Random Variables

Review: Mapping and Functions

Definition: Given two abstract spaces S and A, an A-valued function is defined as

$$f : S \longrightarrow A$$

where, "f" assigns a specific element from A to each element of S

So, assuming that $\omega \in S$, we can write

$$f(\omega) \in A, \quad \forall \omega \in S$$

Within $f : S \to A$ S: Domain → Range when all elements are image
A: Co-Domain

When we study real valued fn^c, the def^n appears as follows: $f : \mathbb{R} \to \mathbb{R}$ | $S = \mathbb{R}$, $A = \mathbb{R}$

Definition: Let's consider two sets $F \subset S$ and, $G \subset A$

we define image of F under f as:

$$f(F) = \{ a \in A : a = f(\omega) \text{ for some } \omega \in F \}$$

and, the pre-image or inverse-image of $G \subset A$ under "f" is defined as:

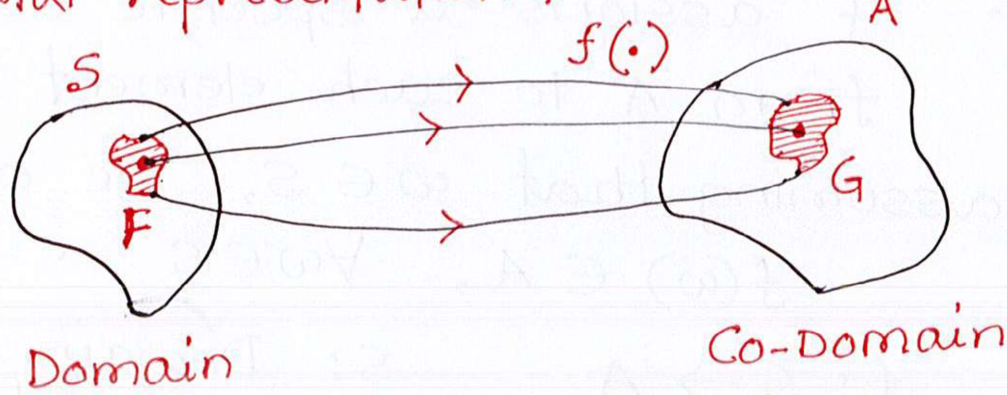
$$f^{-1}(G) = \{ \omega \in S : f(\omega) \in G \}$$

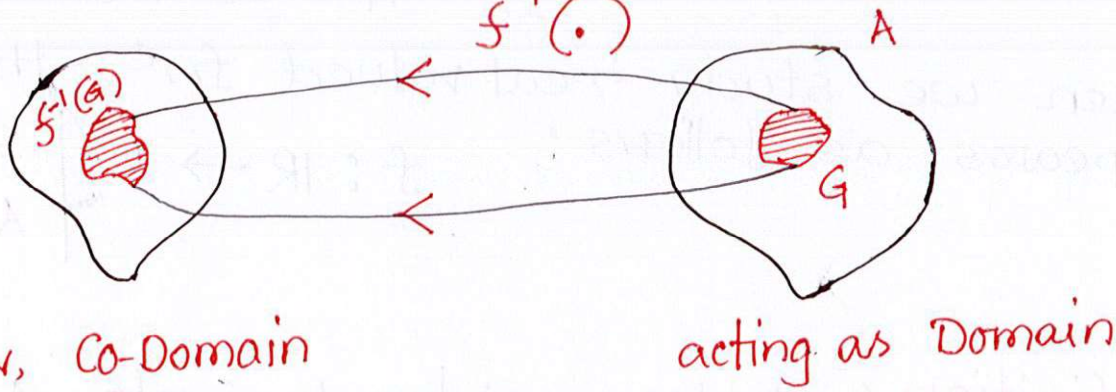
That is,

$f(F)$ is the set of all points in A obtained by mapping the points of F under the definition of "f"

$f^{-1}(G)$ is the set of all points $\omega \in S$ that map to G

Pictorial representation :





⊞ Generally, in calculus we study real-valued functions of a real variable

$$f : \mathbb{R} \longrightarrow \mathbb{R}$$

as we see, $S = \mathbb{R}$ and $A = \mathbb{R}$

▣ Why do we study random variables

We perform experiments to measure a quantity, characterize a system. All the experiments, generally, have different form of outcomes.

