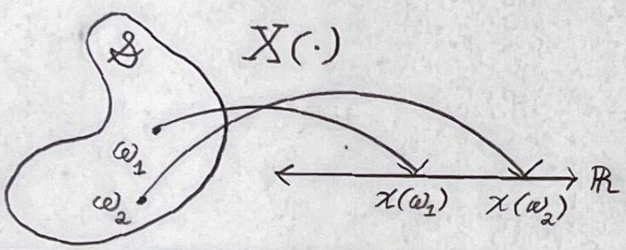
Random Variables

Informal Def: RV X (based on a prob. space $\leftarrow (\mathcal{S}, \mathcal{F}, P)$) is a function mapping $\mathcal{S}$ to the real numbers $\mathbb{R}$, where the set of such functions allowed as RV's is restricted

$$X : \mathcal{S} \mapsto \mathbb{R}$$

Formal Def (Beta version): Given a prob. space $(\mathcal{S}, \mathcal{F}, P)$, a RV X is a function mapping any outcome in a sample space to the corresponding point on the real line



$$X : \mathcal{S} \mapsto \mathbb{R}$$

RV Schedule:

- 1 Random Var
- 2 Random Vars
- n RV's for some finite n
- Random Sequences (inf. seq. of RV's)
- Random Processes (random functions of time)

*Note: Diff btwn X & P (i.e., a RV & a prob. measure)

X : (The RV maps outcomes to the real #'s)

P: (Domain of P is the event space mapping events to the real #'s)

---

* Assign probabilities to the events related to X via a new probability space $\leftarrow (\mathbb{R}, B(\mathbb{R}), P_X)$

$\Rightarrow$ The goal is a new prob. measure that describes X (using the Borel field as the event space since the sample space is the real numbers $\leftarrow B(\mathbb{R})$)

FIND: (?) $P_X(A) = P(X \in A), \forall A \in B(\mathbb{R})$

$\Rightarrow P_X(X \in A) = P(X^{-1}(A)) \leftarrow X^{-1}(A)$ must be in $\mathcal{F}$ (in order to be defined)

*Note: "the event that X is in A"

$\{X \in A\} = \{X(\omega) \in A\} =$

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

the inverse image of A in $\mathcal{S}$ (proper subset)

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

Official Def: Given a prob. space $(\mathcal{S}, \mathcal{F}, P)$, a RV is a function mapping $\mathcal{S}$ to $\mathbb{R}$, satisfying the following property:

$$\boxed{X^{-1}(A) = \{\omega \in \mathcal{S} : X(\omega) \in A\} \in \mathcal{F}, \quad \forall A \in B(\mathbb{R})}$$

$\Leftrightarrow$ **Borel-Measurable Function**

Now we can let:

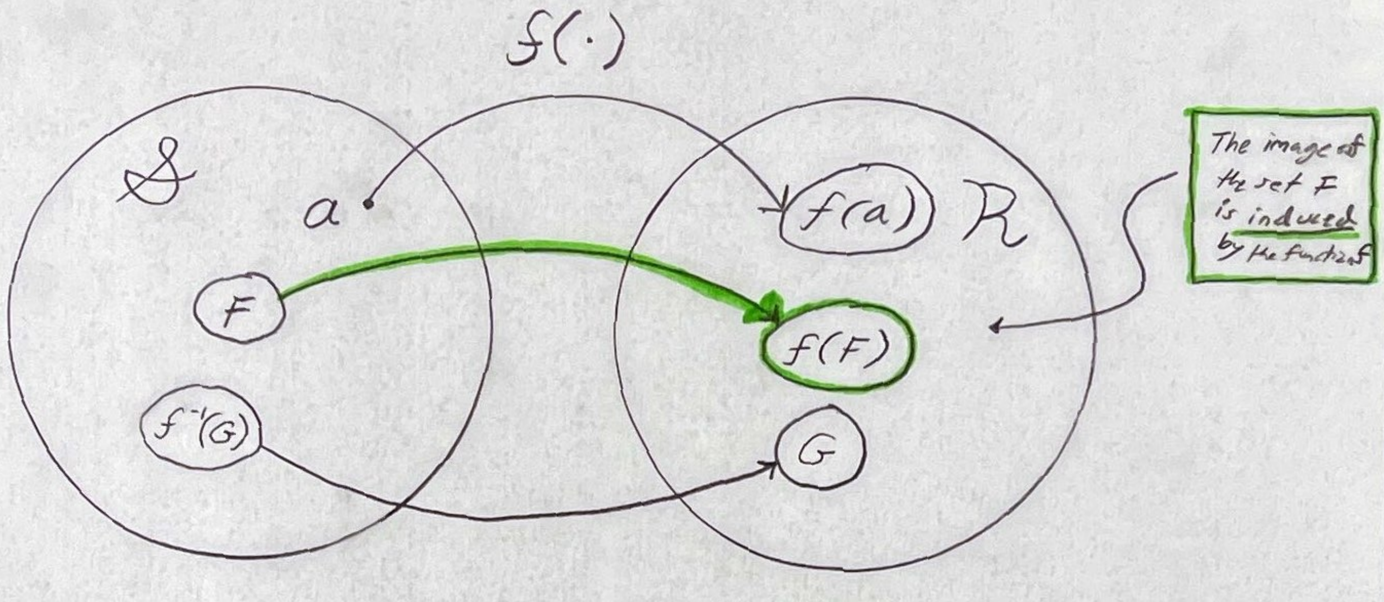
$$\boxed{P_X(A) = P(X^{-1}(A)), \forall A \in B(\mathbb{R})}$$

