TA Session 4

22/08/2018

Agenda:

- Borel σ-algebra of subsets of $\mathbb{R}$
- Random variables - definition and examples
- Pushforward measure, pullback measure and CDF.

Borel σ-algebra on $\mathbb{R}$

Let us recall the defn of a σ-algebra of subsets of a non-empty set Ω.

Defn: A collection $\mathcal{F}$ of subsets of a non-empty set Ω is called a σ-algebra of subsets of Ω if:

a) $\Omega \in \mathcal{F}$

b) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$

c) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

We also saw some examples. Just to recall one of them:

$$\Omega = \{1,2,3,4\}$$

$$\mathcal{F} = \{\phi, \Omega, \{1\}, \{2\}, \{1,2\}, \{3,4\}, \{1,3,4\}, \{2,3,4\}\}.$$

The cases we have studied thus far dealt with finite Ω.

We also looked at one example where $\Omega = \{r \in \mathbb{Q} : 0 \leq r \leq 1\}$, the set of all rationals in $[0,1]$, but we did not touch too much on it.

We will now study the case when $\Omega = \mathbb{R}$. What could be our choice for σ-algebra? One possibility is $2^{\mathbb{R}}$. This is a VERY RICH collection, which poses a problem in terms of probability assignment. There is a theorem which claims that one cannot assign probabilities in a meaningful way if $\mathcal{F} = 2^{\mathbb{R}}$.

So, this rules out $\mathcal{F} = 2^{\mathbb{R}}$ as a possible choice. What is a good choice? We will see a construction by Borel.

<u>Construction</u>: Before we start the construction, we will need a defn.

<u>Defn (open set)</u>: A set $E \subseteq \mathbb{R}$ is said to be an <u>open set</u> if: for every $x \in E$, there exists an $\varepsilon > 0$ such that the interval $(x-\varepsilon, x+\varepsilon) \subset E$.

<u>Examples</u>:
1. Every open interval is an open set. Hence the name "open" interval.
2. $(1,2) \cup (3,4)$ is an open set.
3. $(-\infty, x)$ is an open set $\forall\ x \in \mathbb{R}$
4. ϕ and $\mathbb{R}$ are open sets by convention.

We shall now start the construction process. Let

$$\mathcal{O} = \{E \subseteq \mathbb{R} : E \text{ is open}\}$$

be the collection of all open ~~sets~~ sets of $\mathbb{R}$. Clearly, this collection does not have sets which are complements of open sets. Let us build a σ-algebra of subsets of $\mathbb{R}$ starting from $\mathcal{O}$ by including complements, countable unions and intersections. Denote this collection by $\mathcal{B}(\mathbb{R})$.

Such a collection is very difficult to visualise, but it has some nice properties:

a) $(-\infty, x) \in \mathcal{B}(\mathbb{R})$ for every $x \in \mathbb{R}$

b) $(-\infty, x] \in \mathcal{B}(\mathbb{R})$ $\forall$ $x \in \mathbb{R}$

c) $(x, \infty) \in \mathcal{B}(\mathbb{R})$ $\forall$ $x \in \mathbb{R}$

d) $[x, \infty) \in \mathcal{B}(\mathbb{R})$ $\forall$ $x \in \mathbb{R}$

e) $(a, b), (a, b], [a, b), [a, b] \in \mathcal{B}(\mathbb{R})$ $\forall$ $-\infty < a \leq b < \infty$.

From e) above, it follows that $\{x\} \in \mathcal{B}(\mathbb{R})$ $\forall$ $x \in \mathbb{R}$.

Defn: The complement of an open set is called a closed set.

Thus, $\mathcal{B}(\mathbb{R})$ contains all open subsets of $\mathbb{R}$ and all closed subsets of $\mathbb{R}$. It contains all sets we can visualise and are familiar to us. This is a rich collection of sets, which it turns out, is not as big as $2^{\mathbb{R}}$ in size!