So, it is common that outcomes of experiments are interpreted, and often

We may not be interested in the precise outcome, rather a modified form using functional mapping be of our interest.

generates a number

Examples of experiments:

For instance,

- Suppose, we are measuring induced electric current in antenna due to random thermal motion of charges
- Suppose, we have a coin and we perform N tosses. Now, we are interested in the number of heads in the sequence generated after N-tosses.

interestingly,

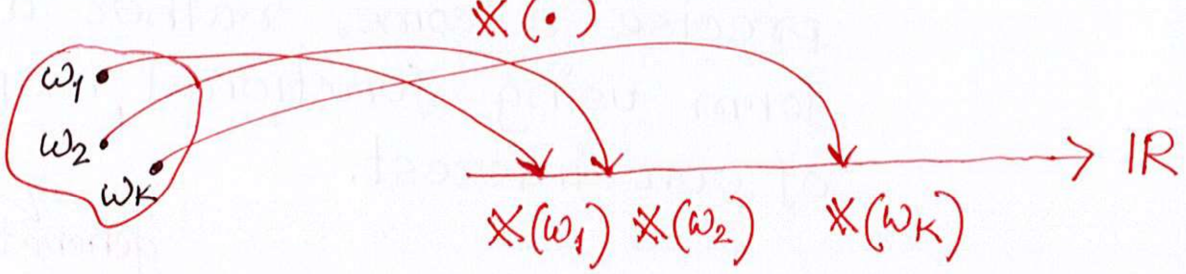
we don't consider the exact sequence

Here, each outcome produces a number from a numerical function

Intuitively,

Given, $(S, \mathcal{F}, P)$, a random variable is a mapping from the sample space S to the real line $\mathbb{R}$. If we denote random variable by X, then

$X : S \to \mathbb{R}$. The pictorial representation is as follows:



Here, $\omega_1, \omega_2 \ldots \omega_K$ are the outcomes and the corresponding numeric values are $X(\omega_1), X(\omega_2) \ldots X(\omega_K)$.

so, the numerical value becomes random

So, in a random experiment, the value of ω is chosen as random because of the possible outcome of the experiment.

Observation:

- Random variable X is a function from sample space S to the real line $\mathbb{R}$
- The term "random" indicates the underlying randomness of choosing an element ω from the sample space

- When the value of ω is fixed, the mapping function $X(\omega)$ generates a fixed value from $\mathbb{R}$
- Interestingly,

  Probability measure is associated with events, whereas

  Random variable is associated with each elementary outcome of the sample space.

In short, the randomness in the observed value $X(\omega)$ for the outcome ω is due to the randomness of $\omega \in S$ in any random experiment defined as:

$$(S, \mathcal{F}, P)$$

And, the mapping $X : S \to \mathbb{R}$ is not random. Instead, it is fixed and deterministic.

⊞ Observation:

Not all subsets of the sample space are considered as events, and the same goes for the random variable. That is, not all functions from S to $\mathbb{R}$ are considered as random variable

So, random variable X is a fn^c with traits as:

The function $X: S \to \mathbb{R}$ has the property that its output inherits a probability measure from P in the probability space defined as $(S, \mathcal{F}, P)$

Let's recall the concept of Borel set again

Given S and a family of subsets $G = \{A_i, i \in I\}$ of S, the σ-field (sigma-field) generated by G, denoted as $\sigma(G)$, is the smallest σ-field containing all the subsets in G.

So, when $S = \mathbb{R}$,
we need the smallest σ-field containing all the open intervals

$$(a, b), \quad a < b, \quad a, b \in \mathbb{R}$$

That is, we want

$$\sigma(G), \text{ where } G = \{(a,b), \forall a, b \in \mathbb{R} \text{ and } a < b\}$$

Interestingly,
$\sigma(G)$ contains all the countable sequences of set operations
- Union $\cup$
- Intersection $\cap$
- complement $-$

along with the open intervals

on any collection of open intervals

With understanding on σ-field, we can now define Borel field and Borel sets.

Borel field : Given the set of real numbers $\mathbb{R}$, the Borel field of $\mathbb{R}$, denoted as $B(\mathbb{R})$, is defined as the σ-field generated by open intervals

$$G = \{(a,b) : \forall a, b \in \mathbb{R} \text{ and } a < b\}$$

Borel Set : The members of $B(\mathbb{R})$ are known as Borel sets.

As the Borel field contains all the open interval, it also contains the below intervals:

$$(-\infty, b) = \bigcup_{n=1}^{\infty} (b-n, b) = \lim_{m\to\infty} (-m, b)$$

$$(a, \infty) = \bigcup_{n=1}^{\infty} (a, a+n) = \lim_{m\to\infty} (a, m)$$

$$\{a\} = \bigcap_{n=1}^{\infty} \left(a-\tfrac{1}{n}, a+\tfrac{1}{n}\right) = \lim_{n\to\infty} \left(a-\tfrac{1}{n}, a+\tfrac{1}{n}\right)$$

$$= \{a\}$$

It also follows that

$$[a,b) = \{a\} \cup (a,b) \in B(\mathbb{R})$$
$$(a,b] = (a,b) \cup \{b\} \in B(\mathbb{R})$$
$$[a,b] = \{a\} \cup (a,b) \cup \{b\} \in B(\mathbb{R})$$

In addition to all these intervals, all finite and countable sequences of set operations of unions ($\cup$), intersections ($\cap$), and complements ($^-$) of those intervals will be in the Borel Field $B(\mathbb{R})$.

