Random variables

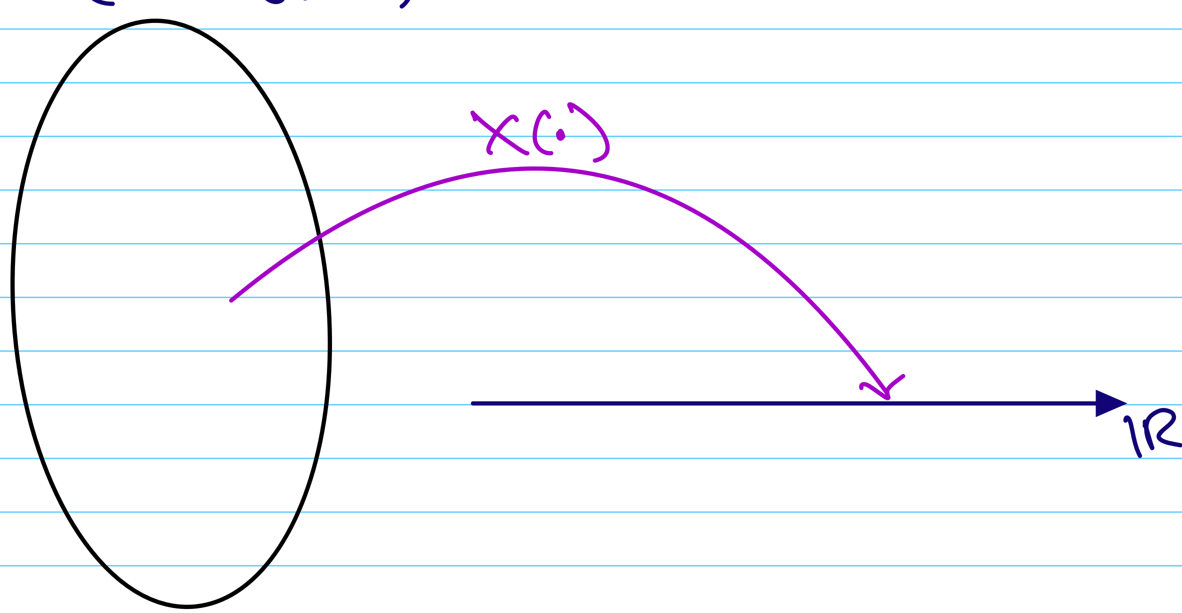
* in a random experiment, in the sample space Ω, we may be interested in the specific elementary outcome $\omega \in \Omega$.
* we don't really care about what specific precise little omega ω occurs.
* we are more interested in some numerical function of the elementary outcome.

* For example. in an experiment involving ten coin tosses, the experiment may only want to know the total number of heads, rather than the precise seq^n of heads and tails.

The term random variable is a misnomer, because a r.v is neither random, nor it is a variable. its not random, its a deterministic function, and its not a variable its a function

Random variables are Just numerical functions of your outcomes. its a function defined on sample space & give a real number in general.

$(\Omega, \mathcal{F}, \mathbb{P})$ ← Probability space



* its a mapping from sample space Ω to Real number Line

* once we fix ω, then the real number is fixed.

* The only source of randomness in the r.v is little omega $\omega \in \Omega$, That ω selected

by mother nature or something. that is the only random thing. after ω realizes there is no randomness in the r.v. its a deterministic mapping, give ω, then it will give you $X(\omega)$

* Not all subset's of Ω are considered event's, in the similar way Not all function's from $\Omega \longrightarrow \mathbb{R}$ are considered Random varibles

* Random variables are certain nice function's called measurable function's

Measurable function:

Let $(\Omega, \mathcal{F})$ be a measurable space. A function $X: \Omega \longrightarrow \mathbb{R}$, is said to be $\mathcal{F}$- measurable function if the Pre-Image of every Borel set is a $\mathcal{F}$-measurable set

$\Rightarrow$ $X : \Omega \longrightarrow \mathbb{R}$ is a $\mathcal{F}$-measurable function, if for every Borel set $B \in \mathcal{B}(\mathbb{R})$ the Pre-image $X^{-1}(B)$ is a $\mathcal{F}$-measurable set.

