Introduction to Statistics (MAT 283)

Dipti Dubey

Department of Mathematics
Shiv Nadar University

August 22, 2022

CUMULATIVE DISTRIBUTION FUNCTION: If X is a discrete random variable, the function $F : \mathbb{R} \to \mathbb{R}$ defined by

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

where $f(t)$ is the value of the probability mass function of X at t, is called the cumulative distribution funtion of X.

• The values $F(x)$ of the distribution function of a discrete random variable X satisfy the conditions:

1. $F(-\infty) = 0$ and $F(\infty) = 1$;
2. if $a < b$, then $F(a) \leq F(b)$ for any real numbers a and b.

- The values $F(x)$ of the distribution function of a discrete random variable X satisfy the conditions:

1. $F(-\infty) = 0$ and $F(\infty) = 1$;

2. if $a < b$, then $F(a) \leq F(b)$ for any real numbers a and b.

- If the range of a random variable X consists of the values $x_1 < x_2 < x_3 < \quad < x_n$, then $f(x_1) = F(x_1)$ and

$$f(x_i) = F(x_i) - F(x_{i-1}) \text{ for } i = 2, 3, \ldots, n$$

EXAMPLE:

Flip a fair coin twice and let X be the number of heads. Then

$$\Omega = \{HH, HT, TH, TT\}$$

and

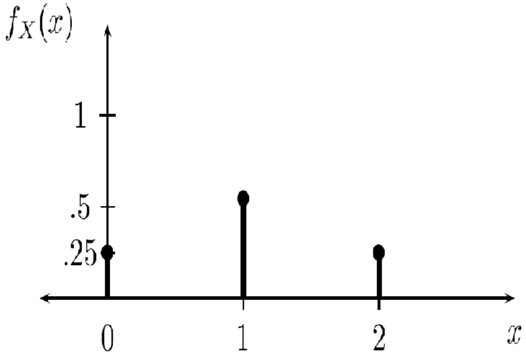
$$R_X = \{0, 1, 2\}.$$

Then,

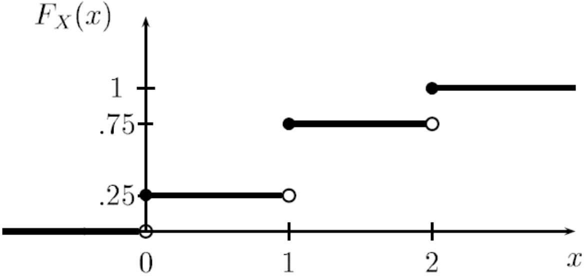
$f(0) = P(\{TT\}) = 1/4$, $f(1) = P(\{HT, TH\}) = 1/2$,
$f(2) = P(\{HH\}) = 1/4$.

The cumulative distribution function (CDF) is

$$F(x) = \begin{cases} 0 & x < 0 \\ 1/4 & 0 \leq x < 1 \\ 3/4 & 1 \leq x < 2 \\ 1 & 2 \leq x \end{cases}$$



Probability function for flipping a coin twice



CDF for flipping a coin twice.

EXAMPLE:

If the probability of a random variable X with space $R_X = \{1, 2, 3, ..., 12\}$ is given by

$$f(x) = k(2x - 1),$$

then, what is the value of the constant k? Also find the cumulative distribution function F of X.

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.

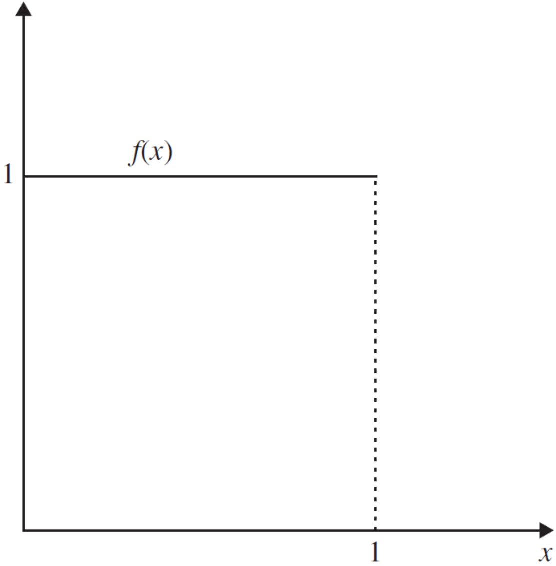
CUMULATIVE DISTRIBUTION FUNCTION: If X is a continuous random variable and the value of its probability density at$t$ is $f(t)$, then the function given by

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

is called the cumulative distribution function of X.

EXAMPLE: Consider a continuous random variable X with pdf

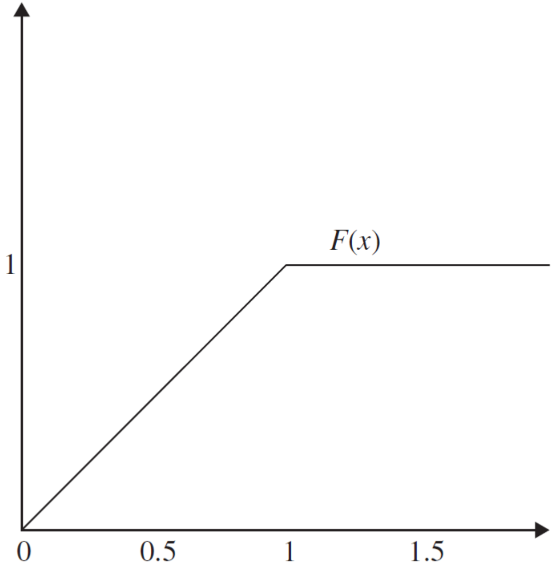
$$f_X(x) = \begin{cases} 1 & \text{if } x \in (0,1) \\ 0 & \text{if otherwise.} \end{cases}$$



$f(x)$
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