Introduction to Statistics (MAT 283)

Dipti Dubey

Department of Mathematics
Shiv Nadar University

Table of Contents

CONTINUOUS RANDOM VARIABLE

MOMENTS OF RANDOM VARIABLES

CONTINUOUS RANDOM VARIABLE:

A random variable X is said to be continuous if and only if there exists a function $f_X : \mathbb{R} \to \mathbb{R}$ such that $f_X(x) \geq 0$, $\int_{-\infty}^{\infty} f_X(x)dx = 1$ and

$$P(a < X < b) = \int_a^b f_X(x)dx$$

for any real constants a and b with $a \leq b$. The function f_X is called the **probability density function** (pdf)of X.

- Note that $f_X(c)$, the value of the probability density of X at c, does not give $P(X = c)$ as in the discrete case. In connection with continuous random variables, probabilities are always associated with intervals and $P(X = c) = 0$ for any real constant c.

• Note that $f_X(c)$, the value of the probability density of X at c, does not give $P(X = c)$ as in the discrete case. In connection with continuous random variables, probabilities are always associated with intervals and $P(X = c) = 0$ for any real constant c.

Theorem: If X is a continuous random variable and a and b are real constants with $a \leq b$, then

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

- For a pdf $f_X(x) > 1$. For example $f_X(x) = 5$ for $x \in [0, 1/5]$ and 0 otherwise.

- For a pdf $f_X(x) > 1$. For example $f_X(x) = 5$ for $x \in [0, 1/5]$ and 0 otherwise.

- A pdf can be unbounded also.

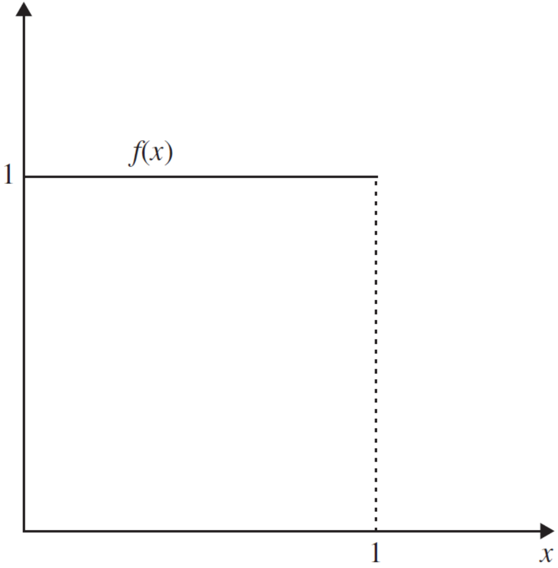
CUMULATIVE DISTRIBUTION FUNCTION: If X is a continuous random variable and the value of its probability density at$t$ is $f(t)$, then the function given by

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

is called the cumulative distribution function of X.

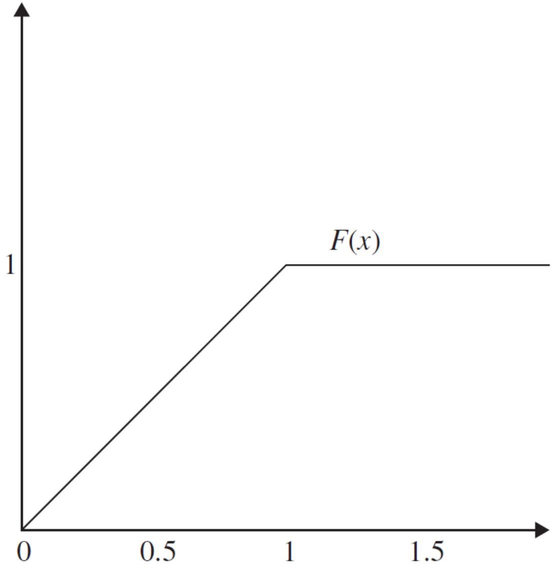
EXAMPLE: Consider a continuous random variable X with pdf

$$f_x(x) = \begin{cases} 1 & \text{if } x \in (0,1) \\ 0 & \text{if } \text{ otherwise.} \end{cases}$$


f(x)
1
1
x

The CDF is given by

$$F_x(x) = \begin{cases} 0 & \text{if } x \leq 0 \\ x & \text{if } x \in (0, 1] \\ 1 & \text{if } x > 1. \end{cases}$$



$F(x)$
1
0
0.5
1
1.5

• If F is the cumulative distribution function of a continuous random variable X, the probability density function f of X is the derivative of F, that is

$$\frac{d}{dx}F(x) = f(x).$$

Table of Contents

CONTINUOUS RANDOM VARIABLE

MOMENTS OF RANDOM VARIABLES

Moments of Random Variables: Let X be a random variable with space R_X and probability density function f. The n^{th} moment about the origin of a random variable X, as denoted by $E(X^n)$, is defined to be

$$E(X^n) = \begin{cases} \sum_{x \in R_X} x^n f(x) & \text{if X is discrete} \\ \\ \int_{-\infty}^{\infty} x^n f(x)\, dx & \text{if X is continuous} \end{cases}$$

for $n = 0, 1, 2, 3, \ldots,$ provided the right side converges absolutely.