$$X^{-1}(B) = \{\omega \in \Omega \mid X(\omega) \in B\}$$

Definition:

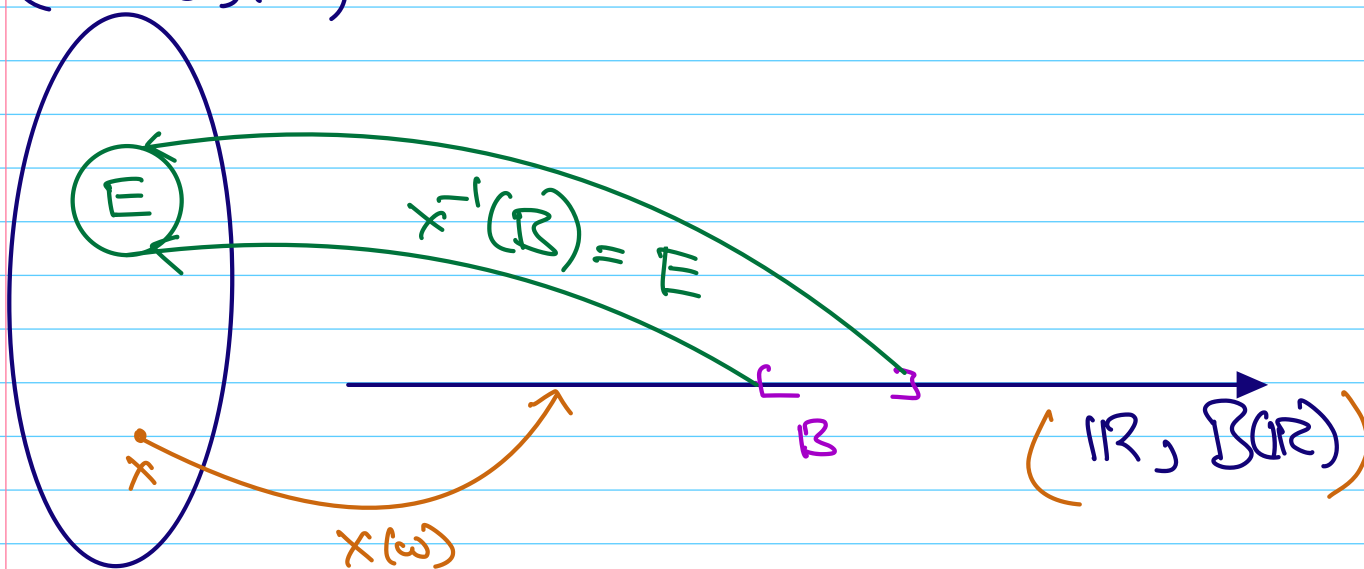
Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a Probability space. then the random variable X on the Proba-bility space is a $\mathcal{F}$-measurable function

$X : \Omega \longrightarrow \mathbb{R}$

* in other word's, for every Borel set B its Pre-image $X^{-1}(B)$ is an event.

* $X : \Omega \to \mathbb{R}$, X r.v maps ω to real number
* B is a Borel set i.e $B \in \mathcal{B}(\mathbb{R})$
* The inverse image of B is a event $E \in \mathcal{F}$

$(\Omega, \mathcal{F}, \mathbb{P})$



why do we insist that we want a r.v to be this very special function, $\mathcal{F}$-measurable function. what is the problem with saying all the function's mapping from $\Omega \to \mathbb{R}$ are random variables?

* Because, ultimately the probability measure is only defined for $\mathcal{F}$-mea-

surable sets. if you give me some arbitrary subsets of omega (Ω), it may not be an event, and we may not define probability measure for it.

* what we ultimately want to do is assign probability measure to Borel sets on $\mathbb{R}$.
* so, if we have a Borel set whose pre-image $X^{-1}(B)$ is not $\mathcal{F}$-measurable, we cannot assign probability to Borel set on $\mathbb{R}$. so a r.v has to be a $\mathcal{F}$-measurable function.

