CHAPTER#03
Random Variables and Probability Distributions

Random Variable

- **Definition 3.1:**
  - A random variable is a function that associates a real number with each element in the sample space.

Random Variable

- **Example 3.1:**
  - Two balls are drawn in succession without replacement from an urn containing 4 red balls and 3 black balls.
  - The possible outcomes and the values y of the random variable Y , where Y is the number of red balls, are:

Sample Space	y
RR	2
RB	1
BR	1
BB	0

Random Variable

- **Example 3.1:**
  - A stockroom clerk returns three safety helmets at random to three steel mill employees who had previously checked them. If Smith, Jones, and Brown, in that order, receive one of the three hats, list the sample points for the possible orders of returning the helmets, and find the value m of the random variable M that represents the number of correct matches.

Random Variable

- **Example 3.1:**

Sample Space	m
SJB	3
SBJ	1
BJS	1
JSB	1
JBS	0
BSJ	0

Random Variable

- **Dummy Variable**
  - There are cases where the random variable is categorical in nature. Variables, often called dummy variables, are used. A good illustration is the case in which the random variable is binary in nature, as shown in the next example.

Random Variable

- **Dummy Variable: Example 3.3:**
  - Consider the simple condition in which components are arriving from the production line and they are stipulated to be defective or not defective. We can define the random variable X as follows:

$$X = \begin{cases} 1, & \text{if the component is defective,} \\ 0, & \text{if the component is not defective.} \end{cases}$$

  - The random variable for which 0 and 1 are chosen to describe the two possible values is called a Bernoulli random variable

Random Variable

- **Example 3.4:**
  - Statisticians use sampling plans to either accept or reject batches or lots of material. Suppose one of these sampling plans involves sampling independently 10 items from a lot of 100 items in which 12 are defective.
  - Let X as the random variable defined as the number of items found defective in the sample of 10.
  - In this case, the random variable takes on the values 0, 1, 2, . . . , 9, 10.

Random Variable

- **Example 3.5:**
  - Suppose a sampling plan involves sampling items from a process until a defective is observed. The evaluation of the process will depend on how many consecutive items are observed. In that regard, let X be a random variable defined by the number of items observed ~~before a defective is found.~~ With **N** a **nondefective** and **D** a **defective**, sample spaces are S = {D} given X = 1, S = {ND} given X = 2, S = {NND} given X = 3, and so on.

Discrete vs Continuous Sample Space

- **Definition 3.2: Discrete Sample Space**
  - If a sample space contains a finite number of possibilities or an unending sequence with as many elements as there are **whole numbers**, it is called a discrete sample space.
  - A random variable is called a discrete random variable if its set of possible outcomes is countable.
  - Achieve by counting.
- **Definition 3.3: Continuous Sample Space**
  - If a sample space contains an infinite number of possibilities equal to the number of points on a line segment, it is called a continuous sample space.
  - Continuous random variable is a random variable whose set of possible values is an entire interval of numbers is not discrete.
  - Achieve by measuring.

Discrete Probability Distributions

- A discrete random variable assumes each of its values with a certain probability.
- In the case of tossing a coin three times, let variable X representing the number of heads.
- The probability distribution of X will be:

x	0	1	2	3
P(X=x)	1/8	3/8	3/8	1/8

Probability Mass Function (PMF)

- **Definition 3.4:**

The set of ordered pairs $(x, f(x))$ is a **probability function**, **probability mass function**, or **probability distribution** of the discrete random variable X if, for each possible outcome x,

1. $f(x) \geq 0$,
2. $\sum_{x} f(x) = 1$,
3. $P(X = x) = f(x)$.

Probability Mass Function (PMF)

- **Example 3.8:**
  - A shipment of 20 similar laptop computers to a retail outlet contains 3 defectives. If a school makes a random purchase of 2 of these computers, find the **probability distribution** for the number of defectives.
- **Solution:**
  - Let X be a random variable whose values x are the possible numbers of

$$f(0) = P(X = 0) = \frac{\binom{3}{0}\binom{17}{2}}{\binom{20}{2}} = \frac{68}{95}, \quad f(1) = P(X = 1) = \frac{\binom{3}{1}\binom{17}{1}}{\binom{20}{2}} = \frac{51}{190},$$

$$f(2) = P(X = 2) = \frac{\binom{3}{2}\binom{17}{0}}{\binom{20}{2}} = \frac{3}{190}.$$

Thus, the probability distribution of X is

x	0	1	2
$f(x)$	$\frac{68}{95}$	$\frac{51}{190}$	$\frac{3}{190}$

Cumulative Distribution Function

– There are many problems where we may wish to compute the probability that the observed value of a random variable X will be **less than or equal** to some real number x. Writing **F(x) = P(X ≤ x)** for every real number x, we define F(x) to be the cumulative distribution function of the random variable X.

