classmate
Date
Page

Probability and Statistics

Random Experiments and Sample Spaces :-

Experiments involving Randomness

Coin toss, dice roll, etc.

$\Omega \rightarrow$ Set of all possible outcomes of a random experiment. It could be finite or infinite.

$\Omega_c = \{Head, Tail\}$, $\Omega_d = \{1,2,3,4,5,6\}$

2-coins $\rightarrow \Omega_{2c} = \Omega_c \times \Omega_c = \{H,T\} \times \{H,T\}$

Outcomes and Events

(omega) $\rightarrow \omega \in \Omega$ is called a sample point or a possible outcome

A subset $A \subseteq \Omega$ is called an event.

Events in the coin toss experiment $C_1 = \{T\}$ $[\subset \Omega_c]$

Events in the die roll $D_1 = \{6\}$, $D_2 = \{1,3,5\}$ $[\subseteq \Omega_d]$

Events are any subset of Ω, even null sets, but $\mathbb{P}(\phi) = 0$

Probability of an event $A = \mathbb{P}(A)$.

It may or may not be possible to measure / assign $\mathbb{P}$ for every subset A.

$\mathbb{P}_c(T) = 0.5$

Probability measure $\mathbb{P}$ is a set function. It acts on sets and measures the probability of such sets

classmate
Date
Page

Set Theory 101:-

A^c = complement of A

$\emptyset$ = denotes empty set. [belongs to every set]

$A \cup B$ = A union B

$A \cap B$ = A intersection B

~~A∆B~~

$A \setminus B$ = A minus B $= A \cap B^c$

~~Symmetric~~

$A \Delta B = (A \setminus B) \cup (B \setminus A)$

ME = Mutually Exclusive

~~HW → Identity Complement,~~

$|A|$ = no. of elements in A = cardinality of A.

Inclusion - Exclusion Principle = $|A \cup B| = |A| + |B| - |A \cap B|$

Countable Sets and Uncountable Sets

Monotone Seq → $A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$ [Increasing seq]

$A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$ [Decreasing]

$I_n = [0, 1 - \frac{1}{n}]$

$D_n = [0, \frac{1}{n}]$

Cartesian product :- $A \times B = \{(a, b) : a \in A, b \in B\}$

Powerset [$\mathcal{P}(A)$] :- Set of all possible subsets of A.

$|\mathcal{P}(A)| = 2^{|A|}$ [Only for discrete elements]

$\mathcal{P}([0,1]) = \{(a,b) : a \leq b, a, b \in [0,1]\}$

↳ range, not just 2 elements

Functions:-

Fns are maps from elements in the domain D to the range R

$f: D \to R$

$f: \mathbb{R} \to \mathbb{R} \Rightarrow f(x) = x$

Set fns are those fns ~~who~~ who act on sets; D is a collect of sets.

$\mathbb{P}$ is a set fn.

Axioms of probability:-

1) $\mathbb{P}(\emptyset) = 0$, $\mathbb{P}(\Omega) = 1$
2) For a set $A \subseteq \Omega$, $\mathbb{P}(A) \in [0,1]$
3) For a disjoint collection of events $A_1, A_2, \cdots A_j$ where $A_i \subseteq \Omega$, then:

$$\mathbb{P}\left(\bigcup_{i=1}^{j} A_i\right) = \sum_{i=1}^{j} \mathbb{P}(A_j)$$ [j can be ∞] → but must be countable

where $A_i \cap A_j \ \forall i,j = \emptyset$

In general, the domain of $\mathbb{P}$ is $\mathcal{P}(\Omega)$.

~~Probability of impossible event = 0~~.

Counter example for $\mathcal{P}(\Omega)$ being satisfactory

1) Pick a number randomly (uniformly) of from the real line.
2) $\Omega = \mathbb{R}$, hence $\mathbb{P}(\mathbb{R}) = 1$.
3) Domain $= \mathcal{P}(\mathbb{R})$
4) $\mathbb{P} : \mathcal{P}(\mathbb{R}) \to [0,1]$
5) $\mathbb{P}$ has the property that sets of equal length have equal probability
6) We know that $\mathbb{R} = \bigcup_{n=-\infty}^{\infty} [n, n+1)$ where $[n, n+1) \in \mathcal{P}(\mathbb{R})$
[Countable union since n is an integer]
7) $\mathbb{P}(\mathbb{R}) = 1 = \sum \mathbb{P}([n, n+1)) = 0 / \infty$ depending
8) This is a contradiction! [Happens when picking segments of unit length]

Powerset is thus a bad choice

Not all set fns can be calibrated to measure all possible subset of the sample space.