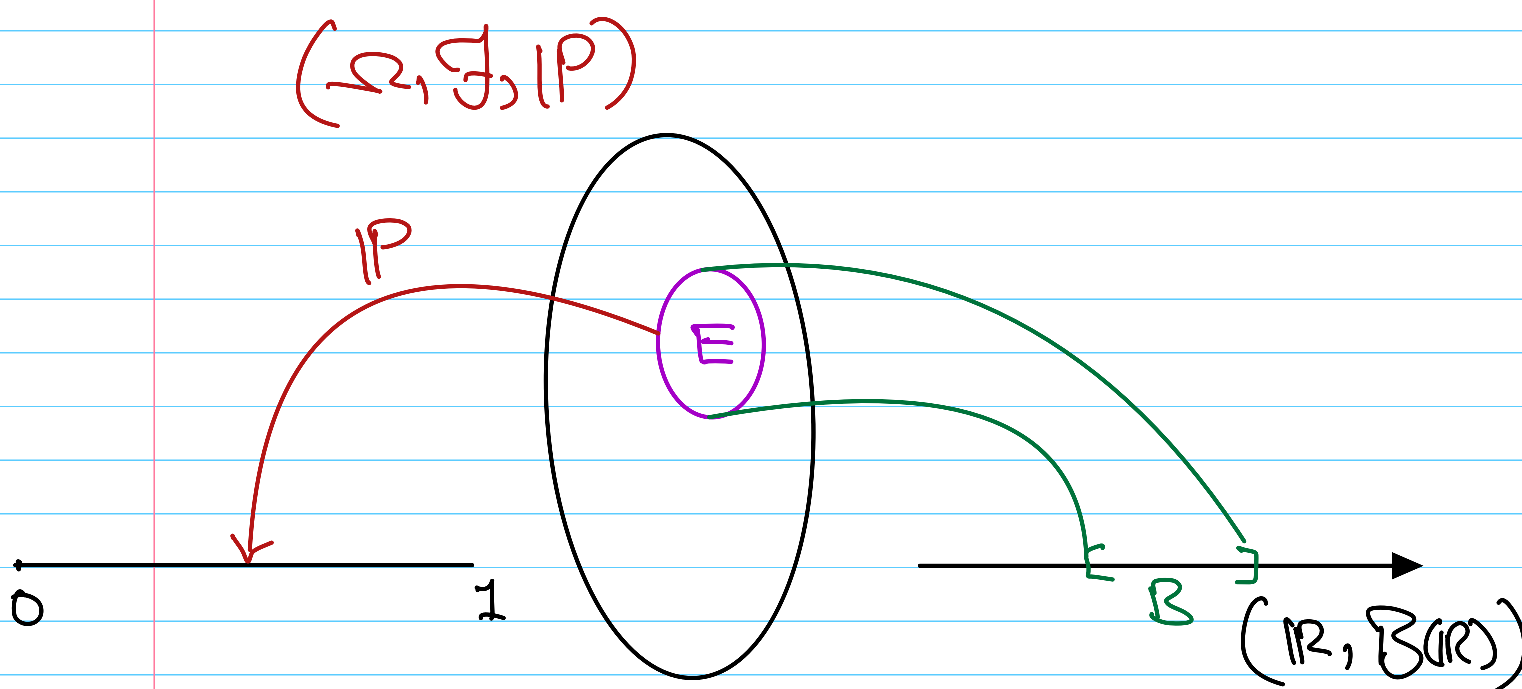
Def: Probability law of random variable X:

The probability law of a r.v X is a function

$$\mathbb{P}_X : \mathcal{B}(\mathbb{R}) \longrightarrow [0,1],$$ which is defined as

$$\mathbb{P}_X(B) \triangleq \mathbb{P}(X^{-1}(B)) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\})$$

original measure on $(\Omega, \mathcal{F})$



two function's are involving here

$$\mathbb{P}_X = \mathbb{P} \circ X^{-1} =$$

$$\mathbb{P}_X(B) = \mathbb{P} \circ X^{-1}(B)$$
$$= \mathbb{P}(X^{-1}(B))$$

$$= \mathbb{P}(\{E\})$$

$\Rightarrow$ let's divide into 2 function's

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

$$\Rightarrow \mathbb{P} \circ X^{-1}(\cdot) : \mathcal{B}(\mathbb{R}) \longrightarrow [0,1]$$

$$\downarrow$$

<u>Step 1</u> $$X^{-1}(\cdot) : \mathcal{B}(\mathbb{R}) \longrightarrow \mathcal{F}$$

