Random Variables

Let $(\Omega, \mathcal{S}, P)$ be a given prob. space. In some situation we may not be directly interested in the sample space Ω; rather we may interested in some numerical aspect of Ω.

Example: A fair coin (head and tail are equally likely) is tossed three times independently, then

$$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

and $$P(\{\omega\}) = \frac{1}{8} \quad \forall \; \omega \in \Omega.$$

Suppose that we are interested in number of heads in three tosses, i.e. we are interested in the function

$X : \Omega \to \mathbb{R}$ defined as

$$X(\omega) = \begin{cases} 0 & \text{if } \omega = TTT \\ 1 & \text{if } \omega \in \{HTT, THT, TTH\} \\ 2 & \text{if } \omega \in \{HHT, HTH, THH\} \\ 3 & \text{if } \omega \in \{HHH\}. \end{cases}$$

Clearly the values assumed by X are random with

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

Hence $Pr(X \in \{0, 1, 2, 3\}) = 1$.

Defn: Let $(\Omega, \mathcal{S}, P)$ be a given probability function. A real valued function $X : \Omega \to \mathbb{R}$ (defined on sample space Ω) is called a random variable (r.v.).

Note: In Rigorus mathematical point of view random variable is ~~not~~ only real valued fun with some technical condition. In this course we are ignoring these technical details. For all practical purpose r.v. is a real valued function defined on Ω.

For a probability space $(\Omega, \mathcal{S}, P)$ and a r.v $X : \Omega \to \mathbb{R}$ note that $\forall \ A \subseteq \mathbb{R}$

$$X^{-1}(A) = \{\omega \in \Omega : X(\omega) \in A\} \in \mathcal{S}$$

Thus one can define a set fun $P_X : \mathcal{B} \to [0,1]$ by

$$P_X(B) = P(X^{-1}(B)) = P(\{\omega \in \Omega : X(\omega) \in B\}), \quad B \in \mathcal{B}$$

where $\mathcal{B}$ is a some class of subsets of $\mathbb{R}$. Here also for all practical purposes we will take $\mathcal{B}$ to be power set of $\mathbb{R}$.

We simply write

$$P_X(B) = P(\{\omega \in \Omega : X(\omega) \in B\}) = P_r(X \in B)$$
$$B \in \mathcal{B}.$$

We have the following scenaio

$$(\Omega, \mathcal{E}, P) \xrightarrow{X} (\mathbb{R}, \mathcal{B}, P_X)$$

__Theorem__ (__Induced Probability space__): $(\mathbb{R}, \mathcal{B}, P_X)$ is a prob space i.e $P_X(\cdot)$ is a prob. function defined on $\mathcal{B}$.

__Proof__: (i) $P_X(\mathbb{R}) = P_r(X \in \mathbb{R}) = P(X^{-1}(\mathbb{R})) = P(\Omega) = 1$

(ii) For any $B \in \mathcal{B}$,

$$P_X(B) = P(X^{-1}(B)) \geq 0$$

(iii) Let $\{B_n\}_{n \geq 1}$ be a collection of mutually exclusive events in $\mathcal{B}$. Then

$$P_X\left(\bigcup_{n=1}^{\infty} B_n\right) = P\left(X^{-1}\left(\bigcup_{n=1}^{\infty} B_n\right)\right) = P\left(\bigcup_{n=1}^{\infty} X^{-1}(B_n)\right)$$

$$= \sum_{n=1}^{\infty} P(X^{-1}(B_n)) = \sum_{n=1}^{\infty} P_X(B_n)$$

__Def^n__: The prob function P_X defined above is called the probability function/measure induced by r.v. X and $(\mathbb{R}, \mathcal{B}, P_X)$ is called the probability space induced by r.v. X.