- **Definition 3.5:**

The **cumulative distribution function** $F(x)$ of a discrete random variable X with probability distribution $f(x)$ is

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t), \quad \text{for } -\infty < x < \infty.$$

Cumulative Distribution Function

- **Example:**
  - Let PDF is as follows:

m	0	1	3
$P(M=m)$	$\frac{1}{3}$	$\frac{1}{2}$	$\frac{1}{6}$

  - The CDF will be:

$$F(m)=\begin{cases} 0, & \text{for } m<0, \\ \frac{1}{3}, & \text{for } 0\le m<1, \\ \frac{5}{6}, & \text{for } 1\le m<3, \\ 1, & \text{for } m\ge 3. \end{cases}$$

Continuous Probability Distributions

- At any given sample a continuous random variable has a probability of 0 of assuming exactly any of its values.
- Continuous probability distribution cannot be given in tabular form but it can be stated as a formula.
- We deals with an interval rather than a point value of our random variable.
  - Examples: at least, at most, between, $P(a<X<b)$, $P(W \geq c)$
- Note that when X is continuous,
  - $P(a < X \leq b) = P(a < X < b) + P(X = b) = P(a < X < b)$.
  - As $P(X = b)$ is nearly zero

Probability Density Function (PDF)

- **Definition 3.6:**

The function $f(x)$ is a **probability density function** (pdf) for the continuous random variable X, defined over the set of real numbers, if

1. $f(x) \geq 0$, for all $x \in R$.
2. $\int_{-\infty}^{\infty} f(x)\, dx = 1$.
3. $P(a < X < b) = \int_a^b f(x)\, dx$.

$$P(a < X < b) = \int_a^b f(x)\, dx.$$

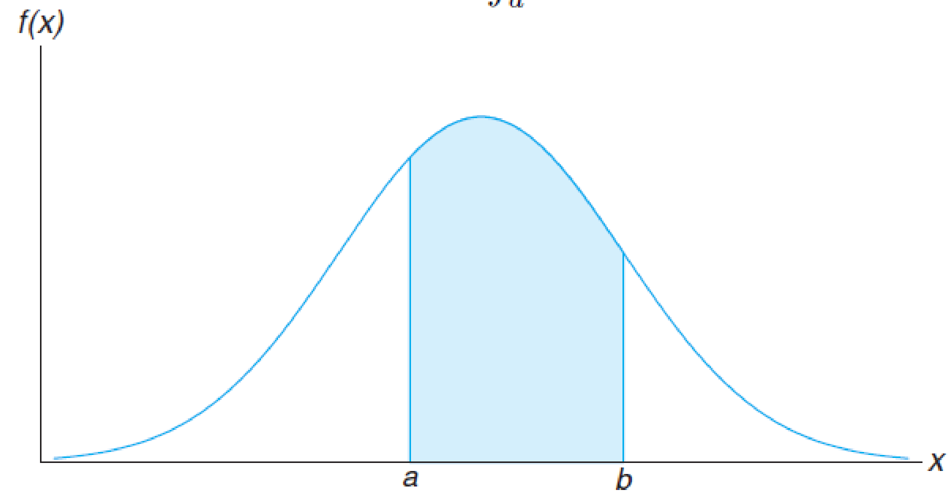

Figure 3.5: $P(a < X < b)$.

Probability Density Function (PDF)

- **Example 3.11:**
  - Suppose that the error in the reaction temperature, in ◦C, for a controlled laboratory experiment is a continuous random variable X having the probability density function.

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Verify that f(x) is a density function.

(b) Find P(0 < X ≤ 1).