Our choice of domain must have certain features.
Let domain = $\mathcal{F}$
then $\phi, \Omega \in \mathcal{F}$; if $B \in \mathcal{F} \to B^c \in \mathcal{F}$ etc;

$\mathcal{F}$
$\uparrow$
Bags that meet these criteria are called sigma-algebras

Sigma algebra / Event space $[\mathcal{F}]$ associated with a set Ω is a collection of subsets of Ω that satisfy
- $\to \emptyset \in \mathcal{F}$ and $\Omega \in \mathcal{F}$
- $\to A \in \mathcal{F} \to A^c \in \mathcal{F}$
- $\to A_1, A_2, \dots, A_n \in \mathcal{F} \to \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

$\to$ The σ-algebra is closed under formation of complements and countable unions and countable intersections (by De Morgans)

$\to$ When Ω is countable and finite, then $\mathcal{P}(\Omega)$ is a σ-algebra

If Ω is countable & finite, we will consider $\mathcal{P}(\Omega)$ as the domain.

Probability Space = $\{\Omega, \mathcal{F}, \mathbb{P}\}$

Probability measure :- A probability measure $\mathbb{P}$ on the measurable space $(\Omega, \mathcal{F})$ is a set fn $\mathbb{P} : \mathcal{F} \to [0,1]$ s.t.
1) $\mathbb{P}(\phi) = 0$, $\mathbb{P}(\Omega) = 1$
2) For a disjoint collection of events $A_1, A_2 \dots$ from $\mathcal{F}$, we have

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

$\mathbb{P}$ space for $U[0,1]$

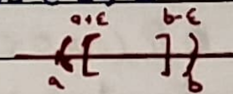
$\Omega = [0,1]$

If we generate a σ-algebra, bet we get a borel-sigma algebra $\mathcal{B}[0,1]$

Borel σ-algebra $[\mathcal{B}[0,1]]$

Defined when $\Omega = [0,1]$, as the σ-algebra generated by closed sets of the form $[a,b]$ where $a \le b$ & $a,b \in [0,1]$

$$(a,b) = \bigcup_{n=1}^{\infty} \left[a + \frac{1}{n},\ b - \frac{1}{n}\right]$$



$$(a,b] = \bigcup_{n=1}^{\infty} \left(a,\ b + \frac{1}{n}\right]$$

$[\mathcal{B}[0,1]]$

Borel σ-algebra ↑: is the σ algebra generated by sets of the form $[a,b]$ or (a,b) or $(a,b]$ or even $[a,b)$ where $a \le b$ & $a,b \in [0,1]$

If $\Omega = \mathbb{R}$, the $\mathcal{B}(\mathbb{R})$ is the σ-algebra generated by open sets of the form (a,b), where $a<b$ and $a,b \in \mathbb{R}$.

How to define $\mathcal{B}(\mathbb{R}^2)$

Consequences of the Probability Axioms:-

i) $P(A^c) = 1 - P(A)$

$$P(A \cup A^c) = P(A) + P(A^c) = 1.$$

ii) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

~~P(A∪B)~~ ~~Prove~~ Prove it

iii) If $A \subseteq B$, prove that $P(A) \le P(B)$ [$A \subseteq B$ can be interpreted as $A \rightarrow B$]

If $A \subseteq B \rightarrow B = A + \lambda$.

iv) $P\left(\bigcup_{i=1}^{\infty} B_i\right) \le \sum_{i=1}^{\infty} P(B_i)$ [Boole's ineq.]

Diff b/w Impossible event vs Zero probability event

ϕ ← Impossible event

Zero probability event → finite sized set

In $U[0,1]$, $\mathbb{P}(\omega = 0.5) = 0$

Every experiment outcome of this experiment is a zero probability event.

This implies that zero outcome events could occur.

ϕ on the other hand, can never occur, and are hence impossible events.

$\mathbb{P}(\omega \in [0, 0.25) \cap [0.75, 1]) = 0$

Limits and Continuity:-

Limits:-

Let $a_1, a_2, \dots, a_n$ be a sequence with limit L.

Then $\forall \varepsilon$, $\exists N_\varepsilon$ s.t $\forall n > N_\varepsilon$, $|a_n - L| \le \varepsilon$

For a fn $f(x)$; $\lim_{x \to c} f(x) = f(c)$

The limit exists only if $\lim_{x \to c^+} f(x) = \lim_{x \to c^-} f(x)$.

All points δ close to c are ε close to $f(c)$.

Continuity:-

LHL = RHL = $f(x)$. If the fn is cont at all pts, it is said to be continuous. As $x \to c$, $f(x) \to f(c)$

But what about $\mathbb{P}$?