In fact, $\mathcal{B}(\mathbb{R})$ can be shown to have a bijection with $\mathbb{R}$, and hence has cardinality equal to that of $\mathbb{R}$.
$\mathcal{B}(\mathbb{R})$ is what is known as a Borel σ-algebra of subsets of $\mathbb{R}$.

As a matter of fact, one can construct a Borel σ-algebra for any non-empty set Ω starting from the collection of all open subsets of Ω.

Hereinafter, whenever we talk of $\mathbb{R}$, we will always consider $\mathcal{B}(\mathbb{R})$ along with it.

Random Variables

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be given. Let us say we are interested in describing events in $\mathcal{F}$ in terms of real-life occurrences.

Examples: 1 Let $\Omega = \{1, \ldots, 6\}$

$$\mathcal{F} = 2^{\Omega}$$

$$\mathbb{P}(\{a\}) = \frac{1}{6} \quad \forall \; a \in \Omega.$$

Here, Ω corresponds to outcomes of a throw of a die.
$\mathbb{P}$ corresponds to a fair die probability assignment.
Let us say we want to associate "die throw resulting in 1, 2, 3, 5" with "I win 20 rupees", and "die throw resulting in 4 or 6" with "I lose 2 rupees".

2. Suppose we consider a communication system having one transmitter and one receiver. Let's say we are transmitting a signal $x(t)$ from the transmitter through a channel, and that the channel adds noise $n(t)$ at time t. The received output $y(t)$ is then modelled by

$$y(t) = x(t) + n(t).$$

We might be interested in modelling the transmission as successful if $y(t) > 100$, and not otherwise.

Such descriptions of events require us to associate numbers with every $\omega \in \Omega$. In other words, what we require is a map

$$f : \Omega \to \mathbb{R}.$$

But this does not do it. There are thousands of such maps we could come up with. The point is that those maps which actually make sure we can assign and talk about probabilities of events of the form "I get 20 rupees" or "$y(t) > 100$" are what we need.

So, can we define maps which help us talk about probabilities of events like these? Yes. Such maps are random variables.

$$(\Omega, \mathcal{F}, P) \xrightarrow{\quad f \quad} (\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

$$A_1 \xleftarrow{\quad f^{-1} \quad} \text{" I get 20 rupees "}$$

$$A_2 \xleftarrow{\quad f^{-1} \quad} \text{" } y(t) > 100 \text{ "}$$

Referring to the figure above, can we build a function f such that we can talk of

$P(A_1)$ = probability of I getting 20 rupees

$P(A_2)$ = probability that $y(t) > 100$.

Can we make sure that A_1 and A_2 are actually in $\mathcal{F}$ so that we can assign probabilities to them? YES!

<u>Def^n</u>: (random variable)

Let $(\Omega, \mathcal{F})$ be given. A function $X : \Omega \to \mathbb{R}$ is called a random variable if:

$$A = \{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F} \quad \forall\, x \in \mathbb{R}$$

$$\begin{tikzcd}
(\Omega, \mathcal{F}) \arrow[r, "X"] & (\mathbb{R}, \mathcal{B}(\mathbb{R})) \arrow[d] \\
A \arrow[u] & (-\infty, x] \arrow[l, "\text{Inverse image}"]
\end{tikzcd}$$

It turns out that this is equivalent to saying that:

$$A = \{\omega \in \Omega : X(\omega) \in B\} \in \mathcal{F} \quad \forall \ B \in \mathcal{B}(\mathbb{R})$$

$$\begin{tikzcd}
(\Omega, \mathcal{F}) \arrow[rr, "X"] & & (\mathbb{R}, \mathcal{B}(\mathbb{R})) \arrow[d] \\
A \arrow[u] & & B \arrow[ll, "\text{Inverse image}"']
\end{tikzcd}$$

<u>Examples</u>:

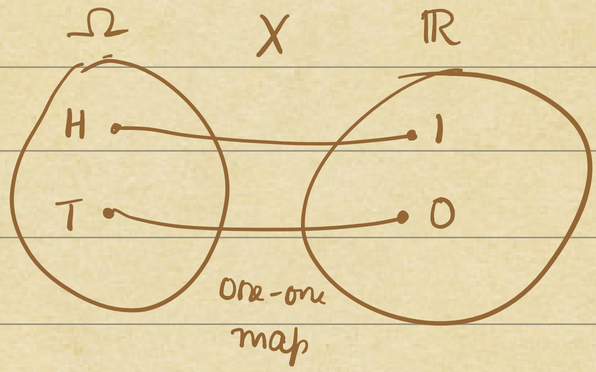
① $\Omega = \{H, T\}$

$\mathcal{F} = 2^{\Omega}$

$X : \Omega \to \mathbb{R}$ given by

$X(H) = 1$

$X(T) = 0.$



$X^{-1}(\mathbb{R}) = \Omega$

$$X^{-1}((-\infty, x]) = \begin{cases} \phi, & \text{if } x < 0 \\ \{T\}, & \text{if } 0 \leq x < 1 \\ \Omega, & \text{if } x \geq 1. \end{cases}$$

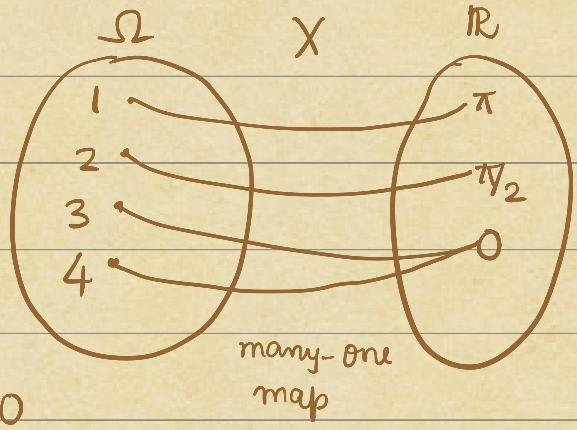
We see that $X^{-1}((-\infty, x]) \in \mathcal{F} \ \forall \ x \in \mathbb{R}$. Thus, X is a r.v.

② $\Omega = \{1, 2, 3, 4\}$

$\mathcal{F} = \{\phi, \Omega, \{1\}, \{2\}, \{1,2\}, \{3,4\}, \{1,3,4\}, \{2,3,4\}\}$

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

$$X(1) = \pi$$
$$X(2) = \pi/2$$
$$X(3) = 0$$
$$X(4) = 0.$$



Here, $$X^{-1}\big((-\infty, x]\big) = \begin{cases} \phi, & \text{if } x < 0 \\ \{3,4\}, & \text{if } 0 \le x < \pi/2 \\ \{2,3,4\}, & \text{if } \pi/2 \le x < \pi \\ \Omega, & \text{if } x \ge \pi. \end{cases}$$

Since $X^{-1}\big((-\infty, x]\big) \in \mathcal{F} \quad \forall\, x \in \mathbb{R}$, X is a rv.

3 We shall tweak the earlier example.

$\Omega = \{1, 2, 3, 4\}$.

$\mathcal{G} = \{\phi, \Omega, \{1\}, \{3\}, \{1,3\}, \{2,4\}, \{1,2,4\}, \{2,3,4\}\}$

X same as in example 2.

Then,

$$X^{-1}\big((-\infty, 0]\big) = \{3,4\} \notin \mathcal{G}$$

Thus, X is not a rv wrt this $\mathcal{G}$.

Hence, σ-algebra lies at the heart of the def^n of a rv.
A function which is a rv wrt a σ-algebra $\mathcal{F}$ need not be a rv to another σ-algebra $\mathcal{G}$. For this reason,

a function $X : \Omega \to \mathbb{R}$ that is a random variable wrt a σ-algebra $\mathcal{F}$ is also referred to as a $\mathcal{F}$-measurable function

Thus, given a σ-algebra $\mathcal{F}$, there are only some types of functions that can be random variables with respect to that σ-algebra.