*note classes of functions are measurable:

1) continuous functions
2) indicator functions
3) Limit of a sequence of measurable functions (i.e., if $X_1, X_2, \ldots$ are measurable, then $\lim_{n \to \infty} X_n$ is measurable)
4) a countable sum of measurable functions

Inverse Images of Sets

(2 spaces, $\mathcal{S}$, & $\mathcal{R}$)



* function $f(\cdot)$ maps $\mathcal{S}$ to $\mathcal{R}$ $\Rightarrow$ $f: \mathcal{S} \mapsto \mathcal{R}$

$\therefore f(a) \in \mathcal{R}, \forall a \in \mathcal{S}$

Definitions:

1) Given a set $F \subset \mathcal{S}$, the <u>image</u> of F under f is:

$$f(F) \equiv \{ b \in \mathcal{R} : b = f(a) \text{ for some } a \in \mathcal{S} \}$$

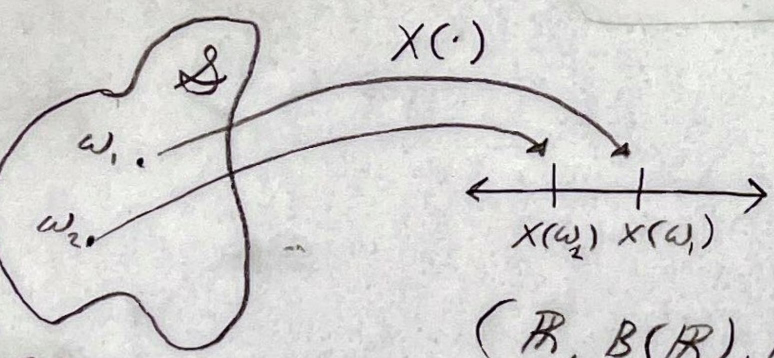
2) Given a set $G \subset \mathcal{R}$, the <u>inverse image</u> of G under f is:

$$f^{-1}(G) \equiv \{ a \in \mathcal{S} : f(a) \in G \}$$

* $f(F)$ is the set of all values/points in $\mathcal{R}$ that are mapped to, from some element in $\mathcal{S}$ by f
* $f^{-1}(G)$ is the set of all points in $\mathcal{S}$ that are mapped to a point in G by the function f
* $f^{-1}(G) \neq f^{-1}(a)$, $a \in \mathcal{R}$ ("inverse img is <u>NOT</u> equal to the inverse function")
* $f(F) \subset \mathcal{R}$, $f^{-1}(G) \subset \mathcal{S}$
* $f^{-1}(a)$ MAY NOT EXIST $\forall a \in \mathcal{R}$ ← (INV FN)
* $f^{-1}(G)$ exists ALWAYS $\forall a \in \mathcal{R}$ ← (INV IMG)

* Note that it is not required that all points in $\mathcal{R}$ be mapped to under f by some element in $\mathcal{S}$
$\Rightarrow$ If so, the inverse image still exists, it is just the empty set $\varnothing$

Prob. Distribution for a RV X



$(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$

$(\mathcal{S}, \mathcal{F}, P)$ $\underbrace{\phantom{(\mathcal{S}, \mathcal{F}, P)}}_{\text{Probability Space}}$

* If the distribution of the RV is known, then you can compute then the probability of the RV for any set of Borel numbers

$$\boxed{P_X(A) = P(X \in A) = P(X^{-1}(A)) \quad \forall A \in \mathcal{B}(\mathbb{R})}$$