For a continuous set fn. S as $A_n \to A$, we have $S(A_n) \to S(A)$

Sequence of sets :-

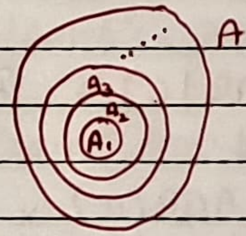
Given $(\Omega, \mathcal{F})$, if $A_1 \subset A_2 \subset \cdots$ is an increasing sequence of events defined on $\mathcal{F}$ and $\bigcup_{n=1}^{\infty} A_n = A \in \mathcal{F}$, then we say that the given seq of sets A_n are increasing to A ($A_n \uparrow A$)

Similarly when $A_1 \supset A_2 \supset \cdots$ is a decreasing seq of events and $\bigcap_{n=1}^{\infty} A_n = A$, then we have $A_n \downarrow A$

Alt notation:- increasing seq of sets $A_n \Rightarrow \lim_{n\to\infty} A_n$ for $\bigcup_{n=1}^{\infty} A_n$

decreasing seq of sets $A_n \Rightarrow \lim_{n\to\infty} A_n = \bigcap_{n=1}^{\infty} A_n$

Lemma:- For a sequence of events of the type $A_n \uparrow A$ or $A_n \downarrow A$, we have $\lim_{n\to\infty} P(A_n) = P(A)$



Proof:-

$$A = \bigcup_{i=1}^{\infty} A_i$$

(i) Consider an increasing seq:

define $F_n = A_n - A_{n-1}$

$$\bigcup_{n=1}^{\infty} A_n = \bigcup_{n=1}^{\infty} F_n = \sum_{n=1}^{\infty} F_n = \lim_{n\to\infty} \sum_{i=1} P(F_i)$$

$$= \lim_{n\to\infty} P\left(\bigcup_{i=1} F_i\right) = \lim_{n\to\infty} P(A_n)$$

Hence Proved.

(ii) Decreasing seq. [HW]

Conditional Probability :-

Ex:- Given the die outcome is odd, what is $P(1)$? [$1/3$]

Given $\bar{\omega} \in [0, 0.5]$, what is $P(\bar{\omega} \in [0, 0.25])$

Defn:- The conditional probability of event A is defined as

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \quad \text{when } \mathbb{P}(A) > 0$$

Theorem:- $\mathbb{P}(A|B)\,\mathbb{P}(B) = \mathbb{P}(B|A)\,\mathbb{P}(A)$

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

Bayes' Rule:- $$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B) \times \mathbb{P}(B)}{\mathbb{P}(A)}$$

(i) What is $\mathbb{P}(A|B \cap C)$

(ii) $\mathbb{P}(A \cap B) = \mathbb{P}(A|B) \times \mathbb{P}(B)$

(iii) $\mathbb{P}(A \cap B|C) = \mathbb{P}(A|B \cap C) \times \mathbb{P}(B|C)$

$$\downarrow$$

$$\frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(\cancel{C})} \times \frac{\mathbb{P}(\cancel{C})}{\mathbb{P}(B \cap C)} \times \frac{\mathbb{P}(B \cap C)}{\mathbb{P}(C)}$$

$$\downarrow$$

$$\mathbb{P}(A|B \cap C) \times \mathbb{P}(B|C)$$

(iv) $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \times \mathbb{P}(B|A) \times \mathbb{P}(C|A \cap B)$

$$\downarrow$$

$$\frac{\mathbb{P}(A \cap B \cap C)}{\mathbb{P}(A \cap B)} \times \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} \times \mathbb{P}(A)$$

(v) $\mathbb{P}(A_1 \cap A_2 \cap \cdots \cap A_n) = \mathbb{P}(A_1) \times \cdots \times \mathbb{P}(A_n | A_1 \cap \cdots \cap A_n)$

classmate

Date
Page

Q) Draw 4 cards without replacement. What is the $\mathbb{P}$(9 club, 8 dia, K sp, K cl)

$$\mathbb{P}(9_c) \times \mathbb{P}(8d \mid 9c) \times \mathbb{P}(K_{sp} \mid 9_c \cap 8d) \times \mathbb{P}(K_{cl} \mid 9_c \cap 8d \cap K_{sp})$$

$$= \frac{1}{52} \times \frac{1}{51} \times \frac{1}{50} \times \frac{1}{49} = \frac{48!}{52!}$$

When every outcome is equally likely in a finite sample space Ω, $\mathbb{P}(B \mid A) = \frac{|A \cap B|}{|A|}$

Law of Total Probability:-

$A = (A \cap B) \cup (A \cap B^c)$, $\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$

Same as $\mathbb{P}(A) = \mathbb{P}(A \mid B)\mathbb{P}(B) + \mathbb{P}(A \mid B^c)\mathbb{P}(B^c)$

Let $B_1, B_2, \dots, B_n$ be the partition of the sample space Ω. Then for any event A we have

$$\mathbb{P}(A) = \sum_{i=1}^{n} \mathbb{P}(A \cap B_i) = \sum_{i=1}^{n} \mathbb{P}(A \mid B_i)\mathbb{P}(B_i)$$