In short,

$B(\mathbb{R})$ includes all the subsets of $S = \mathbb{R}$ that we could be interested in. One might ask, if it is sufficient or not.

Interestingly, proving that $B(\mathbb{R}) \neq \wp(\mathbb{R})$ is always difficult.

Yes, there may be some sets that are not in Borel field of $\mathbb{R}$, but those sets are strange and of no interests in probability problems.

Also, often, we don't need the whole real line $\mathbb{R}$ as the sample space. That is,

$$S = A, \text{ where } A \subset \mathbb{R}$$

example: $A = [0, 1]$

So, the event space should be the Borel field of A, that is $B(A)$

We can cut the $B(\mathbb{R})$ to obtain the $B(A)$

$$B(A) = \{ F \cap A : \forall F \in B(\mathbb{R}) \}$$

With the revision of Borel Field and Borel set, we can now explore the concept of Random Variable.

It is not random

It is not a variable

Rather, it is a fn^c that maps random outcomes "ω" to a real numeric value through a numerical fn^c.

So, random variable, defined as a fn^c follow the mapping $X : S \to \mathbb{R}$ inherits a probability measure from P in the underlying probability space $(S, \mathcal{F}, P)$

Simply, $X(\cdot)$ should have the property that for any $A \in B(\mathbb{R})$, we can compute the probability $P_X(A)$.

And any function $X(\cdot)$ such that we can calculate $P_X(\cdot)$ for all $A \in B(\mathbb{R})$ is called a Borel measurable function

[1] Measurable function

Let's consider $(S, \mathcal{F})$ be a measurable space. Then a function $f: S \rightarrow \mathbb{R}$ is $\mathcal{F}$-measurable function if the pre-image of every Borel set A, is an $\mathcal{F}$-measurable subset of S. $A \in B(\mathbb{R})$

The pre-image of A is defined as:

$$f^{-1}(A) \triangleq \{ \omega \in S \mid f(\omega) \in A \}$$

As per the definition of measurable function, $f: S \rightarrow \mathbb{R}$ is an $\mathcal{F}$-measurable fn^c if $f^{-1}(A)$ is an $\mathcal{F}$-measurable subset of S for every Borel set A.

extending the def^n for probability space.

Given $(S, \mathcal{F}, P)$ a random variable X is an $\mathcal{F}$-measurable function $X: S \rightarrow \mathbb{R}$ from $(S, \mathcal{F})$ to $\mathbb{R}$, with property as follows

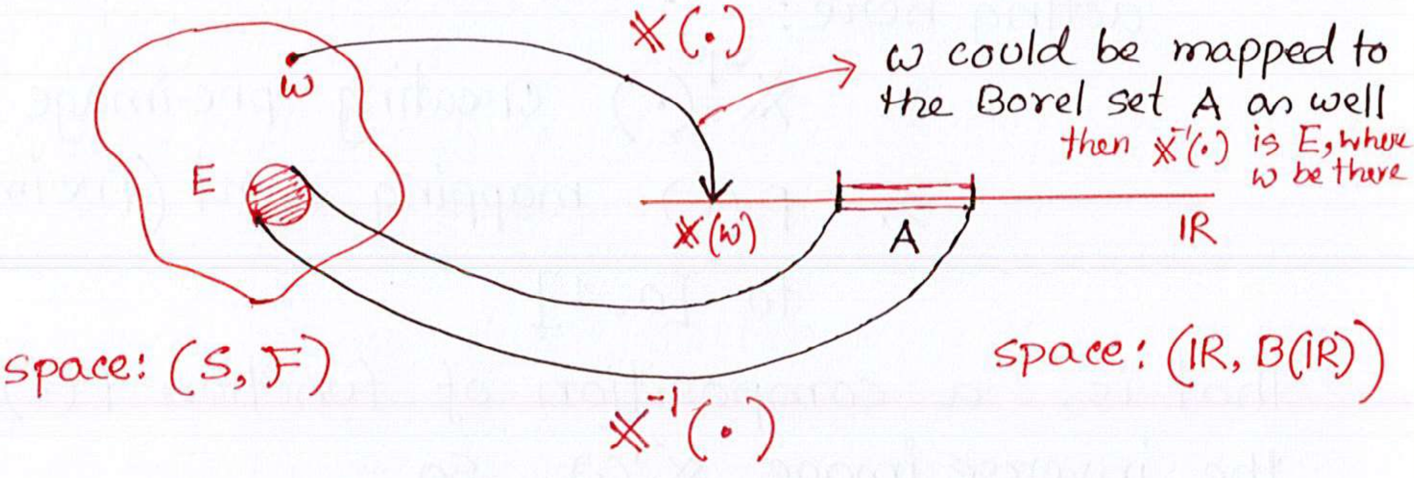
$$X^{-1}(A) = \{ \omega \in S \mid X(\omega) \in A \} \quad \text{for all } A \in B(\mathbb{R})$$

