6. Let $\mathbb{F} : \mathbb{R} \to [0,1]$ be a distribution function and continuous. Then show that the following functions are distribution functions.

$$\mathbb{F}_1(x) = \frac{1}{h} \int_x^{x+h} \mathbb{F}(u)\, du \quad \forall\, x \in \mathbb{R}$$

$$\mathbb{F}_2(x) = \frac{1}{2h} \int_{x-h}^{x+h} \mathbb{F}(u)\, du \quad \forall\, x \in \mathbb{R}.$$

7. Suppose the duration in minutes of a long-distance telephone calls made from a city is found to be a random variable with a probability distribution specified by the distribution function $\mathbb{F}$, given by

$$\mathbb{F}(x) = \begin{cases} 0 & \text{for } x \leq 0. \\ 1 - \left(2/3\right) e^{-x/3} - (1/3)\, e^{-\lfloor x/3 \rfloor} & \text{for } x > 0. \end{cases}$$

where $\lfloor x \rfloor$ is the largest integer less than or equal to x. Sketch the distribution function.

8. Let X be a random ~~rank~~ variable on $(\Omega, 2^{\Omega})$ where $\Omega = \{0, 1, 2, \ldots n\}$. ~~and~~ with probability distribution specified by

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad \text{for all } k = 0, 1, \ldots n$$

where $p \in$ (0,1).

$$\mathbb{F}(x) = \sum_{i=0}^{x} \binom{n}{i} p^i (1-p)^{n-i}$$

$$= \frac{1}{B(n-x, x+1)} \int_0^{1-p} z^{n-x-1} (1-z)^x \, dz.$$

The random variable X is called Binomial random variable. The distribution function of X is related to incomplete Beta function.

9. Let X be a random variable on $(\mathbb{N}_0, 2^{\mathbb{N}_0})$ with probability distribution specified by

$$\mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!} \quad \text{for every } k \in \mathbb{N}_0.$$

The random variable X is called Poisson random variable. Note that

$$\mathbb{F}(x) = \mathbb{P}(X \leq x) = \sum_{i=0}^{x} e^{-\lambda} \frac{\lambda^i}{i!}$$

$$= 1 - \frac{1}{\Gamma(x+1)} \int_0^{\lambda} z^x e^{-z} \, dz \quad \text{for every } x \in \mathbb{N}_0.$$

The distribution function of Poisson random variable is related to incomplete Gamma distribution.

Lecture-V : Discrete and continuous random variables

Definition

A random variable $X : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$ is said to be of discrete type or simply discrete, if there exists a countable set $E \subseteq \mathbb{R}$ such that $\mathbb{P}(X \in E) = \mathbb{P}_X(E) = 1$. (Recall that every countable subset of $\mathbb{R}$ is a Borel set.) The points of E that have positive mass are called <u>jump points</u> or <u>points of increase</u> of the distribution function of X, and their probabilities are called jumps of the distribution function.

Let X be a discrete random variable on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with jump points in E. Let $\{x_i : i \geq 1\}$ be an enumeration of E. Let us define

$$p_i = \mathbb{P}(\{\omega : X = x_i\}) = \mathbb{P}_X(\{x_i\}) \text{ for every } i \in \mathbb{N}.$$

It is immediate that

$$p_i \geq 0 \text{ for all } i \in \mathbb{N} \quad \text{and} \quad \sum_{i \geq 1}^{\infty} p_i = 1.$$

Definition

The collection of numbers $(p_i : i \geq 1)$ satisfying $\mathbb{P}_X(\{i\}) = p_i \geq 0$ for all $i \geq 1$ and $\sum_{i \geq 1}^{\infty} p_i = 1$, is called probability mass function of the discrete random variable X.

The distribution function of the random variable X is given by

$$F(x) = \mathbb{P}(X \leq x) = \sum_{i:\, x_i \leq x} p_i .$$

Let $\mathbb{1}_A$ denotes the indicator function of the set A that is,

$$\mathbb{1}_A(x) = \begin{cases} 1 & \text{if } x \in A \\ 0 & \text{if } x \notin A \end{cases} .$$

We can take $\Omega = \mathbb{R}$ and write down the discrete random variable X as

$$X(\omega) = \sum_{i \geq 1}^{\infty} x_i \mathbb{1}_{\{x_i\}}(\omega) .$$

<u>Theorem</u>

Let $(p_k : k \geq 1)$ be a collection of non-negative real numbers such that $\sum_{k \geq 1}^{\infty} p_k = 1$. Then there exists a random variable X which has probability mass function $(p_k : k \geq 1)$.

We now consider random variables which do not have jump points.

Definition.

