Random Variables

- A random variable is a variable whose value is determined by the outcome of some chance experiment.
- In general, each outcome of an experiment can be associated with a number by specifying a rule of association.
- Associating a number with each point (element) of a sample space is merely another way of saying that we are “defining a function over the points of a sample space”.

- **Definition** If S is a sample space with a probability measure, and *X* is a real-valued function defined over the elements of S, then *X* is called a random variable.
- That is, a random variable *X* is a function of an elementary event $A: X = f(A)$, where $A \in S$. The value of *X* depends on which elementary event *A* occurs as a result of an experiment.

Example 2.1

Two balls are drawn in succession without replacement from a box containing four red balls and three black balls. We want to list the elements of the sample space and the corresponding values x of the random variable X, the total number of red balls drawn

Example 2.2

Consider an experiment in which a car with Greater Accra registration number (*GR* or *GT*) is selected, and define a random variable *X* by

X=1, if the selected car has a Greater Accra registration number and *X*=0, if the selected car does not have a Greater Accra registration number. That is, if a car has registration number *GT 5246 B*, then

while $X(GT\ 5246\ \mathrm{B}) = 1$. Also

while $X(CR\ 9134\ Q) = 0$ $X(WR\ 2376\ D) = 0$

$X(GT\ 4801\ \mathrm{W}) = 1$

- Random variables are usually classified according to the number of values, which they can assume.

- There are two fundamentally different types of random variables-**discrete** and **continuous** random variables.

- *A* discrete random variable is a random variable that assumes its values only at isolated points.

- A continuous random variable is a random variable that, prior to the experiment can conceivably assume any value in some interval or continuous span of real numbers.

Discrete Random Variables

Definition: A variable quantity *X*, that in an experiment assumes one value out of a finite or infinite sequence $x_1, x_2, \ldots x_k$, is called a discrete random variable, if to each value x_k , there corresponds a definite probability P_k that the variable X_k will assume the value x_k .

- That is, a discrete random variable is a variable that assumes its values only at isolated points.
- If X is a discrete random variable, the function given by $f(x) = P(X = x)$ for each *x* within the range of *X* is called the probability distribution of *X*

Theorem: A function can serve as the probability distribution of a discrete random variable *X* if and only if its values, $f(x)$ satisfy the following conditions:

1. $f(x) \geq 0$ for each value within its domain.

2. $\sum_{x} f(x) = 1$ where the summation extends over all the values within its domain.

Example 2.3 Check whether the following functions can serve as probability distribution of discrete random variables.

1. $f(x) = \frac{x}{15}$, for $x = 0,1,2,3,4,5$
2. $g(x) = \frac{x-2}{2}$, for $x =$ 1,2,3,4
3. $h(x) = \frac{x^2}{25}$, for $x = 0,1,2,3,4$
4. $v(x) = \frac{1}{4}$, for $x = 0,1,2,4$

Presentation of Discrete Probability Distributions

1. The functional relationship between the probabilities P_k and the random variables x_k is called the *distribution law* of the random variable *X*.

Values of X	x_1	x_2	…	…	…	x_k
Probability	P_1	P_2	…	…	…	P_k

Example.2.4

The probability of the occurrence of event A in each of an infinite sequence of trials is equal to *P*. The random variable *X* is the number of the trial in which *A* occurred for the first time. Find the law of distribution of *X.*

2. The distribution law can also be represented graphically in the form of a polygon of probability distribution or the frequency polygon.

   The value x_i of the random variable having the greatest probability is called the **mode.**

- The most general form of the distribution law is the distribution function.

- We shall write the probability that *X* takes on a value less than or equal to *x* as $F(x) = P(X \le x)$ and refer to this function defined for all real numbers *x* as the distribution function or the cumulative distribution of X.

- **Definition** If *X* is a discrete random variable, the function given by

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

is called the distribution function or the cumulative distribution of *X*, and *f(t)* is the value of the probability distribution of *X* at *t*.

The distribution function, for any random variable possesses the following properties:

1. $F(-\infty) = 0$ and $F(\infty) = 1$
2. If $a < b$, then $F(a) \le F(b)$ for any real numbers *a* and *b*.

Example 2.5

Find the cumulative distribution function of the total number of heads obtained in three tosses of a balanced coin.

Numerical Characteristics of Discrete Random Variables

- Associated with any random variable *X*, there exist parameters or constants that characterise the random variable over the entire population under study.