$$\& \quad X^{-1}(A) \in \mathcal{F}$$

In other words, every Borel set $A \in B(\mathbb{R})$ is the event E in the event space $\mathcal{F}$.

So, X is a random variable that maps every $\omega \in S$ to the real line $\mathbb{R}$.

We can represent the random variable X definition pictorially as:



Now, the pre-image

$$X^{-1}(A) \triangleq \{\omega \in S \mid X(\omega) \in A\} \in \mathcal{F}$$

indicates that $X^{-1}(A)$ is an event.

means

It has the associated probability measure.

Probability law of a random variable X

Let's denote P_X as the probability law of the random variable X. So,

the probability law P_X of a random variable X is a function P_X

$$P_X : \mathcal{B}(\mathbb{R}) \longrightarrow [0, 1]$$

defined as

$$P_X(A) \triangleq \left(\{\omega \in S \mid X(\omega) \in A\}\right)$$

$$= P(X^{-1}(A))$$

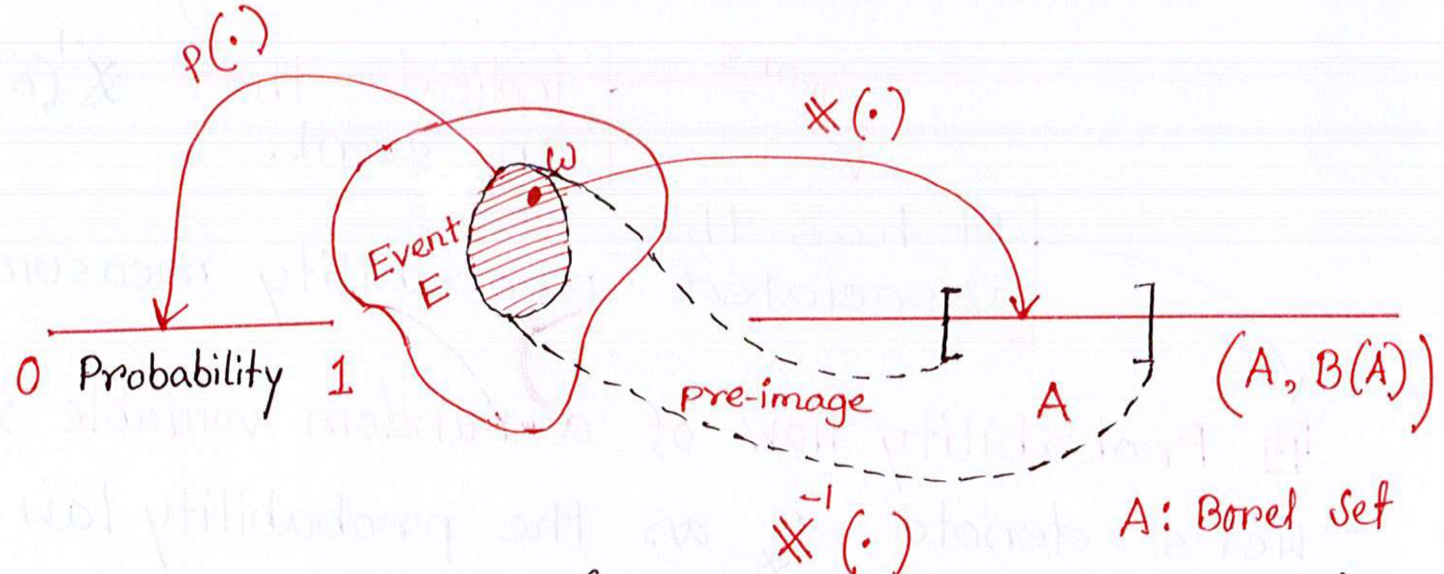
As we observe, we have two functions acting here:

1. $\mathbb{X}^{-1}(\cdot)$ creating pre-image
2. $P(\cdot)$ mapping event (pre-image) to $[0,1]$