" RV X Induces a new sample space $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ "

new prob. measure

In practice, usually not feasible to find P, so we estimate it via modeling but direct modeling is unlikely → use cdf or pdf (density) or pmf (in some cases)

---

cdf for a rv X :

$$
\begin{aligned}
F_X(x) &= P_X((-\infty, x]) \quad \forall x \in \mathbb{R} \\
&= P(X^{-1}((-\infty, x])) \\
&= P(\{\omega \in \mathcal{S} : X(\omega) \le x\})
\end{aligned}
$$

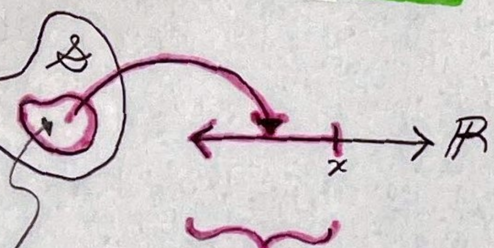
Standard Notation : $F_X(x) = P(X \le x)$

cdf of a rv X : $F_X(x) = P_X((-\infty, x]) \ \forall x \in \mathbb{R}$

$$\boxed{F_X(x) = P(X \leq x), \ \forall x \in \mathbb{R}}$$

$$F_X(x) = P_X((-\infty, x]) \ \forall x \in \mathbb{R}$$
$$= P(X^{-1}((-\infty, x]))$$
$$= P(\{\omega \in \mathcal{S} : X(\omega) \leq x\}), \ \forall x \in \mathbb{R}$$

"big X" (rv to which F's fn belongs)



$\boxed{X^{-1}((-\infty, x])}$ for this set (which is a Borel set)

F_X is the probability of X inverse of interval $(-\infty, x]$

$\Rightarrow$ for each fixed little x, it is the probability of the inverse image

---

Properties Required for a Valid CDF : (make P a valid prob. measure???)

(1) $\lim\limits_{x \to \infty} \left(F_X(x) \right) = 1$ & $\lim\limits_{x \to -\infty} \left(F_X(x) \right) = 0$

(2) For any $x_1, x_2 \in \mathbb{R}$, w/ $x_1 < x_2$, $F_X(x_1) \leq F_X(x_2)$

(F_X must be non-decreasing)

(3) F_X is continuous from the RHS ("continuity from the right")

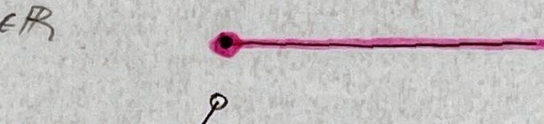
(i.e.,)
$$\underbrace{F_X(x^+) \equiv}_{\text{"}F_X(x^+)\text{ is defined as"}} \lim_{\epsilon \to 0, \ \epsilon > 0} \left(F_X(x+\epsilon) \right) = F_X(x) \ \forall x \in \mathbb{R}$$

(4) $P(X > x) = 1 - F_X(x) \ \forall x \in \mathbb{R}$

(5) If $x_1 < x_2$, then

$P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1)$,

$\forall x_1, x_2 \in \mathbb{R}$



limit from LHS (6) $P(\{X = x\}) = F_X(x) - F_X(x^-)$

where $F_X(x^-) \equiv \lim\limits_{\epsilon \to 0, \epsilon > 0} (F_X(x - \epsilon))$

to be continuous from the RHS when there is a jump discontinuity (Example) requires the value of $F_X(x)$ at that point must be the black filled in dot

Probability Mass Function

[* Recall that a discrete RV X has a cdf that is piecewise constant]

def. PMF (for a discrete RV X)

$$\boxed{p_X(x) \equiv P(X=x),\ \forall x \in \mathcal{R}_X = X(\mathcal{S})}$$

(little p)

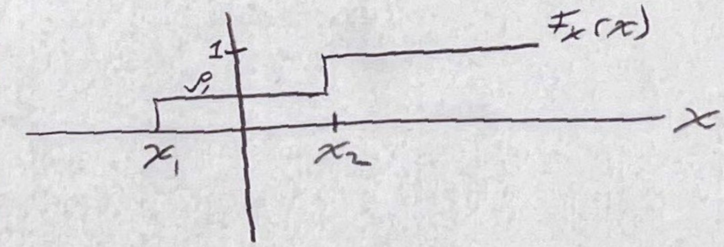
Probability that X equals x for all little x in a set script $\mathcal{R}_X$ equal to the image of the set $\mathcal{S}$ under the random variable/function X