Q) 3 bags with m marbles total. Bag i has R_i red and B_i blue for $i \in [1,3]$. Find $\mathbb{P}$(marble = red from rand bag)

$$\frac{1}{3} \times \sum_{i=1}^{3} \frac{R_i}{R_i + B_i} = \frac{1}{3} \times \sum_{i=1}^{3} \mathbb{P}(\text{Red} \mid B_i) \times \mathbb{P}(B_i)$$

classmate

Date

Page

Independence:-

An event A is independent (w.r.t an event B) if
$\mathbb{P}(A|B) = \mathbb{P}(A)$

Ex:- Simultaneously tossing a coin and rolling a die

$\Omega_1 = \{H, T\}$ $\Omega_2 = \{1,2,3,4,5,6\}$
$\Omega = \Omega_1 \times \Omega_2$
$\mathcal{F} =$ Powerset (Ω)

$\mathbb{P}(\{H,6\}) = \frac{1}{2} \times \frac{1}{6} = \frac{1}{12}$

$\mathbb{P}(\{T, odd\}) = \frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$

$\mathbb{P}(A \cap B) = \mathbb{P}(A)\,\mathbb{P}(B) \rightleftharpoons$ Independence

If A and B are independent

$\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$; $\mathbb{P}(B^c) = 1 - \mathbb{P}(B)$
$\mathbb{P}(A^c) \times \mathbb{P}(B^c) = 1 - \mathbb{P}(A) - \mathbb{P}(B) + \mathbb{P}(A \cap B) = 1 - \mathbb{P}(A \cup B)$
Where to now? $= \mathbb{P}(A \cup B)^c$
$= \mathbb{P}(A^c \cap B^c)$

A^c & B^c are independent

Prove $\mathbb{P}\left(\bigcup_{i=1}^{n} A_i\right) = 1 - \prod_{i=1}^{n} (1 - \mathbb{P}(A_i))$

Mutual and Pairwise independence

A collection of events $\{A_i, i \in I\}$ are said to be mutually independent if the $\mathbb{P}(\bigcap_{j=J} A_j) = \prod_{j \in J} \mathbb{P}(A_j)$ for any subset $J \subseteq I$

A collection of events $\{A_i, i \in I\}$ are said to be pairwise independent if any pair of events from the collection are independent $\Rightarrow$ all pairs.

Ex:- Pick a random number from $\{1, \dots, 10\}$

$\mathbb{P}(A) = \{w \mid w < 7; w \in \{1, 10\}\} = 3/5$

$\mathbb{P}(B) = \{w \mid w < 8; w \in \{1 \dots 10\}\} = 7/10$

$\mathbb{P}(C) = \{w \mid w \% 2 = 0; w \in \{1, \dots 10\}\} = 1/2$

$\mathbb{P}(A \cap B) = \mathbb{P}(A) \rightarrow$ not independent

$\mathbb{P}(A \cap C) = \mathbb{P}(A) \times \mathbb{P}(C) \rightarrow$ independent

$\mathbb{P}(B \cap C) = \frac{3}{10} \neq \mathbb{P}(B) \times \mathbb{P}(C) \rightarrow$ not independent

Correlation b/w events:-

Two events A & B are +vely correlated iff $\mathbb{P}(A \mid B) > \mathbb{P}(A)$

Two events A & B are -vely ~~not~~ correlated iff $\mathbb{P}(A \mid B) < \mathbb{P}(A)$

A & B have same correlation as A^c & B^c (Prove)

A & B have opp correlation as A & B^c (Prove)

Mutual Exclusivity:-

Two events are mutually exclusive if $\mathbb{P}(A \cap B) = 0$

The occurence of one implies the other cannot occur.

Two events with non zero probabilities cannot be both mutually exclusive and independent.

$P(A \cap B) = 0 = P(A)P(B) \rightarrow$ contradiction.

If A & B are ME

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

When $A \subseteq B$, A & B are not ME nor I

Zero Probability Events are always independent

Let $\mathbb{P}(E) = 0$

· Then for any set F, we want to show that $\mathbb{P}(E \cap F) = 0$

But $E \cap F$ has two choices

(a) $E \cap F = \emptyset$

(b) $E \cap F \subseteq E$

In either case $\mathbb{P}(E \cap F) = 0$ or $\mathbb{P}(E \cap F) \leq \mathbb{P}(E) = 0$

$\mathbb{P}(E \cap F) = 0$

Conditional Independence:-

Two 0 events A & B are conditionally independent of C with $\mathbb{P}(C) > 0$ if $\mathbb{P}((A \cap B)/C) = \mathbb{P}(A|C) \cdot \mathbb{P}(B|C)$

It thus follows that $\mathbb{P}(A|B \cap C) = \mathbb{P}(A|C)$

Ex:- Given a fair coin and a fake coin (2 heads). Pick a random coin and toss twice.