- If $n = 1$, then $E(X)$ is called the first moment about the origin. If $n = 2$, then $E(X^2)$ is called the second moment of X about the origin.

- If $n = 1$, then $E(X)$ is called the first moment about the origin. If $n = 2$, then $E(X^2)$ is called the second moment of X about the origin.
- In general, these moments may or may not exist for a given random variable. If for a random variable, a particular moment does not exist, then we say that the random variable does not have that moment.

Expected Value of Random Variables:

Let X be a random variable with space R_X and probability density function f. The expected value $E(X)$ (or mean μ_X) of the random variable X is defined as

$$\mu_X = E(X) = \begin{cases} \sum_{x \in R_X} x\, f(x) & \text{if } X \text{ is discrete} \\ \\ \int_{-\infty}^{\infty} x\, f(x) dx & \text{if } X \text{ is continuous} \end{cases}$$

if the right hand side exists.

EXAMPLE: Tossing a fair coin

$$S = \{H, T\}$$

Define

$$X : S \to \mathbb{R}$$

such that

$$X(H) = 0$$

$$X(T) = 1$$

Then $R_X = \{0, 1\}$, $f(1) = P(x = 1) = 1/2$, $f(0) = P(x = 0) = 1/2$ and

EXAMPLE: Tossing a fair coin

$$S = \{H, T\}$$

Define

$$X : S \to \mathbb{R}$$

such that

$$X(H) = 0$$

$$X(T) = 1$$

Then $R_X = \{0, 1\}$, $f(1) = P(x = 1) = 1/2$, $f(0) = P(x = 0) = 1/2$ and

$$E(x) = 0 \times f(0) + 1 \times f(1) = 0 \times 1/2 + 1 \times 1/2 = 1/2$$

EXAMPLE: Consider a continuous random variable X with pdf

$$f(x) = \begin{cases} 1/5 & \text{if } x \in (2,7) \\ 0 & \text{if } \text{ otherwise.} \end{cases}$$

Then

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x\, f(x) dx \\ &= \int_2^7 x \frac{1}{5} dx \end{aligned}$$

Theorem. Let X be a random variable and $Y = g(X)$.
(a) If X is a discrete random variable with pmf f, then

$$E(g(x)) = \sum_x g(x)\, f(x);$$

(b) Correspondingly, if X is a continuous random variable with pdf f, then

$$E(g(x)) = \int_{-\infty}^{\infty} g(x)\, f(x) dx.$$

Proof. We will prove this result for the discrete case. By definition,

$$E(g(X)) = \sum_i g(x_i) f_{g(x)}(g(x_i)).$$

As g may not be a one-one map, suppose $g(X) = y_i$ when X takes on values $x_{i_1}, x_{i_2}, \ldots, x_{i_{n_i}}$. Then

$$P(g(X)) = y_i = \sum_{j=1}^{n_i} f(x_{i_j})$$

and $g(X)$ can take on values $y_1,\ y_2,\ \ldots, y_m$. Therefore,

$$\begin{aligned} E(g(X)) &= \sum_{i=1}^{m} y_i P[g(X) = y_i] \\ &= \sum_{i=1}^{m} y_i \left(\sum_{j=1}^{n_i} f(x_{i_j}) \right) \end{aligned}$$

$$= \sum_{i=1}^{m} \sum_{j=1}^{n_i} y_i f(x_{i_j})$$

$$= \sum_{x} g(x) f(x),$$

where summation extends over all values of X.

Theorem. Let X be a random variable with pdf f. If a and b are any two real numbers, then

$$E(aX + b) = a\,E(X) + b.$$

Proof: We will prove only for the continuous case.

$$E(aX + b) = \int_{-\infty}^{\infty} (ax + b) f(x) dx$$

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

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

$$= a\,E(X) + b.$$

EXAMPLE: If X is the number of points rolled with a balanced die, find the expected value of $g(X) = 2X^2 + 1$.