Example: Consider the example: tossing a coin three times independently,

$$\Omega = \{HHT, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

$$P(\{\omega\}) = \frac{1}{8}, \quad \forall \omega \in \Omega$$

and

$X : \Omega \to \mathbb{R}$ (number of heads in three tosses)

$X \to 0, 1, 2, 3.$

$$\cancel{P}\{\omega : X(\omega) = 0\} = \{T, T\},$$

$$\{\omega : X(\omega) = 1\} = \{HTT, THT, TTH\}$$

$$\{\omega : X\omega) = 2\} = \{HHT, HTH, THH\}$$

$$\{\omega : X(\omega) = 3\} = \{HHH\}.$$

$X : \Omega \to \mathbb{R}$ is r.v. The induced probability space is $(\mathbb{R}, \mathcal{B}, P_X)$, where

$$P_X(\{0\}) = P(\{TTT\}) = \frac{1}{8}$$

$$P_X(\{1\}) = P(\{HTT, THT, TTH\}) = 3/8$$

$$P_X(\{2\}) = P(\{HHT, HTH, THH\}) = 3/8$$

$$P_X(\{3\}) = P(\{HHH\}) = 1/8$$

Now for any $B \in \mathcal{B}$

$$P_X(B) = P(X^{-1}(B)) = P(\{\omega \in \Omega : X(\omega) \in B\})$$

$$= \sum_{\substack{i=0 \\ i \in B}}^{3} P_X(\{i\}).$$

Distribution Function :

Let X be a r.v. defined on probability space $(\Omega, \mathcal{S}, P)$ and let $(\mathbb{R}, \mathcal{B}, P_X)$ denote the probability space induced by X. Define the function $F_X : \mathbb{R} \to \mathbb{R}$ by

$$F_X(x) = P_r(X \le x) = P_X(X \le x)$$
$$= P_X((-\infty, x]) \quad , \; x \in \mathbb{R}.$$

The function F_X is called the cumulative distribution function (c.d.f) or simply the distribution function (d.f) of r.v. X.

Note: Whenever there is no ambiguity we will drop sub-script X in F_X to represent d.f. of a r.v by F.

Example: In the previous example

$$P_r(X=0) = P_X(\{0\}) = 1/8$$

$$P_r(X=1) = P_X(\{1\}) = 3/8 = P_r(X=2) = P(\{2\})$$

and $$P_r(X=3) = P_X(\{3\}) = 1/8.$$

Then the d.f. of X is obtained as

$$F_X(x) = P_r(X \leq x) = P\{w : X(w) \leq x\} = \sum_{i \leq x} P_X(\{i\})$$

$$= \begin{cases} 0, & x < 0 \\ \frac{1}{8}, & 0 \leq x < 1 \\ \frac{1}{8} + \frac{3}{8} = \frac{1}{2} & 1 \leq x < 2 \\ 7/8 & 2 \leq x < 3 \\ 1 & x \geq 3 \end{cases}$$

~~The following result from calculus will be usefull in studying the properties of d.f.~~

~~Result: Let $-\infty \leq a < b < \infty$ and let $f : (a, b) \to \mathbb{R}$~~

Properties of cdf : (i) $\lim_{x \to -\infty} F(x) = 0$, $\lim_{x \to \infty} F(x) = 1$

(ii) If $x_1 < x_2$ then $F(x_1) \leq F(x_2)$

(iii) $F(x)$ is right continuous at every point i.e.

$$\lim_{h \to 0} F(x+h) = F(x) \quad \text{i.e} \quad F(x+) = F(x).$$

Proof: Let $\{x_n\}$ be a decreasing sequ s.t $\lim x_n = -\infty$

consider $A_n = \{w : X(w) \leq x_n\}$. $\{A_n\}$ is a decreasing sequ^n of events. Then $\lim_{n \to \infty} A_n = \phi$

Now $\lim P(A_n) = P(\lim A_n)$

$\Rightarrow \lim P(X \leq x_n) = P(\phi)$

$\Rightarrow \lim\limits_{x_n \to -\infty} F(x_n) = 0 \Rightarrow F(x) = 0$ as $x \to -\infty$

Take $\{x_n\}$ to be an increasing sequence $\ni$

$\lim x_n = +\infty$

$A_n = P\{\omega : X(\omega) \leq x_n\}$, $\lim A_n = \Omega$.

