Chapter 5

Discrete Random Variables

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.

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: $x_1, x_2, x_3, \ldots$
- 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.

- 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.

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.
- 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:

$$(x, P(X = x)) \tag{5.1}$$

- The cumulative distribution function of X is defined as

$$P(X \leq x) \tag{5.2}$$

- 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,

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

    where the sum extends over all values x_i of X.
- 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:

Table 5.1: Tabular representation of rolling a six-sided die.

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

- Each outcome has an associated probability of $\frac{1}{6}$. Thus, the pairs of values and their probabilities form the probability mass function for X.
- 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

Table 5.2: Tabular cumulative probability distribution of rolling a six-sided die.

x	1	2	3	4	5	6
$P(X = x)$	$\frac{1}{6}$	$\frac{2}{6}$	$\frac{3}{6}$	$\frac{4}{6}$	$\frac{5}{6}$	$\frac{6}{6}$

- The cumulative probability gives the probability of X being less than or equal to x. For example, $P(x \leq 4) = \frac{4}{6} = \frac{2}{3}$.
- A probability distribution may be expressed algebraically.
- For example, for the sid-sided die experiment, the probability distribution of the random variable X is:

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

- Using this formula, we can find

$$P(X = 5) = \frac{1}{6}$$
$$P(X = 7) = 0$$

- 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.
  - ...
  - This is a uniform distribution since the bar heights are all the same.

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

5.2.1 Expected Value

$$\text{Expected Value} \Leftrightarrow \text{Population Mean}$$
$$E(X) \Leftrightarrow \mu$$

- $E(X)$ 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

$$\begin{aligned} E(X) &= \mu \\ &= \sum_i x_i \times P(X = x_i) \end{aligned} \tag{5.3}$$

5.2.2 Variance and Standard Deviation

- For a discrete random variable X with values $x_1, x_2, x_3, \ldots$ that occur with probabilities $P(X = x)$,

$$\begin{aligned} \mathrm{Var}(X) = \sigma^2 &= \sum_i (x_i - \mu)^2 P(X = x_i) \\ &= \sum_i x_i^2 P(X = x_i) - \mu_2 \end{aligned} \tag{5.4}$$

- The standard deviation is the square root of the variance.

$$\mathrm{SD}(X) = \sigma = \sqrt{\sigma^2} \tag{5.5}$$

5.2.3 Risk Neutrality and Risk Aversion

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

- Risk average 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.

- Risk loving consumers:
  - May accept a risky prospect even if the expected gain is negative.

5.2.4 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?

Table 5.3: Painting Value Probabilities

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

5.3 Portfolio Returns

LO 5.5 **Compute summary measures to evaluate a portfolio's return.**

- 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 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.

5.3.1 Properties of random variables useful in evaluating portfolio returns

- Given two random variables X and Y,
  - The expected value of X and Y is
$$E(X+Y) = E(X) + E(Y) \tag{5.6}$$
  - The variance of X and Y is
$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y) \tag{5.7}$$
    where $\text{Cov}(X,Y)$ is the covariance between X and Y.
  - For constants a, b, the formulas extend to
$$E(aX+bY) = aE(X) + bE(Y)$$
$$\text{Var}(aX+bY) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X,Y)$$

5.3.2 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:
$$E(R_p) = w_A E(R_A) + w_B E(R_B) \tag{5.8}$$
  where w_A and w_B are the portfolio weights, $w_A + w_B = 1$, and $E(R_A)$ and $E(R_B)$ are the expected returns on assets A and B, respectively.
- Using the covariance or the correlation coefficient of the two returns, the portfolio variance of return is:
$$\text{Var}(R_p) = w_A^2\sigma_A^2 + w_B^2\sigma_B^2 + 2w_A w_B \rho_{AB}\sigma_A\sigma_B \tag{5.9}$$
  where σ_A^2 and σ_B^2 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, and ρ_{AB} is the correlation coefficient between the returns for Asset A and Asset B.
$$\rho_{AB} = \frac{\sigma_{AB}}{\sigma_A\sigma_B} \tag{5.10}$$

