.**
  - ***With respect to space*—the number of defects in a 50-yard roll of fabric.**

LO 5.7

5.5 The Poisson Probability Distribution

- **A random experiment satisfies a Poisson process if:**
  - **The number of successes within a specified time or space interval equals any integer between zero and infinity.**
  - **The numbers of successes counted in nonoverlapping intervals are independent.**
  - **The probability that success occurs in any interval is the same for all intervals of equal size and is proportional to the size of the interval.**

LO 5.7

5.5 The Poisson Probability Distribution

- **For a Poisson random variable *X*, the probability of *x* successes over a given interval of time or space is**

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

where μ is the mean number of successes and $e \approx 2.718$ is the base of the natural logarithm.

LO 5.7

5.5 The Poisson Probability Distribution

- **For a Poisson distribution:**
  - **The expected value (*E*(*X*)) is:** $E(X) = \mu$
  - **The variance (*Var*(*X*)) is:** $Var(X) = \sigma^2 = \mu$
  - **The standard deviation (*SD*(*X*)) is:** $SD(X) = \sigma = \sqrt{\mu}$

LO 5.7

5.5 The Poisson Probability Distribution

- **Example: Returning to the Starbucks example, Ann believes that the typical Starbucks customer averages 18 visits over a 30-day month.**
  - **How many visits should Anne expect in a 5-day period from a typical Starbucks customer?**

Given the rate of 18 visits over a 30-day month, we can write the mean for the 30-day period as $\mu_{30} = 18$. For this problem, we compute the proportional mean for a 5-day period as $\mu_5 = 3$ because $\frac{18 \text{ visits}}{30 \text{ days}} = \frac{3 \text{ visits}}{5 \text{ days}}$.

  - **What is the probability that a customer visits the chain five times in a 5-day period?**
    - $P(X = 5) = \frac{e^{-3}3^5}{5!} = \frac{(0.0498)(243)}{120} = 0.1008$

End of Chapter