That is, a composition of function $P(\cdot)$ and the inverse image $\mathbb{X}^{-1}(\cdot)$. So,

$$P_{\mathbb{X}}(\cdot) = P(\cdot) \circ \mathbb{X}^{-1}(\cdot)$$

composition notation



With a short-hand notation, the probability law could be written as:

$$P_{\mathbb{X}} = \underbrace{P}_{\text{Original probability measure}} \circ \mathbb{X}^{-1}$$

We just use the interested Borel set right next to it

From the discussion done so far, we can say that,

$\boxed{A}$ function $X : S \to \mathbb{R}$ satisfying the condition:

article

pre-image under X
$$\left[\begin{array}{l} \forall A \in \mathcal{B}(\mathbb{R}), \\ X^{-1}(A) = \{\omega \in S \mid X(\omega) \in A\} \in \mathcal{F} \end{array}\right.$$

is a Borel measurable function from $(S, \mathcal{F})$ to $\mathbb{R}$.

So, measurablity of X insures that we can compute the probability of any Borel set A. That is, we can compute the probability that

$$\{X\ A\} = \{\omega \in S : X(\omega) \in A\} \in \mathcal{F}$$
$$\text{for all } A \in \mathcal{B}(\mathbb{R})$$

Interestingly, we can ask the question that how restrictive this condition is that the mapping be Borel measurable. Specifically,

How restrictive is measurablity on the family of functions that are valid random variables ?

Answer: Not very restrictive.

As long as the event space in the prabability space defined for original experiment, most of the function we can think of is Borel measurable

For instance,

Assume that $(S, \mathcal{F}, P)$ is a probability space and takes shape as follows:

$$S = \mathbb{R}, \quad \mathcal{F} = \mathcal{B}(\mathbb{R})$$

Then, almost any function defined as: $X : \mathbb{R} \to \mathbb{R}$ will be a random variable.

In fact, for a probability space $(\;\; , \mathcal{B}(\mathbb{R}), P)$ when mapped by a random variable $X(\cdot)$ to $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$, that is:

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

all of the followings will be valid random variable:

- Continuous functions
- Polynomials
- Step functions: $K \cdot u(x-a)$
- All the indicator functions

$$1_A(\omega) = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in A \end{cases}$$

for all $A \in \mathcal{B}(\mathbb{R})$

- Trigonometric functions
- Limits of sequences of measurable functions: $f_1(x), f_2(x), \dots \dots f_n(x)$

$$\lim_{n \to \infty} f_n(x) = f(x)$$

- Finite and countable sums of measurable functions $f_1(x), f_2(x), f_3(x) \dots \dots f_n(x) \dots$

So, $$f(x) = \sum_n f_n(x)$$

As we observe, almost any function we can think of is measurable. Precisely,

It is very difficult to describe a fn^c $X(\cdot)$ that is not measurable.

For instance, If the set $A \in \mathcal{P}(\mathbb{R}) - \mathcal{B}(\mathbb{R})$

then $\underbrace{1_A(\omega)}_{}$ is not Borel measurable

↳ indicator function of such a set A

while we know that such set A exists, it is not possible to describe them.

One common and known non-measurable set would be Vitaliset, found by Giuseppe Vitali.

Range space of a fn^c

Let's define function $f: S \to A$, where S, A are domain and co-domain respectively.

Then, the range space of f is the image $f(S) \subset A$.

For instance, $f: \mathbb{R} \to \mathbb{R}$, $f(x) = x^2$

$$f(\mathbb{R}) = [0, \infty) = \text{range space}$$

Given $(S, \mathcal{F}, P)$ and a Random Variable (R.V), the RV takes on values in the range space

$$\mathcal{R} = \underbrace{f(S)}_{\text{Here, } f(\cdot) \text{ is the RV } \mathbb{X}(\cdot)} \subset \mathbb{R}$$

We already have shown that $\mathbb{X}$ is a measurable function, so, all the subsets of the form

$$\mathbb{X}^{-1}(G) = \{\omega \in S \mid \mathbb{X}(\omega) \in G\}$$

where, $G \subset \mathcal{B}(\mathcal{R})$

must be in the event space $\mathcal{F}$ of the given $(S, \mathcal{F}, P)$. That is, we can compute the probability $P(G)$ that $\mathbb{X}(\omega)$ takes on a value in $G \in \mathcal{B}(\mathcal{R})$

Here,
$$\begin{aligned} P_X(G) &= P(\mathbb{X}^{-1}(G)) \\ &= P(\{\omega \in S \mid \mathbb{X}(\omega) \in G\}), \forall G \in \mathcal{B}(\mathcal{R}) \end{aligned}$$