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.
- 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:

$$\begin{aligned} P(X = x) &= \binom{n}{x} p^x q^{n-x} \\ &= \frac{n!}{(n-x)!x!} p^x q^{n-x} \end{aligned} \tag{5.11}$$

    for $x = 0, 1, 2, \ldots, n$.
- For a binomial distribution:
  - The expected value $E(X)$ is:

$$E(X) = \mu = np \tag{5.12}$$

  - The variance $\mathrm{Var}(X)$ is:

$$\mathrm{Var}(X) = \sigma^2 = npq \tag{5.13}$$

  - The standard deviation $\mathrm{SD}(X)$ is:

$$\mathrm{SD}(X) = \sigma = \sqrt{npq} \tag{5.14}$$

5.5 The Poisson Probability Distribution

LO 5.7 **Descrive 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-year roll of fabric.
- A random experiment satisfies a Poisson process if:
  - The number of successes within a specified time or space interval equals any integer between 0 and ∞.
  - The number of successes in non-overlapping intervals are independent.
  - The probability that successes occurs in any interval is the same for all intervals of equal size and is proportional to the size of the interval.
- For a Poisson random variable X, the probability of x successes over a given interval of time or space is:

$$P(X = x) = \frac{e^{-\mu}\mu^x}{x!} \text{ for } x = 0, 1, 2, \ldots \tag{5.15}$$

  where μ is the mean number of successes and $e \approx 2.718$ is the base of the natural logarithm.
- For a Poisson distribution:
  - The expected value $E(X)$ is:

$$E(X) = \mu \tag{5.16}$$

  - The variance $\mathrm{Var}(X)$ is:

$$\mathrm{Var}(X) = \sigma^2 = \mu \tag{5.17}$$

  - The standard deviation $\mathrm{SD}(X)$ is:

$$\mathrm{SD}(X) = \sigma = \sqrt{\mu} \tag{5.18}$$

Chapter 6

Continuous Random Variables

LO 6.1: Describe a continuous random variable.

LO 6.2: Describe a continuous uniform distribution and calculate associated probabilities.

LO 6.3: Explain the characteristics of the normal distribution.

LO 6.4: Use the standard normal table of the z-table.

LO 6.5: **Calculate and interpret probabilities or a random variable that follows the normal distribution.**

LO 6.6: **Calculate and interpret probabilities or a random variable that follows the exponential distribution.**

LO 6.7: **Calculate and interpret probabilities or a random variable that follows the lognormal distribution.**

6.1 Continuous Random Variables and the Uniform Probability Distribution

LO 6.1 **Describe a continuous random variable.**

- Remember that 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.

- When computing probabilities for a continuous random variable, keep in mind that $P(X = x) = 0$.
  - We cannot assign a nonzero probability to each infinitely uncountable value and still have the probabilities sum to one.

– Thus, since $P(X = a)$ and $P(X = b)$ both equal zero, the following holds true for continuous random variables:

$$P(a \leq X \leq b) = P(a < X < b) = P(a \leq X < b) = P(a < X \leq b)$$

6.1.1 Probability Density Function $f(x)$ of a continuous random variable X

- Describes the relative likelihood that X assumes a value within a general interval (e.g., $P(a \leq X \leq b)$), where
  – $f(x) > 0$ for all possible values of X.
  – The area under $f(x)$ over all values of x equals 1.

6.1.2 Cumulative Density Function $F(x)$ of a continuous random variable X

- For any value x of the random variable X, the cumulative distribution function $F(x)$ is computes as:

$$F(X) = P(X \leq x)$$

- As a result:

$$P(a \leq X \leq b) = F(b) - F(a)$$

6.1.3 The Continuous Uniform Distribution

LO 6.2 **Describe a continuous uniform distribution and calculate associated probabilities.**

- Describe a random variable that has an equally likely chance of assuming a value within a specified range.
- Probability density function:

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x \leq b, \text{ and} \\ 0 & \text{for } x < a \text{ or } x > b \end{cases} \tag{6.1}$$

where a and b are the lower and upper limits, respectively.

- The expected value and standard deviation of X are:

$$E(X) = \mu = \frac{a+b}{2} \tag{6.2}$$

$$\begin{aligned} \text{SD}(X) &= \sigma \\ &= \sqrt{\frac{(b-a)^2}{12}} \end{aligned} \tag{6.3}$$

6.1.4 Graph of the continuous uniform distribution

- The values of a and b on the horizontal axis represent the lower and upper limits, respectively.
- The height of the distribution does not directly represent a probability.
- It is the area under $f(x)$ that corresponds to probability.

Cumulative function:

$$\begin{aligned} P(X > x_1) &= \text{base} \times \text{height} \\ &= (b - x_1) \times \frac{1}{b-a} \end{aligned}$$

6.2 The Normal Distribution

- For a random variable X with mean μ and variance σ^2:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{6.4}$$

6.2.1 The Normal Distribution

LO 6.3 **Explain the characteristics of the normal distribution.**

- Symmetric
- Bell-shaped
- Closely approximates the probability distribution of a wide range of random variables, such as the
  - Heights and weights of newborn babies
  - Scores on SAT
  - Cumulative debt of college graduates
- Serves as the cornerstone of statistical inference.

6.2.2 Characteristics of the Normal Distribution

- Symmetric about its mean
  - Mean = Median = Mode
- Asymptotic–that is, the tail gets closer and closer to the horizontal axis but never touches it.
- The normal distribution is completely described by two parameters: μ and σ^2.