Recall that the image of set $\mathcal{S}$ under the RV function X, $X(\mathcal{S})$ is:

$$X(\mathcal{S}) = \{x \in \mathbb{R} : X(\omega) = x, \text{ for some } \omega \in \mathcal{S}\}$$

(* PMF is only defined on the subset of real number $\mathcal{R}_X$, which is the image of X)

Example:



$$F_X(x) = p_1\, u(x-x_1) + p_2\, u(x-x_2),\ \forall x \in \mathbb{R}$$

* All info in the PDF is carried by the values x_1, x_2, p_1

PDF $\Rightarrow$ $f_X(x) = p_1\,\delta(x-x_1) + p_2\,\delta(x-x_2),\ \forall x \in \mathbb{R}$

(PMF uses the same information just in a diff form)

$\mathcal{R}_X \leftarrow$ Subset of the real #'s on which the pmf is defined

---

The PMF is NOT defined on $\mathbb{R}$, it is defined on a subset, $\mathcal{R}_X$ ($\mathcal{R}_X \subset \mathbb{R}$)

$\therefore$ New probability space

$$(\Omega, \mathcal{B}, \mathbb{P}) \rightarrow (\mathcal{R}_X, \mathcal{P}(\mathcal{R}_X), P_X)$$

(Domain is powerset of $\mathcal{R}_X$)

(prob measure on the space induced by X)

cdf/pdf for discrete RV, prob. space: $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$

pmf for discrete RV, prob. space: $(\mathcal{R}_X, \mathcal{P}(\mathcal{R}_X), P_X)$

In general PDF for a discrete RV is

$$f_X(x) = \sum_i p_i\,\delta(x-x_i),\ \forall x \in \mathbb{R}$$

$\Downarrow$

PMF is then

where $\mathcal{R}_X = \{1, \dots, n\}$ (finite) or $\{1, 2, 3, \dots\}$ (countable)

$$p_X(x_i) = p_i,\ \forall x_i \in \mathcal{R}_X$$

Probability Density Function

*NORMALLY USE PMF FOR DISCRETE RANDOM VARIABLES

For any $A \in B(\mathbb{R})$, $P(X \in A)$ may be written in terms of $F_X(x)$ (cdf) but sometimes this is not natural (or may be very ugly)

... Use pdf instead (a.k.a. the density function)

Prob. Density Function of a RV X is the derivative of the cdf

$$\Rightarrow \quad f_X(x) = \frac{d}{dx} F_X(x), \ \forall x \in \mathbb{R} \qquad \boxed{\text{PDF}}$$

* At points where F_X is not differentiable, use the dirac-delta function to determine $f_X(x)$ (Jump discontinuities

---

def. Dirac Delta Function: (impulse)

i) $\delta(x) = 0, \ \forall x \neq 0$

ii) $\int_{-\infty}^{\infty} \delta(x)\,dx = \int_{-\varepsilon}^{\varepsilon} \delta(x)\,dx = 1, \ \forall \varepsilon > 0$

(requirement to be satisfied to be defined as a DD fn)

PDF Properties of RV X:

cdf is non decreasing, & since density fn is deriv. of cdf, the slope of the cdf will NEVER be negative, thus the density function will never be negative

1) $f_X(x) \geq 0, \ \forall x \in \mathbb{R}$

2) $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

Sufficient to prove a valid pdf

3) $P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f_X(x)\,dx$

*NOTE: $f_X(x)$ is NOT a probability for a fixed x (Unlike the CDF)

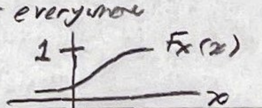
& in general $P(X \in A) = \int_A f_X(x)\,dx, \ \forall A \in B(\mathbb{R})$

* The Fundamental Thm of Calc: $F_X(x) = \int_{-\infty}^{x} f_X(r)\,dr, \ \forall x \in \mathbb{R}$

Every Random Variable has a VALID CDF

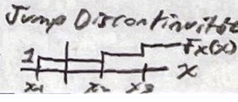
* A RV w/ a CDF that is continuous everywhere

$\Rightarrow$ CONTINUOUS RV Example:



* A RV w/ a CDF that is piecewise constant with a finite or countable # of Jump Discontinuities

$\Rightarrow$ DISCRETE RV Example:



Conditional Distributions for a RV (via cdf/pdf/pmf)

$*$ **Recall:**

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$\forall A \in \mathcal{F}, B \in \mathcal{F}, P(B) > 0$