Probability Density Function (PDF)

- **Example 3.11:**

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

(a) Verify that f(x) is a density function.

$$\int_{-\infty}^{\infty} f(x)\, dx = \int_{-1}^{2} \frac{x^2}{3} dx = \frac{x^3}{9} |_{-1}^{2} = \frac{8}{9} + \frac{1}{9} = 1.$$

(b) Find $P(0 < X \leq 1)$.

$$P(0 < X \leq 1) = \int_0^1 \frac{x^2}{3} dx = \left. \frac{x^3}{9} \right|_0^1 = \frac{1}{9}.$$

Cumulative Distribution Function

- **Definition 3.7:**

The **cumulative distribution function** $F(x)$ of a continuous random variable X with density function $f(x)$ is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt, \quad \text{for } -\infty < x < \infty.$$

Cumulative Distribution Function

- Example 3.12:
  - For the density function of Example 3.11:

$$f(x) = \begin{cases} \frac{x^2}{3}, & -1 < x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

  - find F(x), and use it to evaluate $P(0 < X \leq 1)$.

Cumulative Distribution Function

For $-1 < x < 2$,

$$F(x) = \int_{-\infty}^{x} f(t)\, dt = \int_{-1}^{x} \frac{t^2}{3} dt = \frac{t^3}{9}\bigg|_{-1}^{x} = \frac{x^3+1}{9}.$$

Therefore,

$$F(x) = \begin{cases} 0, & x < -1, \\ \frac{x^3+1}{9}, & -1 \leq x < 2, \\ 1, & x \geq 2. \end{cases}$$

The cumulative distribution function $F(x)$ is expressed in Figure 3.6. Now

$$P(0 < X \leq 1) = F(1) - F(0) = \frac{2}{9} - \frac{1}{9} = \frac{1}{9},$$

Exercises

- **3.2**
  - An overseas shipment of 5 foreign automobiles contains 2 that have slight paint blemishes. If an agency receives 3 of these automobiles at random, list the elements of the sample space S, using the letters B and N for blemished and non-blemished, respectively; then to each sample point assign a value x of the random variable X representing the number of automobiles with paint blemishes purchased by the agency.

Exercises

- **3.3**
  - Let W be a random variable giving the number of heads minus the number of tails in three tosses of a coin. List the elements of the sample space S for the three tosses of the coin and to each sample point assign a value w of W.

Exercises

- **3.4**
  - A coin is flipped until 3 heads in succession occur. List only those elements of the sample space that require 6 or less tosses. Is this a discrete sample space? Explain.

Exercises

- **3.5:**
  - Determine the value c so that each of the following functions can serve as a probability distribution of the discrete random variable X:

$$\text{(a) } f(x) = c(x^2 + 4), \text{ for } x = 0, 1, 2, 3;$$

$$\text{(b) } f(x) = c\binom{2}{x}\binom{3}{3-x}, \text{ for } x = 0, 1, 2.$$

Exercises

- **3.6:**
  - The shelf life, in days, for bottles of a certain prescribed medicine is a random variable having the density function

$$f(x) = \begin{cases} \frac{20,000}{(x+100)^3}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

  - Find the probability that a bottle of this medicine will have a shell life of

(a) at least 200 days;

(b) anywhere from 80 to 120 days.

Exercises

- **3.6:**
  - The shelf life, in days, for bottles of a certain prescribed medicine is a random variable having the density function

$$f(x) = \begin{cases} \frac{20{,}000}{(x+100)^3}, & x > 0, \\ 0, & \text{elsewhere.} \end{cases}$$

  - Find the probability that a bottle of this medicine will have a shell life of

  (a) at least 200 days;

  (b) anywhere from 80 to 120 days.

Exercises

- **3.7**
  - The total number of hours, measured in units of 100 hours, that a family runs a vacuum cleaner over a period of one year is a continuous random variable X that has the density function.

$$f(x) = \begin{cases} x, & 0 < x < 1, \\ 2 - x, & 1 \le x < 2, \\ 0, & \text{elsewhere.} \end{cases}$$

  - Find the probability that over a period of one year, a family runs their vacuum cleaner

(a) less than 120 hours;

(b) between 50 and 100 hours.

Exercises

- **3.8**
  - Find the probability distribution of the random variable W in Exercise 3.3, assuming that the coin is biased so that a head is twice as likely to occur as a tail.
- 3.12
- 3.13