That is, $$(S, \mathcal{F}, P) \xrightarrow{\mathbb{X}(\cdot)} (\mathcal{R}, \mathcal{B}(\mathcal{R}), P_X)$$

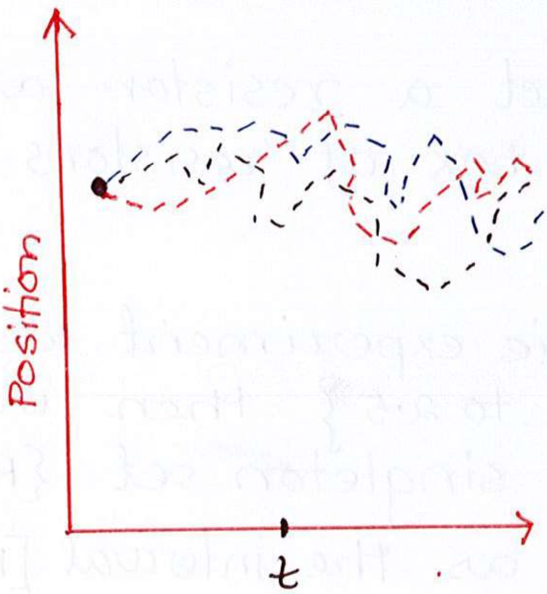
As a result, we can focus on the new random experiment $(\mathcal{R}, \mathcal{B}(\mathcal{R}), P_X)$

We can ask a fundamental question

Why do we describe a random variable (RV) as a mapping in the first place?

- Insight on physical problems, such as induced current on a placed antenna because of thermal movement of charges at any given instance of time
- Secondly, we can define multiple random variables on a random experiment
- In fact, we can generalize the simple random variables to random processes or stochastic processes.

Sequence of RVs at a given time point when the experiment involves dynamic evolution in time. For instance, observing a particle moving randomly in time.



If the position is tracked and multiple observations are made, position of the particle at a given time point will be a sequence of RVs.

A few examples:

- Let $(S, \mathcal{F}, P)$ be a probability space and assume that event $A \in \mathcal{F}$. We can define random variable
$$X(\omega) = 1_A(\omega)$$

- Again, suppose a die has six colored side and the die is rolled as the random experiment. So,
$$S = \{\text{Red, Green, Yellow, Blue, Cyan, Magenta}\}$$
$$\mathcal{F} = \mathcal{P}(S)$$
We can define a random variable (RV) as follows:
$$X(R) = 1, \; X(G) = 2, \; X(Y) = 3, \; X(B) = 4,$$
$$X(C) = 5, \; X(M) = 6$$
measurable RV

- Let's say we select a resistor at random from a box of resistors and measure its value.

For instance, for the die experiment we choose a Borel set $\{1.5 \text{ to } 2.5\}$ then in pre-image we will have singleton set $\{R\}$
If the Borel set is taken as the interval [1.5
Then, pre-image of that would be
$\{\text{Red, Green}\}$, which will be in the power set

That is, measurable

So, the discussion shows that

we can think of a random variable in two ways:

Mathematically, a random variable is a measurable function from $(S, \mathcal{F}, P)$ to $\mathbb{R}$.

↳ precise mathematical way to think about a RV

Intuitively, Something that takes on real values at random.

The outcomes of a random experiment, where real values are generated at random

$(\mathcal{C}, \mathcal{B}(\mathcal{C}), P)$, where $\mathcal{C} \subset \mathbb{R}$

↳ set of all possible values that can be taken of at random

Given a random variable with range space $\mathcal{C} \subset \mathbb{R}$, we can construct a probability space $(\mathcal{C}, \mathcal{B}(\mathcal{C}), P_X)$.

Then, the question is –

How do we describe P_X?

Two cases
- → Discrete
- → Continuous

Discrete case : For P_X, we can use the probability mass function (pmf), denoted as

$$P_X(x) = P_X(\{x\}), \ \forall x \in \mathcal{C}$$

with properties : (i) $P_X(x) \geq 0$, $\forall x \in \mathcal{C}$

(ii) $\sum_{x \in \mathcal{C}} P_X(x) = 1$

Continuous case : For continuous case, we use the probability density function (pdf) to assign

$$P_X(G) = \int_G f_X(x)\,dx, \quad \forall G \in \mathcal{B}(\mathcal{C})$$

with properties as : (i) $f_X(x) \geq 0$, $\forall x \in \mathcal{C}$

(ii) $\int_{\mathcal{C}} f_X(x)\,dx = 1$

continuous case is the uncountable sample space

However, even in the discrete case we have enormously large number of events and assigning probabilities to all those are not simple task. To make the process easier different mathematical tools are used and one such tool is the