A:- First toss $\rightarrow$ H $= 3/4$

B:- Second toss $\rightarrow$ H $=$ Still $3/4$

~~Event~~ C :- First coin is chosen = 1/2

$\mathbb{P}(A \cap B) = \mathbb{P}(A \cap B \mid C) + \mathbb{P}(A \cap B \mid C^c)$

$= \underbrace{\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}}_{\text{Fair coin}} \qquad \underbrace{\frac{1}{2}}_{\text{Unfair coin}}$

$= \frac{5}{8} \neq \frac{3}{4} \times \frac{3}{4}$ → Not independent

C : coin is chosen and tossed twice; not one coin being chosen and tossed twice

A & B are conditionally independent on C, not C^c

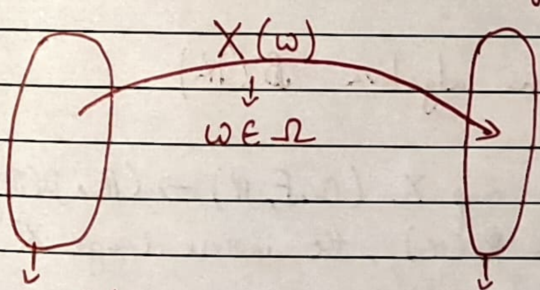
Random Variables

Given a random experiment with associated $(\Omega, \mathcal{F}, \mathbb{P})$, it is sometimes difficult to deal directly with $\omega \in \Omega$

ω is not a number, it is a sequence (vector) of nos.

Random variables are devices that help us make a mapping from $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\Omega', \mathcal{F}', \mathbb{P}_X)$, as this would help analyze functions of sample pts rather than any single sample pt.

Assume you want to count no. of 6s in 10 rolls.



Ω = set of all 10 vectors

$\mathbb{P}$

$\Omega' = \{1, \dots, 10\}$; $\mathcal{F}'$ is the Powerset (Ω')

$\mathbb{P}_X$

A random variable X is a fn $X: \Omega \to \Omega'$ that transforms the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ to $(\Omega', \mathcal{F}', \mathbb{P}_X)$ and is $(\mathcal{F}, \mathcal{F}')$-measurable

The map $X: \Omega \to \Omega'$ implies $X(\omega) \in \Omega'$ $\forall$ $\omega \in \Omega$

For the event $B \in \mathcal{F}'$, the preimage $X^{-1}(B)$ is defined as $X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\}$

$\rightarrow$ for every $B \in \mathcal{F}'$, we have $X^{-1}(B) \in \mathcal{F}$

X^{-1} is a one to many mapping from $\mathcal{F}' \to \mathcal{F}$

$\mathbb{P}_X(\{B\})$ = Probability that X results in $\{B\}$ while ~~being dacted~~ acting upon $\omega \in \Omega$

= ~~Probability of~~ $\mathbb{P}(\omega)$ where $\omega \in \Omega$; $X(\omega) \in B$

$\mathcal{F}$ could have elements that have no images in $\mathcal{F}'$. But $\forall f \in \mathcal{F}'$, f has a pre image in $\mathcal{F}$

$\mathbb{P}_X$ is called induced probability measure

Conventions:-

i) $\Omega' = \mathbb{R}$

ii) $\mathcal{F}'$ = Borel sigma algebra $\mathcal{B}(\mathbb{R})$

A rand var X is a map $X: (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}), \mathbb{P}_X)$ such that for each $B \in \mathcal{B}(\mathbb{R})$, the inverse image $X^{-1}(B) = \{\omega \in \Omega ; X(\omega) \in B\}$ satisfies

$\therefore X^{-1}(B) \in \mathcal{F}$

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

If Ω' is countable, then the random variable is called a discrete random variable. Only here is $\mathcal{F}'$ used as the powerset

Capital letters X Y Z etc for Rand var, and x, y, z etc are the realizations of said rand vars

Discrete Random Variables :-

Ex:- Roll two dice and calculate the sum

Ω = set of all tuples (x,y) where $x,y \in [1,6]$
Ω' = integers $[2,12]$

$\mathcal{F}$ & $\mathcal{F}'$ are Powersets

$\{X=3\} \in \mathcal{F}' \rightarrow \mathbb{P}_x(3) = \mathbb{P}(\{1,2\},\{2,1\})$

In general for $x \in \Omega'$ $\mathbb{P}_x(x) = \mathbb{P}(\{a,b\} \mid \{a,b\} \in \Omega, a+b=x)$

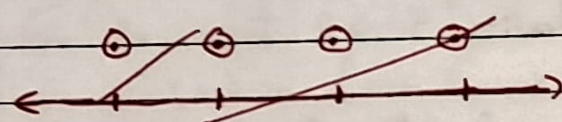
$p_X(x) = \mathbb{P}_X(\{x\})$
→ fn [not a probability measure] → makes x a set and passes as arg to $\mathbb{P}_X$.
→ Probability mass fn.

