29/09/23

Post Midsem Class

→ Motivation for independence of 2 events.

→ Suppose you have 2 events:-

Case 1:- $P(B) = 0 \Rightarrow P(A \cap B) = 0$

$P(A \cap B) = P(A) \cdot P(B)$

Case 2:- $P(B) > 0$

- Here, A is independent of B should mean occurrence/non occurrence of B should not affect the occurrence of event A $\Rightarrow P(A|B) = P(A)$

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

$$P(A \cap B) = P(A) \cdot P(B)$$

→ Fix $B_0 \in \mathcal{P}(\Omega)$ st $P(B_0) > 0$.

→ $P_{B_0} : \mathcal{P}(\Omega) \rightarrow \mathbb{R}$:- $P_{B_0}(A) = P(A|B_0) \quad \forall A \in \mathcal{P}(\Omega)$

Theorem:- P_{B_0} is a probability on Ω

clearly $\forall A \in \mathcal{P}(\Omega)$, $P_{B_0}(A) = \frac{P(A \cap B_0)}{P(B_0)} \geq 0$

$\Rightarrow$ axiom 1 is fulfilled

$$P_{B_0}(\Omega) = \frac{P(\Omega \cap B_0)}{P(B_0)} = \frac{P(B_0)}{P(B_0)} = 1 \quad [\because P(B_0) > 0]$$

$\Rightarrow$ axiom 2 is fulfilled.

We take pairwise disjoint events $A_1, A_2, \ldots \in \mathcal{P}(\Omega)$

$$P_{B_0}\left(\bigcup_{i=1}^{\infty} A_i\right) = P\left(\bigcup_{i=1}^{\infty} A_i \,\Big|\, B_0\right)$$

$$= \frac{P\left(\bigcup_{i=1}^{\infty} (A_i \cap B_0)\right)}{P(B_0)}$$

$$= \frac{\sum_{i=1}^{\infty} (P(A_i \cap B_0))}{P(B_0)}$$ [$\because$ A_i are pairwise disjoint, so are $A_i \cap B_0$'s.

$$= \sum_{i=1}^{\infty} \frac{P(A_i \cap B_0)}{P(B_0)}$$

$$= \sum_{i=1}^{\infty} P(A_i | B_0)$$

$$= \sum_{i=1}^{\infty} P_{B_0}(A_i)$$ $\Rightarrow$ Axiom 3 is also verified

Hence P_{B_0} is a Probability.

- All other properties of a probability are also inherited by P_{B_0}

Let A, B, C be events, $P(B \cap C) > 0$
$\Rightarrow$ $P(A | B \cap C)$ can also be written as $P(A | B, C)$

Theorem :- $$P(A | B, C) = \frac{P(A \cap B | C)}{P(B | C)}$$

"Probability of a conditional event can be conditioned further."

Proof:- RHS :- $$\frac{P(A \cap B | C)}{P(B | C)} = \frac{\frac{P(A \cap B \cap C)}{\cancel{P(C)}}}{\frac{P(B \cap C)}{\cancel{P(C)}}} = \frac{P(A \cap (B \cap C))}{P(B \cap C)} = P(A | B \cap C) = \text{LHS}$$

$[\because P(C) \neq 0]$

→ If A_1, A_2, A_3 are 3 independent events,

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2) \cdot P(A_3)$$

→ If they are not independent

$$P(A_1 \cap A_2 \cap A_3) = P(A_1) \cdot P(A_2 | A_1) \cdot P(A_3 | A_1, A_2) \quad \left[\text{Chain rule of Prob}\right]$$

if $\underline{\underline{P(A_1 \cap A_2) > 0}}$

Theorem (Chain Rule of Probability)

Let $A_1, A_2, \ldots, A_n$ are events such that $P(A_1 \cap A_2 \cap \cdots \cap A_{n-1}) > 0$,

then $P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n) = P(A_1) \cdot P(A_2 | A_1) \cdots P(A_n | A_1 \cap A_2 \cdots A_n)$

Proof:-

RHS :- $\cancel{P(A_1)} \cdot \dfrac{\cancel{P(A_1 \cap A_2)}}{\cancel{P(A_1)}} \cdot \dfrac{P(A_1 \cap A_2 \cap A_3)}{\cancel{P(A_1 \cap A_2)}} \cdots \dfrac{P(A_1 \cap A_2 \cap \cdots \cap A_n)}{P(A_1 \cap A_2 \cdots \cap (A_{n-1}))}$

$= P(A_1 \cap A_2 \cap A_3 \cap \cdots \cap A_n)$

$= \underline{\underline{\text{LHS}}}$

→ Composite Random Experiments:-

- It is a random experiment with two or more stages such that the outcome in the second stage depends on the outcome of the first stage and so on

Example:- ① SRSWOR

② Polya's Urn Scheme :- Fix $b, r, c \in \mathbb{N}$

'b' black balls
'r' red balls

all balls are indistinguishable except by color.

i) Choose a ball at random from the urn., note its colour and put it back

ii) Add 'c' balls of the same colour into the urn.