Cumulative Distribution Function, shortly known as CDF

Definition of CDF

Given a random variable X defined on $(S, \mathcal{F}, P)$ that includes a new probability space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$, the cumulative distribution function "cdf" of X is defined as

$$F_X(\alpha) = P_X((-\infty, \alpha]) = P_X(\{x : x \leq \alpha\})$$

alpha

$$= P(\{\omega \in S \mid X(\omega) \leq \alpha\})$$

$$= P(X^{-1}((-\infty, \alpha])), \quad \alpha \in \mathbb{R}$$

As observed, the cdf notion accomodates alternative options.

For instance, $\underbrace{\{X \leq \alpha\}}_{\text{event}}$ or $P(\{X \leq \alpha\})$

represents $\underbrace{\{\omega \in S \mid X(\omega) \leq \alpha\}}_{\text{set of all } \omega \text{ in the sample space as defined in } (S, \mathcal{F}, P)} \in \mathcal{F}$

Essence of $F_X(\alpha)$: If we know $F_X(\alpha)$ it completely specifies the probability measure $P_X(\cdot)$

We know that. It is possible to construct any set $F \in \mathcal{B}(\mathbb{R})$, where F acts as the event, using countable sequence of set operations on the intervals of the form $(-\infty, x_n]$.

Although semi-close semi-open intervals are considered, it is actually possible. Because,

we already have seen that any event F in the Borel field by performing a countable sequence of set operations on the open interval of the form (α_n, β_n)

we saw that it is not hard to get intervals of the form $(-\infty, x_n)$ from (α_n, β_n) types of intervals.

For instance,

$$[a, b) = \{a\} \cup (a, b) \in B(\mathbb{R})$$

$$(a, b] = (a, b) \cup \{b\} \in B(\mathbb{R})$$

$$[a, b] = \{a\} \cup (a, b) \cup \{b\} \in B(\mathbb{R})$$

So, we write $F_X(x_n) = P_X((-\infty, x_n])$

That is, starting from cdf $F_X(\alpha)$, it is possible to get $P_X(\cdot)$ for any set (event) we are interested in.

means that CDF is a complete probabilistic description of any random experiment

So, the cdf is defined as:

$$F_X(x) = P(\underbrace{\{X \leq x\}}_{\text{event stating that } X \leq x}), \quad \forall x \in \mathbb{R}$$

$$= P_X((-\infty, x]), \quad \forall x \in \mathbb{R}$$

$$= P(\{\omega \in S \mid X(\omega) \leq x\}), \quad \forall x \in \mathbb{R}$$

Properties of cdf:

1. $F_X(+\infty) = 1$ and $F_X(-\infty) = 0$

Explanation: $$F_X(+\infty) = P(\underbrace{\{X \leq +\infty\}}_{\text{becomes } S})$$

$$= P(S) = 1$$

$$F_X(-\infty) = P(\underbrace{\{X \leq -\infty\}}_{\text{impossible event } \emptyset})$$

$$= P(\emptyset) = 0$$

2. If $x_1 < x_2$, then $F_X(x_1) \leq F_X(x_2)$

Explanation: $$F_X(x_1) = P_X(-\infty, x_1])$$

$$F_X(x_2) = P_X((-\infty, x_2])$$

Observing the real line, we can write

$$(-\infty, x_1] \subset (-\infty, x_2]$$

we can write $(-\infty, x_2]$ as the union of two disjoint sets

$$(-\infty, x_2] = (-\infty, x_1] \cup (x_1, x_2]$$

$$\Rightarrow P_{X}((-\infty, x_2]) = P_{X}\left((-\infty, x_1] \cup (x_1, x_2]\right)$$

$$= P_{X}((-\infty, x_1]) + P_{X}((x_1, x_2])$$

$$\Rightarrow F_{X}(x_2) = F_{X}(x_1) + \underbrace{P_{X}((x_1, x_2])}_{\text{Which is greater than zero}}$$