- The most common of these are the mean (μ), the variance , (σ^2) and the standard deviation (σ)

Mathematical Expectation of a Discrete Random Variable

Definition: The mathematical expectation (or simply expectation) of a random variable *X* denoted by $E[x]$ or E_x is the sum of the products of all possible values of the random variable by the probabilities of these variables.

Thus

$$E[X] = x_1 P_1 + x_2 P_2 + \ldots + x_n P_n$$

Or more compactly

$$E[X] = \sum_{k=1}^{n} x_k P_k$$

where

$$\sum_{k=1}^{n} P_k = 1$$

The Variance and the Standard Deviation of a Discrete Random Variable

- **Definition:** The variance of a random variable denoted by $V[x]$ or σ^2 is the expectation of the square of the difference between *X* and its expectation.
- Mathematically, the variance of a random variable is defined as

$$V[X] = E\left[(X - E_x)^2\right]P_k = E\left[X^2\right] - (E[X])^2$$

$$= \sum_{k=1}^{n} (x_k - E_x)^2 P_k$$

Definition: The mean square deviation or the standard deviation of a random variable X denoted by σ , is the square root of its variance.

Mathematically we write

$$\sigma = \sqrt{\sum_{k=1}^{n} (x_k - E_x)^2 P_k}$$

Example 2.6: A random variable is specified by the following distribution table

X	2	3	4
P_k	0.3	0.4	0.3

Determine

1. The expectation
2. The variance
3. The standard deviation

Example.2.7

The probability that Mr Sekyi-Baidoo will sell a piece of property at a profit of 3 million cedis is 0.15. The probability that he will sell it at a profit of 1.5 million cedis is 0.35. The probability that he will break even is 0.35 and the probability that he will lose 1.5 million cedis is 0.15. What is his expected profit?

Continuous Random Variables

- **Definition** A continuous random variable is one that can assume any value in some interval of real numbers, and the probability that it can assume any specific value is zero.
- We designate the probability distribution by the functional notation $f(x)$.

Probability Density Function (p.d.f)

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

$$P(a \leq x \leq b) = \int_a^b f(x)dx$$

For $f(x)$ to be a legitimate p.d.f, it must satisfy two conditions:

1. $f(x) \geq 0.$ for all x

2. $$\int_{-\infty}^{\infty} f(x)dx = 1$$

That is, the area under the entire graph of is equal to 1

Example 2.8

If a random variable has the p.d.f

$$f(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0 & , x \leq 0 \end{cases}$$

find the probability that it will take on a value

(a) between 1 and 3

(b) greater than 0.5

Example 2.9

If *X* has a probability density function of the form

$$f(x) = \begin{cases} ke^{-3x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

Determine

$$P(0.5 \leq x \leq 1)$$

The Distribution Function

As in the discrete case, we shall write as $F(x)$ the probability that a random variable with p.d.f $f(x)$ takes one value less than or equal to *x* and we shall refer to the corresponding function $F(x)$ as the distribution function or the cumulative distribution of the random variable.

Definition: If *x* is a continuous random variable and the value of its p.d.f at *t* is $f(t)$ then the function given by

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

is called the distribution function, or the cumulative distribution of *X*.

Example 2.10

If a random variable has the p.d.f,

$$f(x) = \begin{cases} 2e^{-2x}, x > 0 \\ 0, \quad x \leq 0 \end{cases}$$

find $F(x)$ and hence evaluate $P(x \leq 1)$.

Example 2.11

Consider the function

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

1. Verify that this function defines a p.d.f.

2. Find $P(-1 < x \leq 1)$

3. Find $F(x)$ and use it to evaluate $P(x \leq 1)$

4. What can you deduce from your answers in (2) and (3)?

Numerical Characteristics of a Continuous Random Variable

Definition: Let be a continuous random variable with probability density function $f(x)$. The mathematical expectation $E[X]$ or μ of X is given by

$$E[x] = \int_{-\infty}^{\infty} xf(x)dx.$$

Definition: The variance σ^2 of a continuous random variable *X* with probability density function $f(x)$ and mean μ is given by

$$\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx$$

$$= E[(x-\mu)^2] = E[(x^2)] - (E[x])^2$$

Definition: The standard deviation of a continuous random variable is the square root of its variance.

Example 2.12

The probability density function of weekly gravel sales is given by

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

Find:

1. The expectation
2. The variance
3. The standard deviation

Example2.13

If a random variable has the p.d.f

$$f(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

Find the mean, the variance and the standard deviation of the given random variable.