<u>Step 2</u> $$\mathbb{P}(\cdot) : \mathcal{F} \longrightarrow [0,1]$$

<u>Example:</u>

$$B \in \mathcal{B}(\mathbb{R}) \quad \& \quad X^{-1}(B) = E \in \mathcal{F}$$

$$\Rightarrow \mathbb{P}_X(B) = \mathbb{P}\left(X^{-1}(B)\right) = \mathbb{P}(E)$$

$$\mathbb{P}_X = \mathbb{P} \circ X^{-1}$$

($\mathbb{P}$ composition X^{-1})

$\mathbb{P}$ = original probability measure on $(\Omega, \mathcal{F})$

$\mathbb{P}_X$ = Induced probability measure on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$

Theorem: $(\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$ is a probability measure.

Proof: (1)

$$\mathbb{P}_X(\phi) = \mathbb{P} \circ X^{-1}(\phi)$$
$$= \mathbb{P}(X^{-1}(\phi))$$
$$= \mathbb{P}(\phi)$$
$$= 0$$

(2)

$$\mathbb{P}_X(\mathbb{R}) = \mathbb{P} \circ X^{-1}(\mathbb{R})$$
$$= \mathbb{P}(X^{-1}(\mathbb{R}))$$

$$= \mathbb{P}(\Omega)$$

$$= 1$$

Q. What does a random variable do?
Why are we defining it?

<u>Ans</u>: it effectively pushes, the measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ onto $(\mathbb{R}, \mathbb{B}(\mathbb{R}))$ and $\mathbb{P}_X$ is the probability measure induced on the Real Line.

* Once we have $(\mathbb{R}, \mathbb{B}(\mathbb{R}), \mathbb{P}_X)$ and we can throw away $(\Omega, \mathcal{F}, \mathbb{P})$ Because we are only interested in X.

<u>Example</u>: if we are interested in Number of Heads in 3 coin tosses

then

$$\Omega = \left\{ \begin{matrix} HHH \\ HHT \\ HTH \\ THH \end{matrix} , \begin{matrix} TTT \\ TTH \\ THT \\ HTT \end{matrix} \right\}, \; |\Omega| = 8$$

$$\mathcal{F} = 2^{\Omega}, \quad |\mathcal{F}| = 2^8$$

$$\mathbb{P}(\{\omega\}) = \frac{1}{8}$$

or.V $\;$ X: Number of Heads in 3 coin tossed.

then the new sample space

$$\Omega_{new} = \{0,1,2,3\}, \quad |\Omega_{new}| = 4$$

$$\mathcal{F}_{new} = 2^{\Omega_{new}}, \quad |\mathcal{F}_{new}| = 2^4$$

$$\mathbb{P}_x(\{0\}) = \frac{1}{8}, \; \mathbb{P}_x(\{1\}) = \frac{3}{8} \; \mathbb{P}_x(\{2\}) = \frac{3}{8}$$

$$\mathbb{P}_x(\{3\}) = \frac{1}{8}$$

Now we can throw $(\Omega, \mathcal{F}, \mathbb{P})$

* So, it appear's as though, in order to give you a complete description of a r.v, we need to have the $\mathbb{P}_X(\cdot)$ for every Borel set. it turn's out that we don't need $\mathbb{P}_X(\cdot)$ for all Borel sets.

* There are multiple way's to generate Borel set's on Real line $\mathbb{R}$.

Ex: ① Using open interval's

② Using open closed interval's

③ Using semi - infinite interval's

$(-\infty, x]$ where $x \in \mathbb{R}$

$$\sigma\left((-\infty, x] : x \in \mathbb{R} \right) = \mathcal{B}(\mathbb{R})$$

* The Probability Law $\mathbb{P}_X$ that defined on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ for all Borel sets must be defined for semi-interval subsets (The generating class)

* Note that: sets of form $(-\infty, x]$ where $x \in \mathbb{R}$ are Borel sets.