$\Rightarrow \lim\limits_{x \to \infty} F(x) = 1$.

(ii) Now $x_1 < x_2$ then $\{\omega : X(\omega) \leq x_1\} \subseteq \{\omega : X(\omega) \leq x_2\}$

$\Rightarrow P(\{\omega : X(\omega) \leq x_1\}) \leq P(\{\omega : X(\omega) \leq x_2\})$

$\Rightarrow F(x_1) = F(x_2)$

(iii) Let $\{x_n\}$ be a decreasing sequence s.t. $x_n \geq x$

and $\lim x_n = x$.

Consider $A_n = \{\omega : X(\omega) \leq x_n\}$

Then $\{A_n\}$ is a decreasing then $\lim A_n = \cap A_n$

$= (-\infty, x]$.

Now

Now

$$\lim P(A_n) = P(\lim A_n)$$

$$\Rightarrow \lim_{n\to\infty} F(x_n) = P\{(-\infty, x]\} = F(x)$$

Since the above holds for any sequence $\{x_n\}$ s.t. $\{x_n\}$ is decreasing to x. $\Rightarrow \lim_{h\to 0} F(x+h) = F(x)$

i.e. $F(x+) = F(x)$.

So $F(x)$ is right continuous.

Theorem: Given a prob function Q on $(\mathbb{R}, \mathcal{B})$ $\exists$ a cdf F satisfying $Q(-\infty, x] = F(x)$ $\forall x \in \mathbb{R}$. Conversely given a function F satisfying the three properties there exists a unique prob. function Q on $(\mathbb{R}, \mathcal{B})$ s.t. $Q(-\infty, x] = F(x)$.

Proof: Proof this result is a part of advanced theory of probability.

Remark (I) From the calculus we know that any monotone function is either continuous on $\mathbb{R}$ or it has atmost countable number of discontinuities. Thus any cdf $F_X(x)$ is either continuous on $\mathbb{R}$ or has only countable number of discontinuities.

(ii) We have for every $x \in \mathbb{R}$

$$F_X\left(x-\tfrac{1}{n}\right) \leq F_X(x) = F_X(x+), \quad n \in \mathbb{N}$$

$$\lim_{n\to\infty} F_X\left(x-\tfrac{1}{n}\right) \leq F_X(x) = F_X(x+)$$

i.e. $F_X(x-) \leq F_X(x) = F_X(x+)$

A distribution function F is continuous at x iff $F(a)=F(a-)$

(iii) For $x \in \mathbb{R}$

$$P_r(X<x) = P_X\left(\left(-\infty, x\right)\right)$$

$$= P_X\left(\bigcup_{n=1}^{\infty}\left(-\infty, x-\tfrac{1}{n}\right]\right)$$

$$= \lim_{n\to\infty} P_X\left(\left(-\infty, x-\tfrac{1}{n}\right]\right)$$

$$= \lim_{n\to\infty} F_X\left(x-\tfrac{1}{n}\right) = F_X(-x)$$

So $P_r\left(\{X<x\}\right) = F_X(x-), \ \forall x \in \mathbb{R}.$

(iv) For $-\infty < a < b < \infty$

$$P_r(X \leq b) = P_r\left(\{X \leq a\}\right) + P_r\left(\{a < X \leq b\}\right)$$

$$P_r\left(\{a < X \leq b\}\right) = P\left(\{X \leq b\}\right) - P\left(\{X \leq a\}\right)$$

$$= F_X(b) - F_X(a)$$

Similarly, for $-\infty < a < b < \infty$

$$\Pr(\{a < X \le b\}) = \Pr(\{X < b\}) - \Pr(\{X \le a\})$$
$$= F_X(b-) - F_X(a)$$

