BIT

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 continuous random variables are and how they are used.
- define the Cumulative Distribution Function of a continuous random variables .
- compute the mean and the variance of a continuous random variable.
- use and interpret some continuous probability distributions.

List of sub topics

3.6 Probability distribution of a continuous random variable

3.6 Probability distribution of a continuous random variable

- A continuous random variable can take any values anywhere in some interval of the real line, e.g. $[0, \infty)$ or $(0, 1)$. Very often quantities such as time, weight, height etc. are commonly considered as continuous random variables.
- Recall that, for a discrete random variable X, the probability distribution lists all values that X can take, and give their probabilities. For a continuous random variable X, it is impossible to list all the values that X can take. It is also impossible to think of the probability that X takes any one specific value.
- E.g.: Even between the values 0.99999999 and 1.00000001, there are so many values that the probability of each value is negligible. In fact, we write $P(X = x) = 0$ for any x, when X is continuous. Instead, we work with intervals for continuous random variables:
- E.g. $P(X = 0) = 0$, but $P(0.999 \leq X \leq 1.001)$ can be > 0.

3.6.1 Definition of a Probability distribution of a continuous random variable

Definition

X is said to be a *continuous random variable* if there exists a function *f*, called the probability density function (pdf) of X, satisfying the following conditions:

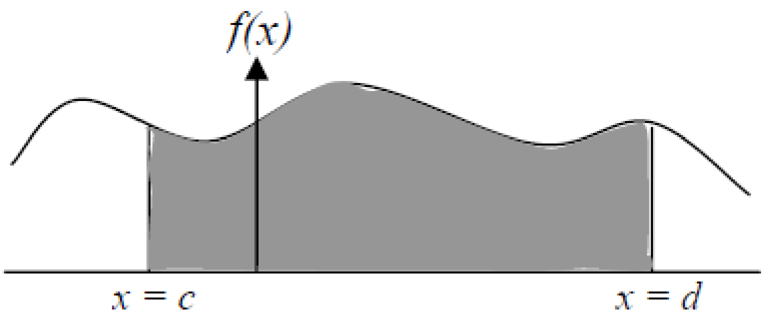
(a) $\int_{-\infty}^{+\infty} f(x)dx = 1$

(b) $f(x) \geq 0$ for all x

3.6 Probability distribution of a continuous random variable

Note:

P($c \leq x \leq d$) represents the shaded area under the graph in the following figure of the probability density function between $x = c$ and $x = d$.



$$\text{So, } P(c \leq X \leq d) = \int_c^d f(x)dx = F_X(d) - F_X(c).$$

3.6 Probability distribution of a continuous random variable

Properties of the Cumulative Distribution Function (when X is Continuous)

1. $F(-\infty) = 0$, $F(+\infty) = 1$
2. $F(x)$ is a non-decreasing continuous function of x. This means that $F(x) < F(y)$ if $x < y$.
3. $P(a \leq X \leq b) = F(b) - F(a)$

- It makes no difference whether we say $P(a < X \leq b)$ or $P(a \leq X \leq b)$ because $P(a \leq X \leq b) = P(X = a) + P(a < X \leq b) = P(a < X \leq b)$ since $P(X = a) = 0$.
- i.e. for a continuous random variable, $P(a < X < b) = P(a \leq X \leq b)$.
- However, this is not true for a discrete random variable.

3.6 Probability distribution of a continuous random variable

Properties of the Cumulative Distribution Function (when X is Continuous)

1. $F(-\infty) = 0$, $F(+\infty) = 1$
2. $F(x)$ is a non-decreasing continuous function of x. This means that $F(x) < F(y)$ if $x < y$.
3. $P(a \leq X \leq b) = F(b) - F(a)$

3.6 Probability distribution of a continuous random variable

Definition:

Let X be a continuous random variable with cumulative distribution function FX(*x*). The probability density function (pdf) of X is:

$$f(x) = \frac{dF(x)}{dx}$$

Note : If *f(x)* is the pdf for a continuous random variable, then

$$F(x) = \int_{-\infty}^{x} f(x)\,dx$$

This is true only if X is a continuous r.v.

3.6.2 Mean and Variance of a Continuous Random Variable

Definition:

Let X be a continuous random variable with cumulative distribution function FX(*x*). The probability density function (pdf) of X is:

$$f(x) = \frac{dF(x)}{dx}$$

Note : If *f(x)* is the pdf for a continuous random variable, then

$$F(x) = \int_{-\infty}^{x} f(x)\,dx$$

This is true only if X is a continuous r.v.

3.6.2 Mean and Variance of a Continuous Random Variable

The mean and variance of a continuous random variable are given by the following formulas.

$$\mu = E(X) = \int_{-\infty}^{+\infty} x.f(x)dx$$

where $f(x)$ is the probability density function.

$$\sigma^2 = V(X) = E(X - E(X))^2 \quad \text{or}$$
$$= E(X^2) - [E(X)]^2 \quad \text{or}$$
$$= \int_{-\infty}^{+\infty} x^2.f(x)dx - \mu^2$$

Note : The properties of variance for continuous r.v.'s are exactly the same as for discrete r.v.'s.

3.6.2 Mean and Variance of a Continuous Random Variable

Example:

Suppose that the random variable *X* has p.d.f. given by

$$f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & Otherwise \end{cases}$$

(a) Verify whether *f(x)* is a p.d.f.

(b) Find the cumulative distribution function of X

(c) Find the mean and variance of X

3.6.2 Mean and Variance of a Continuous Random Variable

Solution:

(a) $f(x) \geq 0$ for all x because $f(x)=2x \geq 0$ when $0 \leq x \leq 1$, and $f(x) = 0$ when x takes all the other values.

Also, $\int_0^1 f(x)dx = \int_0^1 2x.dx = [x^2]_0^1 = 1$

Therefore, $f(x)$ satisfies the two conditions to become a p.d.f.

(b) The Cumulative distribution function

$$F_X(x) = \int_0^1 f(x)dx = \begin{cases} 0 & \text{when } x < 0 \\ x^2 & \text{when } 0 \leq x \leq 1 \\ 1 & \text{when } x > 1 \end{cases}$$

(c) Mean $= E(X) = \int_0^1 x.2x.dx = 2\left[\frac{x^3}{3}\right]_0^1 = \frac{2}{3}$

$$E(X^2) = \int_0^1 x^2.2x.dx = 2\left[\frac{x^4}{4}\right]_0^1 = \frac{1}{2}$$

So, $V(X) = E(X^2) - [E(X)]^2 = 1/2 - 4/9 = 1/18.$