$\Rightarrow$ $\mathbb{P}_X(\cdot)$ is defined well on generating class $(-\infty, x]$

$\Rightarrow$ $\mathbb{P}_X((-\infty, x])$ is well defined

$$\Rightarrow \mathbb{P}_X\big((-\infty, x]\big) \triangleq \mathbb{P}\big(\{\omega \in \Omega \mid X(\omega) \le x\}\big)$$

this is called CDF

$$F_X(x) = \mathbb{P}_X\big((-\infty, x)\big]$$

$$= \mathbb{P}\big(\{\omega \in \Omega \mid X(\omega) \le x\}\big)$$

in short hand notation

$\mathbb{P}(\{\omega \in \Omega \mid X(\omega) \leq x\})$ is written

as $\mathbb{P}_X(X \leq x)$. its a absusive notation, but its a short hand.

$$\mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\})$$
$$= \mathbb{P}_X(X \in B)$$

absusive notion but we use this

π- system:

A collection $\mathcal{G}$ of subsets of Ω is said to be a π- system if

$$A, B \in \mathcal{G} \implies A \cap B \in \mathcal{G}$$

only finite intersection, No compliments

No countable union's & intersection.

⇒ this is weaker than algebra also.

one of the most commonly used π-system on $\mathbb{R}$ is the class of all closed semi-infinite intervals

$$\pi(\mathbb{R}) \triangleq \{(-\infty, x] : x \in \mathbb{R}\}$$

The σ-algebra generated by $\pi(\mathbb{R})$ is the Borel σ-algebra $\mathcal{B}(\mathbb{R}) = \sigma(\pi(\mathbb{R}))$

* there is a non-trivial result in measure theory. if we uniquly specify a measure on a π-system, its gets uniquly specified on σ-algebra genered by $\mathcal{G}$.

theorem: Uniqueness of extension, π-system's.

Let $\mathcal{G}$ be a π-system on Ω, and $\mathcal{F} = \sigma(\mathcal{G})$ Let $\mathbb{P}_1$ & $\mathbb{P}_2$ are two Probability measure on $(\Omega, \mathcal{F})$ that 'agree' on $\mathcal{G}$ i.e $\mathbb{P}_1 = \mathbb{P}_2$ on $\mathcal{G}$, then $\mathbb{P}_1 = \mathbb{P}_2$ on $\mathcal{F}$.

__Explanation:__

$$\mathcal{F} = \sigma(\mathcal{G})$$

where $\mathcal{G}$ is a π-system.

* Now let's take two Probability space

$$(\Omega, \mathcal{F}, \mathbb{P}_1) \;\&\; (\Omega, \mathcal{F}, \mathbb{P}_2)$$

* these both agree on $\mathcal{G}$ ie

$$\mathbb{P}_1(\mathcal{G}) = \mathbb{P}_2(\mathcal{G})$$

* Then

$$\mathbb{P}_1(\mathcal{J}) = \mathbb{P}_2(\mathcal{J})$$

* if we uniquely specify measure on π-system then the measure is uniquely specified on the entire σ-algebra.

* why is this useful. because π-system are easy, σ-algebra is difficult.

Now, $\pi(\mathbb{R}) \triangleq \{(-\infty, x] \mid x \in \mathbb{R}\}$

is a π-system

$$\sigma(\pi(\mathbb{R})) = \mathcal{B}(\mathbb{R})$$

* So, if we specify CDF (which is simply specifying the Probability measure on the $\pi(\mathbb{R})$) then it get uniquely specified on entire Borel σ-algebra.

Theorem: The CDF $F_X(\cdot)$ is uniquely specifies the probability law $\mathbb{P}_X$ for any r.v X

$\Rightarrow$ if you tell me CDF of a r.v, we can unique specify probability measure $\mathbb{P}_X$ on entire Borel σ-algebra.

* in general $\mathbb{P}_X$ is the complete discription of r.v X, But in practice we use $F_X(\cdot)$, because it captures all the information $\mathbb{P}_X(\cdot)$.