3 : Basic Statistics

IT5506 – Mathematics for Computing II

Level III - Semester 5

Intended Learning Outcomes

At the end of this lesson, you will be able to;

- define what random variables are and how they are used.
- define the Cumulative Distribution Function of a random variable.
- define what is meant by the distribution of a discrete random variable and a continuous random variable.
- compute the mean and the variance of a discrete and a continuous random variable.

List of sub topics

3.3 Probability distribution of a discrete random variable

3.3.1 Definition

3.3.2 Mean and Variance

3.3 Probability distribution of a discrete random variable

Probability Distribution of Random Variables

As the value of a random variable cannot be predicted in advance, it may be useful to find the probabilities that correspond to the possible values of the random variable.

Definition:

The **probability distribution** of a random variable provides the possible values of the random variable and their corresponding probabilities. A probability distribution can be in the form of a table, graph, or mathematical formula.

Let X be a discrete random variable. The most basic question we can ask is: what is the probability that X takes the value x? In other words, what is $P(X=x)$?

3.3.1 Definition

Definition

Let X be a discrete random variable. Suppose X takes countable number of values $x_1, x_2, x_3, \ldots$. With each possible value x_i, we associate a number $p(x_i) = P(X = x_i)$, then we call $p(x_i)$, i=1,2,3, … the probability of x_i if they satisfy the following conditions.

1. $\sum P(X = xi) = 1$
2. $0 \leq P(X = xi) \leq 1$ for all i

3.3.2 Mean and Variance of a Discrete Random Variable

- Having identified the distribution of a discrete random variable, it may now be important to introduce the centre and spread of a random variable.
- We usually use the mean to describe the centre of a random variable.
- The mean of a random variable is often called the *expected value* of the random variable.
- The variance and standard deviation are used to describe the spread of a random variable.
- Variance is in fact a measure of how spread-out the values are around their mean.

3.3.2 Mean and Variance of a Discrete Random Variable

- The mean and variance of a discrete random variable are given by the following formulas. Note that the mean (or expected value) is denoted by μ (or $E(X)$), and the variance is denoted by σ^2 or $V(X)$.

- $\mu = E(X) = \sum[x.P(X = x)]$
- where x is the value of the random variable and $P(X = x)$ is the probability that X takes the value x. Note that μ is the population mean because the sum ∑ is taken over all values of the r.v.
- $\sigma^2 = V(X) = E(X - E(X))^2$. Here σ is the population standard deviation.

3.3.2 Mean and Variance of a Discrete Random Variable

Note:

- The expected value of X always lies between the smallest and largest values of X.
- The variance of X is never negative.

- To find the standard deviation of the random variable, take the square root of the variance.
- When computing the V(X), it may be easy to use be shown as follows.
- $V(X) = E(X^2) - [E(X)]^2$ which can
- By definition,

$$V(X) = E(X - E(X))^2$$
$$= E(X^2 - 2X.E(X) + [E(X)]^2)$$
$$= E(X^2) - 2E(X).E(X) + [E(X)]^2$$
$$= E(X^2) - [E(X)]^2$$

3.3.2 Mean and Variance of a Discrete Random Variable

Example

Consider the previous example 3, and consider *X* as the number of heads. What are the expected value and variance of *X*?

Answer

We have seen that the probability distribution of X was:

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

E(X) = 0.(1/8) + 1.(3/8) + 2.(3/8) + 3.(1/8) = 3/2.

V(X) = 02.(1/8) + 12 (3/8) + 22.(3/8) + 32.(1/8) - [3/2]2 = 3/4,

since V(X) = E(X2) – [E(X)]2

3.3.2 Mean and Variance of a Discrete Random Variable

Now, consider the following theorems with regard to the expected value and the variance. The proofs of the theorems are not given here.

- Theorem 1: If ‘a’ is a constant, then

  $E(a) = a$
- Theorem 2: If ‘a’ is a constant and X is any random variable, then

  $E(aX) = a\,E(X)$
- Theorem 3: If ‘a’ and ‘b’ are constants, then

  $E(aX + b) = a\,E(X) + b$
- Theorem 4: If ‘a’ is a constant, then

  $V(a) = 0$
- Theorem 5: If ‘a’ is a constant and X is any random variable, then

  $V(aX) = a^2\,V(X)$
- Theorem 6: If ‘a’ and ‘b’ are constants, then

  $V(aX + b) = a^2\,V(X)$

Note that these theorems are valid whether X is a discrete or a continuous random variable.