Constructing random variables from σ-algebras

Let $\Omega = \{1, 2, 3, 4\}$

① $\mathcal{F} = \{\phi, \Omega\}$.

In this case, only constant functions can be random variables, i.e., functions of the form

$X : \Omega \to \mathbb{R}$ s.t.

$X(\omega) = c \quad \forall \ \omega \in \Omega$, where $c \in \mathbb{R}$ is some constant.

② $\mathcal{F} = \{\phi, \Omega, \{1,2\}, \{3,4\}\}$

Here, functions of the form

$$X(\omega) = \begin{cases} c_1, & \text{if } \omega = 1,2 \\ c_2, & \text{if } \omega = 3,4 \end{cases}$$

are random variables, where $c_1, c_2 \in \mathbb{R}$ are constants which may or may not be equal.

In general, for any $A \subseteq \Omega$, if

$$\mathcal{F} = \{\phi, \Omega, A, A^c\},$$

then only those functions that are of the form

$$X(\omega) = \begin{cases} c_1 & \text{if } \omega \in A \\ c_2 & \text{if } \omega \in A^c \end{cases}$$

are random variables.

(3) $\mathcal{F} = 2^{\Omega}$. Here, <u>every function</u> $X : \Omega \to \mathbb{R}$ will be a random variable.

<u>CDF - Cumulative Distribution function</u>

We saw the construction of $\mathcal{B}(\mathbb{R})$, the Borel σ-algebra of subsets of $\mathbb{R}$. Now, let us consider $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

Let $\mathbb{P}^*$ be a probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

That is,

(1) $\mathbb{P}^*(\mathbb{R}) = 1$, and

(2) If $A_1, A_2, \dots,$ is a countable collection of disjoint in $\mathcal{B}(\mathbb{R})$, then

$$\mathbb{P}^*\left(\biguplus_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}^*(A_i).$$

Now, let's look at $\mathbb{P}^*((-\infty, x])$, $x \in \mathbb{R}$. (Note that, by construction of $\mathcal{B}(\mathbb{R})$, $(-\infty, x] \in \mathcal{B}(\mathbb{R})$ $\forall$ $x \in \mathbb{R}$)

$\mathbb{P}^*((-\infty, x])$ is a function of $x \in \mathbb{R}$. Let us denote this as F. That is

$$F(x) = \mathbb{P}^*((-\infty, x]), \quad x \in \mathbb{R}.$$

F is what is called a cumulative distribution function, or CDF in short. It is a function from $\mathbb{R}$ to $[0,1]$.

Examples:

① $$F(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1, & \text{if } x \geq 0 \end{cases}.$$

② $$F(x) = \begin{cases} 0, & \text{if } x < 0 \\ p, & \text{if } 0 < x < 1 \\ 1, & \text{if } x \geq 1 \end{cases}$$

where $p \in [0,1]$ is called a Bernoulli CDF with parameter p.

③ $$F(x) = \begin{cases} 1 - e^{-\lambda x}, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

is called the exponential CDF with parameter $\lambda > 0$.

④ $$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-t^2/2} \, dt$$

is called a standard normal CDF.

We now relate CDF to the concept of random variables. We have the following two worlds: one of $(\Omega, \mathcal{F})$ and the other of $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$, as depicted in the figure below. Let $X : \Omega \to \mathbb{R}$ be a rv wrt $\mathcal{F}$.

$$
\begin{tikzcd}
(\Omega, \mathcal{F}) \arrow[r, "X \text{ rv}"] & (\mathbb{R}, \mathcal{B}(\mathbb{R})) \arrow[d] & \\
\{\omega \in \Omega : X(\omega) \le x\} \arrow[u] & (-\infty, x] \arrow[l, "\text{Inverse image}"'] & \forall x \in \mathbb{R}
\end{tikzcd}
$$