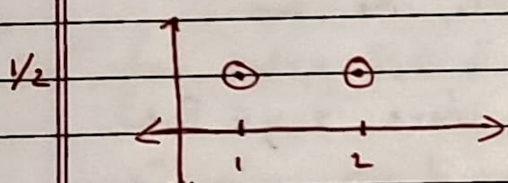
$$p_X(x) = \begin{cases} x-1/36 & \forall x \in \text{int}[2,7] \\ 13-x/36 & \forall x \in \text{int}[8,12] \end{cases}$$

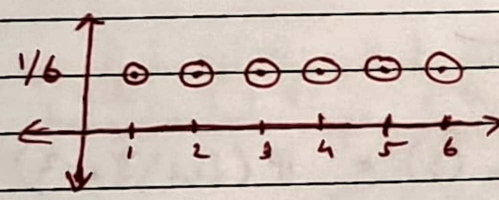
Multiple X_is can have the same Probability Mass fn.
For example, for the exp of rolling a die four times, then divide and conquer using $p_X(x)$ mentioned above
Only if X_is are independent

The fn $p_X(x) := \mathbb{P}_X(\{x\})$ for $x \in \Omega'$ is called as a Probability Mass Function (PMF) of a random variable X.

Q) Plot PMFs for X corresponding to coin tosses or dice rolls.



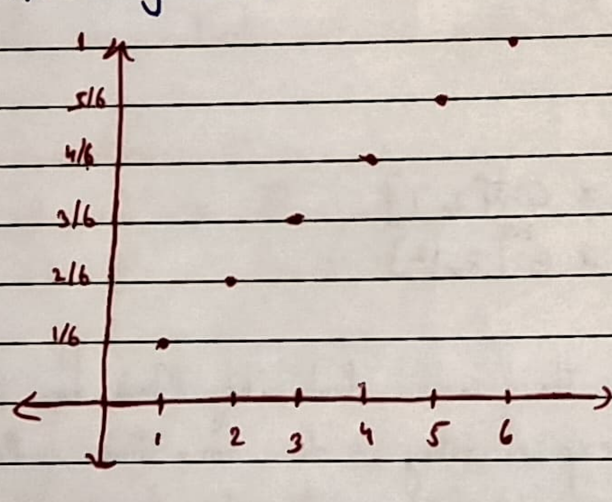




The Cumulative Distribution function (CDF) is defined as

$$F_X(x_1) = \sum_{x \le x_1} p_X(x) = \mathbb{P}\{\{\omega \in \Omega : X(\omega) \le x_1\}\}$$

F_X (Rolling dice)



The X axis is only X not Ω. Here ~~the mapping~~ X is identical to Ω. Ω'

Expectation and Moments:-

The mean or expectation of a random variable X is denoted by $E(x)$ and is given by $E(x) = \sum_{x \in \Omega'} x \, p_x(x)$

The expected value need not be part of Ω

Depends on how X is defined.

First moment of a rand var.

The n^{th} moment of a random variable X is denoted by $E(X^n)$ and is given by $E(x^n) = \sum_{x \in \Omega'} x^n p_x(x)$

For a fn $g(\cdot)$ of a rand var X, $E(g(X)) = \sum_{x \in \Omega'} g(x) \, p_x(x)$

fns of random variables are random variables with their own probability mass fns.

Consistency of PMF:-

PMF: $p_x(x) = \mathbb{P}(\{\omega \in \Omega : X(\omega) = x\})$ for $x \in \Omega'$

To prove $\sum_{x \in \Omega'} p_x(x) = 1$

$$\sum_{x \in \Omega'} p_x(x) = \sum_{x \in \Omega'} \mathbb{P}(\{\omega \in \Omega : X(\omega) = x\})$$

$$= \mathbb{P}\Big(\bigcup_{x \in \Omega'} \{\omega \in \Omega : X(\omega) = x\}\Big)$$

$$= \mathbb{P}(\Omega) = 1$$

Expectance is Linear :-

Let $Y = aX + b$

Then $E(Y) = E(aX+b)$

$$= \sum_{x \in \Omega'} (ax+b)\, p_x(x)$$

$$= a\left(\sum_{x \in \Omega'} x\, p_x(x)\right) + b$$

$$= a\,E(X) + b$$

The fn is OTO, then you could still ~~aptly~~ apply this

$y_i = ax_i + b$

$\mathbb{P}(y_i) = \mathbb{P}(x_i) \rightarrow PMF(y_i) = PMF(x_i)$

$$E[Y] = \cancel{\sum g\, p_y(y)} \sum_{y \in \Omega''} y_i\, p_Y(y)$$

$$= \sum_{x \in \Omega'} (ax_i+b)\, p_Y(ax_i+b)$$

$$= \sum_{x \in \Omega'} (ax_i+b)\, p_Y(x) = \sum_{x \in \Omega'} (ax_i+b)\, p_X(x)$$

$$= a\,E[X] + b$$

What ~~$y = g($~~ is $Y = g(X)$ is MTO?