Let X be a random variable defined on $(\Omega, \mathcal{F}, \mathbb{P})$ with distribution function F. Then X is said to be of the continuous type (or simply, continuous) if F is absolutely continuous that is, if there exists a non-negative function $f: \mathbb{R} \to [0, \infty)$ such that for any real number x, we have

$$\mathbb{F}(x) = \int_{-\infty}^{x} dt \ f(t).$$

The function f is called probability density function.

It follows from the properties of distribution function that

$$\int_{\mathbb{R}} dt \ f(t) = \int_{-\infty}^{\infty} dt \ f(t) = 1.$$

Theorem

Let $f: \mathbb{R} \to [0, \infty)$ with $\int_{-\infty}^{\infty} dx \ f(x) = 1$. Then there exists a unique probability $\mathbb{P}$ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that

$$\mathbb{P}(B) = \int_{B} dx \ f(x) \quad \text{for every } B \in \mathcal{B}(\mathbb{R}). \longrightarrow (*_1)$$

That is, there exists a random variable X on a probability space $(\Omega, \mathcal{F}, \tilde{\mathbb{P}})$ with probability density function f.

In Riemann integral, we have only taken f to be continuous and B to be an open interval. So $(*_1)$ can not be supported by the theory of Riemann integration as we have never mentioned density function to be continuous or restrict B to be an interval. To support our claim we need tools and results from measure and integration theory which of course not the aim of this course and so we skip it.

$*$ In the discrete case $p_i = \mathbb{P}(X = \{i\})$ is a probability. In continuous case, $f(x)$ does not stand for $\mathbb{P}(\{X = x\})$ and is not a probability.

Theorem

Let X be any random variable. Then

$$\mathbb{P}(\{X = a\}) = \lim_{t \uparrow a} \mathbb{P}(X \in (t, a]).$$

Proof Left as an exercise. Hint: use continuity property of probability.

Exercise Show that $\mathbb{P}(X = a) = F(a) - F(a-)$ for every $a \in \mathbb{R}$ where F is the distribution function of X.

Remark If X is a continuous random variable with absolutely continuous distribution function F. Then for every $a \in \mathbb{R}$, $\mathbb{P}(X = a) = 0$. More generally, if X is a random variable with continuous (might not be absolutely continuous) distribution function, we have $\mathbb{P}(X=a)=0$. for every $a \in \mathbb{R}$.

Question What is the difference between the continuous and absolutely continuous distribution function?

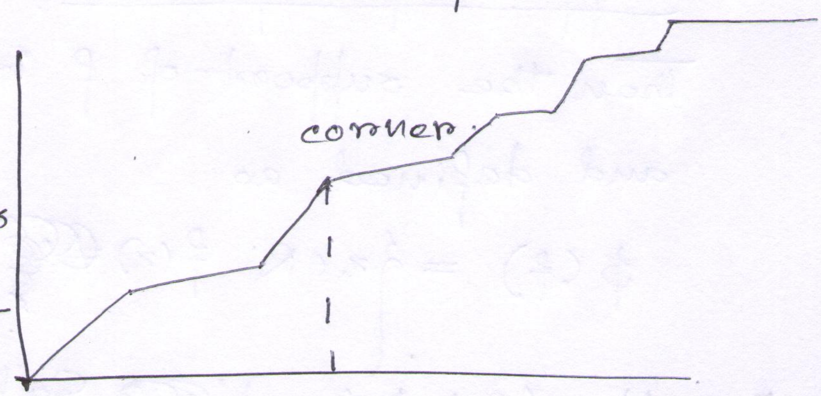
According to the definition of distribution function, any function F which satisfies.

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

ii) $x \mapsto F(x)$ is continuous.

iii) $x \mapsto F(x)$ is non-decreasing

is a distribution function. Note that these conditions ~~are not~~ do not gurantee existance of derivative of F. If we sketch the graph of F, then the graph may have corners and hence, does not have a derivative at that point.



Then we can write down F as a sum of continuous functions which are supported on the ~~disjoint~~ intervals. We then call that X has a piecewise continuous density function. We can use Reimann integral to integrate each of them. to get the distribution function. We shall see such examples.

Remark :- Suppose that $\mathbb{F}$ be an absolutely continuous distribution function that is, there exists a ~~non~~ continuous function $f : \mathbb{R} \to [0, \infty)$ such that

$$\mathbb{F}(x) = \int_{-\infty}^{x} du\, f(u). \text{ for every } x \in \mathbb{R}.$$

Then the density function can be obtained as

$$f(x) = \frac{d}{du} \mathbb{F}(u) \Big|_{u=x}.$$

using fundamental theorem of calculus.