Let us redraw the above picture with $\mathbb{P}^*$ on the Right side world:

$$
\begin{tikzcd}
(\Omega, \mathcal{F}) \arrow[r, "X \text{ rv}"] & (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}^*) \arrow[d] \\
\{\omega \in \Omega : X(\omega) \le x\} \arrow[u] & (-\infty, x],\ \mathbb{P}^*\big((-\infty, x]\big) \arrow[l]
\end{tikzcd}
$$

We spoke of random variables with an intention of speaking about probabilities of events like "I win 20 rupees" or "$y(t) > 100$". If we want to do this, we need to convert event "I win 20 rupees" which are on the $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ side to an event on the $(\Omega, \mathcal{F})$ side and then speak about its probability.

Towards that end, let $\mathbb{P}$ be a probability measure on $(\Omega, \mathcal{F})$. If X is a $\mathcal{F}$-measurable rv, then we know from the definition of random variable that

$$\{\omega \in \Omega : X(\omega) \leq x\} \in \mathcal{F} \qquad \forall \; x \in \mathbb{R}.$$

Thus, the quantity $\mathbb{P}\left(\{\omega \in \Omega : X(\omega) \leq x\}\right)$ makes sense.

The picture now is as follows:

$$\begin{tikzcd}
(\Omega, \mathcal{F}, \mathbb{P}) \arrow[rr, "X \text{ rv}"] & & (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}^*) \arrow[d] \\
\{\omega \in \Omega : X(\omega) \leq x\} \arrow[u] & & (-\infty, x] \arrow[ll, "\text{Inverse image}"'] \\
\mathbb{P}\left(\{\omega \in \Omega : X(\omega) \leq x\}\right) & & \mathbb{P}^*\left((-\infty, x]\right).
\end{tikzcd}$$

In order to maintain consistency, we make sure that $\mathbb{P}$ and $\mathbb{P}^*$ are related such that

$$\mathbb{P}\left(\{\omega \in \Omega : X(\omega) \leq x\}\right) = \mathbb{P}^*\left((-\infty, x]\right) = F(x) \qquad \forall \; x \in \mathbb{R}.$$

Once we define $\mathbb{P}$ and $\mathbb{P}^*$ this way, we can choose to speak about probabilities of events either in $(\Omega, \mathcal{F}, \mathbb{P})$ terms or $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}^*)$ terms according to our convenience.

In the literature, $\mathbb{P}$ is called the <u>pullback measure of $\mathbb{P}^*$</u> and $\mathbb{P}^*$ is called the <u>pushforward measure of $\mathbb{P}$</u>.

Exercises:

1. Let Ω be a non-empty set. Let $A_1, \dots, A_n$ be a partition of Ω, i.e.,

$$\bigcup_{i=1}^{n} A_i = \Omega \quad \text{and} \quad A_i \cap A_j = \phi \ \ \forall\, i \neq j.$$

Construct a σ-algebra $\mathcal{F}$ as

$$\mathcal{F} = \left\{ \bigcup_{i \in J} A_i \ : \ J \subseteq \{1, \dots, n\} \right\}.$$

What kind of functions $X : \Omega \to \mathbb{R}$ are random variables in this setting?

2. Let $(\Omega, \mathcal{F})$ be given, and let $X : \Omega \to \mathbb{R}$ be a random variable. Define four new functions $S : \Omega \to \mathbb{R}$, $Y : \Omega \to \mathbb{R}$, $Z : \Omega \to \mathbb{R}$ and $T : \Omega \to \mathbb{R}$ as:

$$S(\omega) = |X(\omega)|, \quad \omega \in \Omega$$
$$Y(\omega) = -X(\omega) \quad \forall\, \omega \in \Omega$$
$$Z(\omega) = \max\{X(\omega), 0\} \quad \forall\, \omega \in \Omega$$
$$T(\omega) = -\min\{X(\omega), 2\} \quad \forall\, \omega \in \Omega$$

Prove (using def^n of random variable) that S, Y, Z and T are also random variables with respect to $(\Omega, \mathcal{F})$.