(1 of 2)

(i) Cond. cdf of A given B: where $A = \{X \le x\}$ $\forall x \in \mathbb{R}$

$*$ Define the cdf of A conditioned on B

$$F_X(x|B) = P(X \le x|B) = \frac{P(\{X \le x\} \cap B)}{P(B)}, \quad \forall B \in \mathcal{F},\ P(B) > 0$$

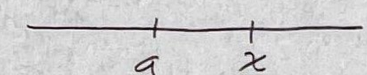
(ii) Cond. pdf of X given B is: $f_X(x|B) = \frac{d}{dx} F_X(x|B)$ (wherever the cdf F_X is differentiable)

Example: let B = the event that x is greater than A for some fixed A in $\mathbb{R}$

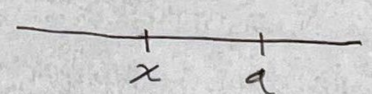
$\Rightarrow B = \{X > a\}$, for some $a \in \mathbb{R}$

or $F_X(x|B)$

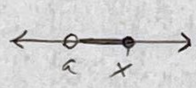
$$\Rightarrow F_X(x|X > a) = \frac{P(\{X \le x\} \cap \{X > a\})}{P(X > a)}$$



2 Cases:
$$\begin{cases} 1)\ x > a: & F_X(x|B) = P(X \le a)/P(X \le a) = 1 \\ 2)\ x \le a: & F_X(x|B) = P(X \le x)/P(X \le a) = \frac{F_X(x)}{F_X(a)} \end{cases}$$



1) $x > a$;



$$F_X(x|B) = P(a < X \le x) / P(X > a)$$
$$= \frac{F_X(x) - F_X(a)}{1 - F_X(a)}$$

2) $x \le a$:

(Empty set)

$$F_X(x|B) = \frac{P(\varnothing)}{1 - F_X(a)} = 0$$

$\therefore$
$$F_X(x|X > a) = \begin{cases} \frac{F_X(x) - F_X(a)}{1 - F_X(a)}, & x > a \\ 0, & x \le a \end{cases}$$

$\therefore$
$$f_X(x|X > a) = \begin{cases} \frac{f_X(x)}{1 - F_X(a)}, & x > a \\ 0, & x \le a \end{cases}$$

"looking for the set of all omega that X maps to some value less than or equal to x but greater than a"

$\Rightarrow$ Empty set (not possible) b/c they are disjoint regions

Bayes Theorem for conditional distributions of one RV:

$$F_X(x|B) = \frac{P(B|X \le x) F_X(x)}{P(B)}$$

$$\& \quad TPL: \quad F_X(x) = \sum_{i=1}^{n} F_X(x|A_i) P(A_i)$$

iff this is true: $A_1, \dots, A_n$ form a partition with $P(A_i) > 0$ $\forall i$

Instead of $P(B|X \le x)$, $\forall x \in \mathbb{R}$ we often want $P(B|X=x)$

$$P(B|X=x) = \frac{P(B \cap \{X=x\})}{P(X=x)}, \Rightarrow \text{RHS is } \frac{0}{0} \text{ for any continuous } X$$

(BAD 8)

Thus, Solution:

$$P(B|X=x) \equiv \lim_{\Delta x \to 0} \left(\frac{P(x < X \le x + \Delta x | B) P(B)}{P(x < X \le x + \Delta x)} \right)$$

$$= P(B) \lim_{\Delta x \to 0} \left(\frac{F_X(x + \Delta x | B) - F_X(x|B)}{F_X(x + \Delta x) - F_X(x)} \right)$$

(written in terms of cdfs)

$$= P(B) \frac{\lim_{\Delta x \to 0} \left(\frac{F_X(x + \Delta x | B) - F_X(x|B)}{\Delta x} \right)}{\lim_{\Delta x \to 0} \left(\frac{F_X(x + \Delta x) - F_X(x)}{\Delta x} \right)}$$

$$= P(B) \frac{f_X(x|B)}{f_X(x)} \Longrightarrow \boxed{P(B|X=x) = \frac{f_X(x|B) P(B)}{f_X(x)}}$$

Bayes Thm for $P(B|X=x)$

* If X is discrete RV & $B \in \mathcal{F}$ with $P(B) > 0$, use the conditional prob. mass fn (pmf) of X given B

$$\Rightarrow P_X(x|B) = \frac{P(B|X=x) P_X(x)}{P(B)}$$

$\forall x \in \mathbb{R}_X$, with $P(X=x) > 0$