Ex:- $Y = |X|$

Function of Random Variables :-

Consider $Y = |X|$, $X = (\text{int})[-4, 4]$

$p_X = \frac{1}{9} \quad \forall x \in X = \{(\text{int})[-4,4]\}$

$$p_Y(2) = \sum_{x:|x|=2} p_X(x) = p_X(-2) + p_X(2) = \frac{2}{9}$$

Prove:- If $Y = g(X)$ and PMF of $X = p_X(x)$, then
$$p_Y(y) = \sum_{\{x \mid g(x)=y\}} p_X(x)$$

Theorem:- Suppose $Y = g(X)$ and X is discrete with PMF $p_X(x)$. Then $E[Y] = \sum_x g(x)\, p_X(x)$

Proof:- $$E[Y] = \sum_y y\, p_Y(y)$$

$$= \sum_y \sum_{x: g(x)=y} g(x)\, p_X(x)$$

This is nothing but the summation over all x in domain

$$= \sum_x g(x)\, p_X(x)$$

Variance:-

Consider $g(X) = (X - E[X])^2$

$g(X)$ quantifies the deviation of X from the mean.

This is called the variance.

$\sqrt{E(g(X))}$ quantifies the deviation

$$Var(X) = E[(X - E[X])^2] = E[x^2] - E[x]^2$$

$$E[x^2] + E[x^2 - 2x\,E[x] + E[x]^2] = E[x^2] + E[-2x\,E[x]] + E[E[x]^2]$$
$$= E[x^2] - 2\,E[x]^2 + E[x]^2 = E[x^2] - E[x]^2$$

$\sigma_X = \sqrt{Var(X)}$ is called the Standard Deviation of X

When $Y = aX+b$, $Var(Y) = a^2[Var(X)]$

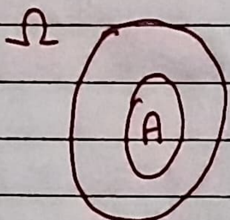
no (+b) cause variance is not affected by const.

a^2 cause variance squares

substitute and verify.

Examples of DRV

(i) Indicator Random Variables

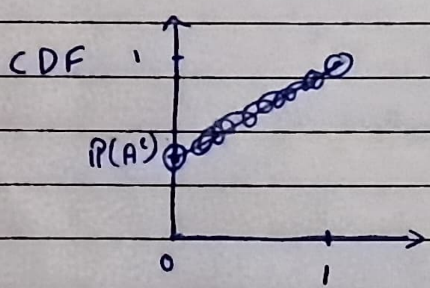


We want to know if A has occured.

$I_A(\omega) = \begin{cases} 1, & \text{if } \omega \in A \subseteq \Omega \\ 0, & \text{else} \end{cases}$

$p_{I_A}(1) = \mathbb{P}(A)$

$p_{I_A}(0) = \mathbb{P}(A^c)$



Mean = $\mathbb{P}(A)$ (?)

(ii) Bernoulli Random Variable

$$X = \begin{cases} 1, & \text{with } P = p \\ 0, & \text{else} \end{cases}$$

~~Used~~ Benoulli Bandits

Mean $= p$, $E[X^n] = p$

(iii) Binomial Random Variables $B(n,p)$

Treat as n Bernoulli Random Variable

Consider a biased coin ($P(\text{head}) = p$) and toss it n times.

Denote head $= 1$, tail $= 0$

Let N denote no of heads in n tosses

$$P_N(k) = \underbrace{{}^nC_k}_{\text{no of ways to choose } k \text{ coin}} \cdot \underbrace{p^k}_{\text{of } k \text{ heads}} \cdot \underbrace{(1-p)^{n-k}}_{\text{remaining } n-k \text{ tails}}$$

~~Mean $= \sum k p_k$~~ $E[N]$, $E[N^2]$, $Var(X)$

(iv) Geometric Random Variable

Consider a biased coin ($P(H) = p$) and keep tossing till you achieve the first head.

Random Variable N = no. of tosses reqd for 1st head

$$P_N(k) = (1-p)^{k-1} \times p$$

$\bar{F}_N(k) = 1 - F_N(k) = P(N > k)$

$P(N > k)$ = k tails = $(1-p)^k$

If head doesn't occur in k tosses, the probability that we need m more tosses ~~of one~~ is independent of k. This is the memoryless property of Geometric Random Variables.