* <u>Support of a function.</u> Let $f : \mathbb{R} \to \mathbb{R}$ be a function. Then the support of f is denoted by $\$(f)$. and defined as

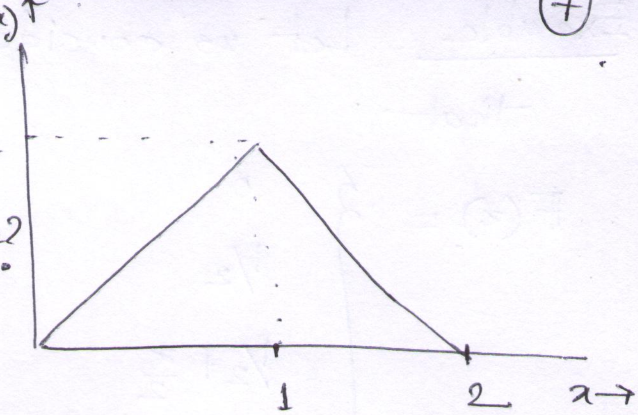
$$\$(f) = \{ x \in \mathbb{R} : f(x) \neq 0 \}.$$

<u>Example</u> (Rohatgi and Saha)

Let the random variable X have a triangular p.d.f. (probability density function)

$$f(x) = \begin{cases} x & 0 < x \leq 1 \\ 2-x & 1 \leq x \leq 2 \\ 0 & \text{otherwise} \end{cases}$$

Q^n How can we check that a non-negative function is a probability density function?

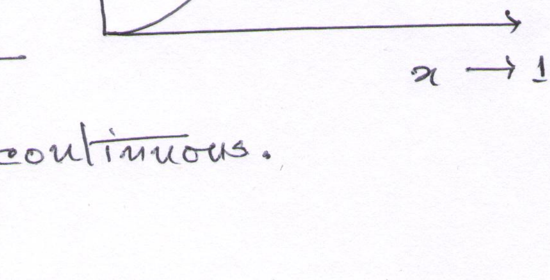


Although, we do not derived any properties of the p.d.f., we can integrate the function given and should check whether the integral satisfies the properties of distribution function.

$$F(x) = \int_{-\infty}^{x} du\, f(u) = \begin{cases} 0 & \text{if } x \leq 0 \\ \int_0^x du\, u = \frac{x^2}{2} & \text{if } 0 < x \leq 1 \\ \int_0^1 du(u) + \int_1^x du\,(2-u) & \text{if } 1 < x \leq 2 \\ \int_0^1 du\,(u) + \int_1^2 du\,(2-u) & \text{if } u > 2 \end{cases}$$

$$= \begin{cases} 0 & \text{if } x \leq 0 \\ \frac{x^2}{2} & \text{if } 0 < x \leq 1 \\ \frac{1}{2} + 2(x-1) - \frac{1}{2}(x^2-1) = 2x - \frac{1}{2}x^2 - 1 & \text{if } 1 < x \leq 2 \\ 1 & \text{if } x \geq 2 \end{cases}$$

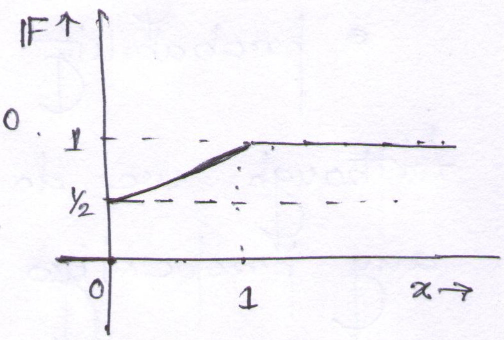
It is easy to check that F is a distribution function.



* This is the example where f the density function is continuous.

<u>Example</u> Let us consider a function $F : \mathbb{R} \to [0,1]$ such that

$$F(x) = \begin{cases} 0 & \text{if } x < 0. \\ 1/2 & \text{if } x = 0. \\ 1/2 + \frac{x}{2} & \text{if } 0 < x < 1. \\ 1 & \text{if } 1 \le x. \end{cases}$$



$F(0) - F(0-) = \frac{1}{2}$, 0 is the point of discontinuity of F.

It is easy to check that F is a distribution function. Therefore, there exists a random variable X which has the distribution function F and probability distribution $\mathbb{P}_X$. It is clear that X is not a continuous random variable as

$$\mathbb{P}_X(\{0\}) = F(0) - F(0-) = \frac{1}{2}. \longrightarrow (*_1)$$

If $x \in (0,1)$, then $F(x)$ (is absolutely ~~contin~~ continuous)

$$F(x) = \frac{1}{2} + \frac{1}{2}\int_0^x du. \longrightarrow (*_2)$$

From $(*_1)$ and $(*_2)$, we conclude that X is neither discrete nor continuous.