  μ is the population mean which describes the central location of the distribution.
  σ^2 is the population variance which describes the dispersion of the distribution.

6.2.3 Probability Density Function of the Normal Distribution

- For a random variable X with mean μ and variance σ^2:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \tag{6.5}$$

6.2.4 The Standard Normal (Z) Distribution

LO 6.4 **Use the standard normal table of the z-table.**

- A special case of the normal distribution:
  - Mean μ is qual to zero ($E(X) = 0$).
  - Standard deviation σ is equal to 1 (SD(Z) = 1).

6.2.5 Standard Normal Table (Z-Table)

- Gives the cumulative probabilities $P(Z \leq z)$ for positive and negative values of z.
- Since the random variable Z is symmetric about its mean of 0,

$$P(Z < 0) = P(Z > 0) = 0.5$$

- To obtain the $P(Z < z)$, read down the z-column first, then across the top.

6.2.6 Finding the Probability for a Given z-Value

- Transform normally distributed random variables into standard normal random variables and use the z-table to compute the relevant probabilities.
- The z-table provides cumulative probabilities $P(Z \leq z)$ for a given z.

6.3 Solving Problems with the Normal Distribution

LO 6.5 **Calculate and interpret probabilities or a random variable that follows the normal distribution.**

6.3.1 The Normal Transformation

- Any normally distributed random variable X with mean μ and standard deviation σ can be transformed into the standard normal random variable Z as:

$$Z = \frac{X-\mu}{\sigma} \text{ with corresponding values } z = \frac{x-\mu}{\sigma} \tag{6.6}$$

- As constructed: $E(Z) = 0$ and $\mathrm{SD}(Z) = 1$.
- A z-value specifies by how many standard deviations the corresponding x value falls above ($z > 0$) or below ($z < 0$) the mean.
  - A positive z indicates by how many standard deviations the corresponding x lies above μ.
  - A zero z indicates that the corresponding x equals μ.
  - A negative z indicates by how many standard deviations the corresponding x lies below μ.

6.3.2 Use the Inverse Transformation to Compute Probabilities for Given x values

- A standard normal variable Z can be transformed to the normally distributed random variable X with mean μ and standard deviation σ as

$$X = \mu + Z\sigma \text{ with corresponding values } x = \mu + z\sigma \tag{6.7}$$

6.4 Other Continuous Probability Distributions

LO 6.6 Calculate and interpret probabilities or a random variable that follows the exponential distribution.

6.4.1 Exponential Distribution

- A random variable X follows the exponential distribution if its probability density function is:

$$f(x) = \lambda e^{-\lambda x} \text{ for } x \geq 0 \tag{6.8}$$

  where λ is the rate parameter and $E(X) = \mathrm{SD}(X) = \frac{1}{\lambda}$.
- The cumulative distribution function is:

$$P(X \leq x) = 1 - e^{-\lambda x} \tag{6.9}$$

6.4.2 The Lognormal Distribution

LO 6.7 Calculate and interpret probabilities or a random variable that follows the lognormal distribution.

- Defined for a positive random variable, the lognormal distribution is positively skewed.
- Useful for describing variables such as

- Income
- Real estate values
- Asset prices

- Failure rate may increase or decrease over time.
- Let X be a normally distributed random variable with mean μ and standard deviation σ. The random variable $Y = e^X$ follows the lognormal distribution with a probability density function as

$$f(y) = \frac{1}{y\sigma\sqrt{2\pi}} \exp\left(-\frac{(\ln(y) - \mu)^2}{2\sigma^2}\right) \text{ for } y > 0 \tag{6.10}$$

- The lognormal distribution is clearly positively skewed for $\sigma > 1$. For $\sigma < 1$, the lognormal distribution somewhat resembles to normal distribution.

6.4.3 Expected values and standard deviations of the lognormal and normal distributions

- Let X be a normal random variable with mean μ and standard deviation σ and let $Y = e^X$ by the corresponding lognormal variable. The mean μ_y and standard deviation σ_Y or Y are derived as:

$$\mu_Y = \exp\left(\frac{2\mu + \sigma^2}{2}\right) \tag{6.11}$$

$$\sigma_Y = \sqrt{(\exp(\sigma^2) - 1)\exp\left(2\mu + \sigma^2\right)} \tag{6.12}$$

- Equivalently, the mean and standard deviation of the normal variable $X = \ln(Y)$ are derived as

$$\mu = \ln\left(\frac{\mu_Y^2}{\sqrt{\mu_Y^2 + \sigma_Y^2}}\right) \tag{6.13}$$

$$\sigma = \sqrt{\ln\left(1 + \frac{\sigma_Y^2}{\mu_Y^2}\right)} \tag{6.14}$$