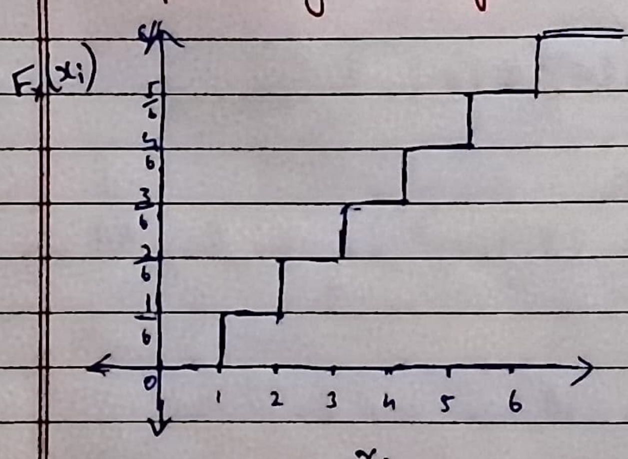
Continuous Random Variables

Consider a random variable X that maps (Ω, F, P) to $(\mathbb{R}, \mathcal{B}(\mathbb{R}), P_X)$ where $\mathcal{B}(\mathbb{R})$ = Borel σ algebra of $\mathbb{R}$

We then define the CDF by
$$F_X(x) = P\{\omega \in \Omega : X(\omega) \le x_1\}$$

Take the set $(-\infty, x_1]$. $P_X((-\infty, x_1]) = F_X(x_1)$

This holds for discrete random variables as well, but the probability of range is the $\sum P(x)\ \forall x \in$ range.



CDF for die roll assuming the above Rand Variable lt

classmate

Date

Page

If the CDF graph has discontinuous, the random variable is discrete and vice versa

CDFs must be monotonically increasing (or staying const) and it must hit 1 at some point.

Mixed Random Variables have cont. & disc. parts in the CDF graph

All CDFs for continuous rand variables (we consider) is differentiable.

Derivative of the CDF = Density fn (How much is conc. over a len)

Ex of CRV → Pick a no. from $[a,b]$

↳ Level of water in a dam

↳ Server work load

Probability Density fn:- (unit → probability per unit len)

$$\text{PDF} \Rightarrow f_X(x) = \lim_{\Delta \to 0^+} \frac{P_X(x < X \leq x+\Delta)}{\Delta} \rightarrow \mathbb{P}(A \cap B^c)$$

$$\downarrow$$

$$\mathbb{P}(A) - \mathbb{P}(B)$$

Any value works, we just normalize

$$= \lim_{\Delta \to 0^+} \frac{F_X(x+\Delta) - F_X(x)}{\Delta}$$

$$= \frac{d F_X(x)}{dx} \quad \text{(if derivative exists)}$$

Equivalently $F_X(x) = \int_{u=-\infty}^{x} f_X(u)\, du$

Properties:-

(i) $P_X(\mathbb{R}) = \int_{-\infty}^{\infty} f_X(u)\,du = 1$

(ii) $P_X([a,b]) = F_X(b) - F_X(a) = \int_a^b f_X(u)\,du$

(iii) In general, for any $B \subseteq \mathbb{R}$, $P_X(B) = \int_{u \in B} f_X(u)\,du$

(iv) $P_X(\{a\}) = 0$ (no mass at any point)
↳ Contradiction proof
↳ or $\int_a^a f_X(u)\,du = 0$ by defn

(v) $P_X([a,b]) = P_X((a,b]) = P_X([a,b)) = P_X((a,b))$

Mean, Variance & Moments:-

$E[X] = \int_{-\infty}^{\infty} u f_X(u)\,du$

$E[X^n] = \int_{-\infty}^{\infty} u^n f_X(u)\,du$

$E[g(X)] = \int_{-\infty}^{\infty} g(u) f_X(u)\,du$

$Var[X] = E[g(X)]$ & $g(X) = (X - E[X])^2$

For $Y = aX + b$
(i) $E[Y] = aE[X] + b$
(ii) $F_Y[y] = F_X\left(\frac{y-b}{a}\right)$ (Only $a > 0$)
$a < 0 \rightarrow 1 - F_X\left(\frac{y-b}{a}\right)$

Examples:-

(i) Picking uniformly from $[a, b]$

$$f_X(x) = \frac{1}{b-a} \quad \forall\ x \in [a, b]$$

$$\cancel{CDF = \int_{-\infty}^{a} \frac{1}{b-a}\, du = \frac{1}{b-a}\, u \Big|_{-\infty}^{a}}$$

$$CDF \Rightarrow F_X(x) = \begin{cases} 0 \ ; & x < a \\ \frac{x-a}{b-a} \ ; & x \in [a, b] \\ 1 \ ; & \text{otherwise} \end{cases}$$

$$E[X] = \frac{a+b}{2} \qquad Var = \frac{(b-a)^2}{12}$$

(ii) Exponential Random Variable $(Exp(\lambda))$

$$f_X(x) = \lambda e^{-\lambda x} \quad \text{for } x \geq 0$$

$$CDF = F_X(x) = 1 - e^{-\lambda x} \quad \forall\ x \geq 0$$

$$E[X] = \frac{1}{\lambda} \qquad Var = \frac{1}{\lambda^2} \qquad E[X^n] = \frac{n!}{\lambda^n}$$

Building block of Markov Chains.
Are Memoryless.

$$P(X > a+h \mid X > a) = \frac{e^{-\lambda(a+h)}}{e^{-\lambda(a)}} = e^{-\lambda(h)} = P(X > h)$$