$$\Rightarrow F_{X}(x_2) \geqslant F_{X}(x_1)$$

3. $P(\{X > \alpha\}) = 1 - F_{X}(\ )$

The event $\{X > \alpha\} \equiv \overline{\{\ \leqslant \alpha\}}$

As, $\{X > \alpha\}$ and $\{X < \alpha\}$ are disjoint their union

$$\{X > \alpha\} \cup \{X \leqslant \alpha\} = S$$

$$\Rightarrow P(\{X > \alpha\}) + P(\{X \leqslant \alpha\}) = P(S)$$

$$\Rightarrow P(\{X > \alpha\}) = 1 - P(\{X \leqslant \alpha\})$$

$$\Rightarrow P(\{X > \alpha\}) = 1 - F_{X}(\alpha)$$

4. If $x_1 < x_2$, then $P(\{x_1 < \ \leqslant x_2\}) = F_{X}(x_2) - F_{X}(x_1)$

We can write $(-\infty, x_2]$ as the union of two disjoint set

$$(-\infty, x_2] = (-\infty, x_1] \cup (x_1, x_2]$$

$$\Rightarrow P(\{-\infty < X \leqslant x_2\}) = P\left(\{-\infty < X \leqslant x_1\} \cup \{x_1 < X \leqslant x_2\}\right)$$

Using the definition, we can write

$$F_{\mathbb{X}}(x_2) = F_{\mathbb{X}}(x_1) + P(\{x_1 < \mathbb{X} \leqslant x_2\})$$

$$\Rightarrow P(\{x_1 < \mathbb{X} \leqslant x_2\}) = F_{\mathbb{X}}(x_2) - F_{\mathbb{X}}(x_1)$$

Proved

5. Right continuity of CDF

If $F_{\mathbb{X}}(\cdot)$ is right continuous, it means that

$$\forall x \in \mathbb{R}, \quad \lim_{\underbrace{\varepsilon \downarrow 0}_{\text{epsilon is approaching zero}}} F_{\mathbb{X}}(x+\varepsilon) = F_{\mathbb{X}}(x)$$

Let's consider that $\varepsilon_n \geqslant 0$ is any non-negative sequence, such that $\varepsilon \downarrow 0$ as $n \to \infty$

Now, consider the following

$$\underbrace{\lim_{n\to\infty} F_{\mathbb{X}}(x+\varepsilon_n)}_{} = \lim_{n\to\infty} P(\{\mathbb{X} \leqslant x+\varepsilon_n\})$$

we use ε_n here, as a trick so that we can deal with countable unions & intersections

As in, $\lim_{\varepsilon\downarrow 0} F_{\mathbb{X}}(x+\varepsilon)$ we have uncountable ε as it is continuous and goes to zero

But, we cannot deal with uncountable unions or intersections

That is,

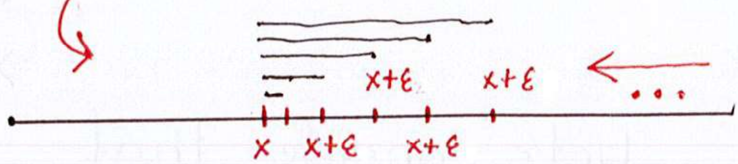
$$\lim_{\varepsilon \downarrow 0} F_{\mathbb{X}}(x+\varepsilon) = \lim_{\varepsilon \downarrow 0} P(\{\mathbb{X} \leq x+\varepsilon\})$$

$$= \lim_{n \to \infty} P(\{\mathbb{X} \leq x+\varepsilon_n\})$$

We can write it in terms of ω as :

$$= \lim_{n \to \infty} P(\{\omega \mid \mathbb{X}(\omega) \leq x+\varepsilon_n\})$$

$$= P\Big(\bigcap_{n \in \mathbb{N}} \{\omega \mid \mathbb{X}(\omega) \leq x+\varepsilon_n\}\Big)$$



so, intersection keeps the smallest part from all the possibilities

$$= P(\{\omega \mid \mathbb{X}(\omega) \leq x\}) = P(\{\mathbb{X} \leq x\})$$

$$= F_{\mathbb{X}}(x)$$

Generally, CDF need not be continuous, but the right-continuity must be satisfied. Any function that satisfies the right-continuity, monotonicity as in 2, the properties shown in 1, is eligible to be the CDF of some random variable.

Another CDF property:

Assume, $P(\{X = x_0\})$ stands for probability that X is equal to the particular number x_0. Then,

$$P(\{X = x_0\}) = F_X(x_0) - F_X(x_0^-)$$

where,
- x_0^- is infinitesimally smaller point than x_0
- $F_X(x_0^-) = \lim_{\varepsilon \downarrow 0} F_X(x_0 - \varepsilon)$

We can prove it as follows:

$$F_X(x_0) - F_X(x_0^-)$$

$$= P_X((-\infty, x_0]) - \lim_{\varepsilon \to 0} F_X(x_0 - \varepsilon)$$

$$= P_X((-\infty, x_0]) - \lim_{\varepsilon \to 0} P_X((-\infty, x_0 - \varepsilon])$$

$$= \lim_{\varepsilon \downarrow 0} P_X((x_0 - \varepsilon, x_0]) = P_X(\{X = x_0\})$$