$$\Pr(\{a \le X \le b\}) = \Pr(\{X \le b\}) - \Pr(\{X < a\})$$
$$= F_X(b) - F_X(a-).$$

$$\Pr(\{a \le X < b\}) = \Pr(\{X < b\}) - P(\{X < a\})$$
$$= F_X(b-) - F_X(a-)$$

$$\Pr(\{X > b\}) = 1 - \Pr(\{X \le b\}) = 1 - F_X(b)$$

$$\Pr(\{X \geq b\}) = 1 - \Pr(\{X < b\}) = 1 - F_X(b-).$$

(V) For any $a \in \mathbb{R}$

$$\Pr(\{X = a\}) = \Pr(\{X \le a\}) - \Pr(\{X < a\})$$
$$= F_X(x) - F_X(-x).$$

Example: Consider the function $G : \mathbb{R} \to \mathbb{R}$ defined by

$$G(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{x}{3} & \text{if } 0 \le x < 1 \\ 1/2 & \text{if } 1 \le x < 2 \\ 2/3 & \text{if } 2 \le x < 3 \\ 1 & \text{if } x \geq 3 \end{cases}$$

(a) Show that G is d.f. of some r.v. X,

(b) Find $Pr(X = a)$ for various values of $a \in \mathbb{R}$

(c) Find $Pr(X < 3)$, $Pr(X \geq 1/2)$, $Pr(2 < X \leq 4)$, $Pr(1 \leq X < 2)$, $Pr(2 \leq X \leq 3)$ and $Pr(\frac{1}{2} < X < 3)$

solu^n clearly G is non-decreasing in $(-\infty, 0)$, $(0,1)$, $(1,2)$, $(2,3)$ & $(3,\infty)$. However

$$G(0) - G(0-) = 0 \geq 0$$

$$G(1) - G(1-) = \frac{1}{2} - \frac{1}{3} > 0$$

$$G(2) - G(2-) = \frac{2}{3} - \frac{1}{2} > 0$$

$$G(3) - G(3-) = 1 - 2/3 > 0$$

It follows that G is non-decreasing.

Now clearly G is continuous (and hence right continuous on $(-\infty, 0)$, $(0,1)$, $(1,2)$, $(2,3)$ & $(3,\infty)$. Moreover

$G(0+) - G(0) = 0 - 0 = 0$

$G(1+) - G(1) = 1/2 - 1/2 = 0$ $\Rightarrow$ G is right continuous on $\mathbb{R}$

$G(2+) - G(2) = 2/3 - 2/3 = 0$

$G(3+) - G(3) = 1 - 1 = 0$

$\lim_{x \to \infty} G(x) = 1$ & $\lim_{x \to -\infty} G(x) = 0$.

$\Rightarrow$ G is a d.f of some random variable X.

(b) The set of discontinuity points of F is

$$D = \{1, 2, 3\}$$

Thus $Pr(X = a) = G(a) - G(a-) = 0 \ \forall a \neq 1, 2, 3$

$Pr(X=1) = G(1) - G(1-) = \frac{1}{2} - \frac{1}{3} = 1/6$

$Pr(X=2) = G(2) - G(2-) = \frac{2}{3} - \frac{1}{2} = 1/6$

$Pr(X=3) = G(3) - G(3-) = 1 - 2/3 = 1/3$

(c) $Pr(X < 3) = G(3-) = 2/3$

$Pr(X \geq 1/2) = 1 - G(\frac{1}{2}-) = 1 - \frac{1}{6} = 5/6$

$Pr(2 < X \leq 4) = G(4) - G(2) = 1 - 2/3 = 1/3$

$Pr(1 \leq X < 2) = G(2-) - G(1-) = \frac{1}{2} - \frac{1}{3} = 1/6$

$Pr(2 \leq X \leq 3) = G(3) - G(2-) = 1 - 1/2 = 1/2$

$Pr(\frac{1}{2} < X < 3) = G(3-) - G(1/2) = 2/3 - 1/6 = 1/2$.