→ Run it under a for(or) while loop.

Events R_i, B_i :- $R_i :=$ i^{th} ball is red

$B_i :=$ i^{th} ball is black

$E :=$ exactly 1 of 3 balls drawn is red.

$$E = (R_1 \cap B_2 \cap B_3) \cup (B_1 \cap R_2 \cap B_3) \cup (B_1 \cap B_2 \cap R_3)$$

$$P(R_1 \cap B_2 \cap B_3) = P(R_1) \cdot P(B_2 | R_1) \cdot P(B_3 | R_1, B_2)$$

$$= \frac{r}{r+b} \cdot \frac{b}{r+b+c} \cdot \frac{b+c}{r+b+2c}$$

$$P(B_1 \cap R_2 \cap B_3) = \frac{b}{r+b} \cdot \frac{r}{b+r+c} \cdot \frac{b+c}{r+b+2c}$$

$$P(B_1 \cap B_2 \cap R_3) = \frac{b}{r+b} \cdot \frac{b+c}{r+b+c} \cdot \frac{r}{r+b+2c}$$

$$\therefore P(E) = \frac{3(rb(b+c))}{(r+b)(r+b+c)(r+b+2c)}$$

→ Generalisation for 'n' stages :- $(n \geq 2)$

$n_b + n_r = n$

↳ no. of draws

$P(E)$ where $E :=$ n_r many red balls are drawn.

$$P(E) = {}^{n}C_{n_r} \frac{r(r+c) \cdots (r+(n_r-1)c) \cdot b(b+c) \cdots (b+(n_b-1)c)}{(r+b)(r+b+c) \cdots (r+b+(n-1)c)}$$

[Both n_r, n_b non zero]

→ Case 2 ($n_b = 0$, $n_r = n$)

$$P(E) = \frac{1 \times (r)(r+c) \cdots (r+(n-1)c)}{(r+b)(r+b+c) \cdots (r+b+(n-1)c)}$$

[replace r's ~~root~~ by b's in numerator to get case 3 ($n_b = n$, $n_r = 0$)]

Back to Example ①:- SRSWOR

$\Omega = \{(i_1, i_2, \ldots, i_k)$ Each $i_j \in \{1, 2, \ldots, N\}$, i_j are all distinct$\}$

$|\Omega| = N(N-1)(N-2) \cdots (N-k+1)$

Fix $l_1, l_2, \ldots, l_k$ st $i_j \in \{1, 2, \ldots, N\}$

Let $E = \{(l_1, l_2, l_3, \ldots, l_k)\}$

$$P(E) = \frac{1}{N} \cdot \frac{1}{(N-1)} \cdot \frac{1}{(N-2)} \cdots \frac{1}{(N-k+1)} = \frac{1}{|\Omega|}$$

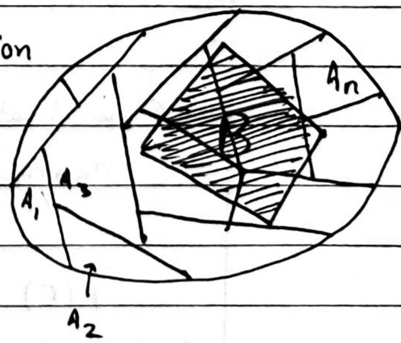
LAW OF TOTAL PROBABILITY

Def^n :- Events $A_1, A_2, \ldots, A_n$ form a partition of Ω if

(i) $A_1, A_2, \ldots, A_n$ are pairwise disjoint

(ii) $A_1, A_2, \ldots, A_n$ are exhaustive,

i.e., $\bigcup_{i=1}^{n} A_i = \Omega$



[Applies to a countably infinite no. of events]

Theorem:-

- Let Ω be a ctble sample space, and P be a probability on it. Suppose the events $A_1, A_2, \ldots, A_n$ form a partition of Ω. Then for any event B, we have

$$P(B) = \sum_{i=1}^{n} P(A_i)\, P(B \mid A_i) \qquad \text{[law of total prob.]}$$

proof:- $B = B \cap \Omega = B \cap \left[\bigcup_{i=1}^{n} A_i \right]$

$$= \bigcup_{i=1}^{n} B \cap A_i$$

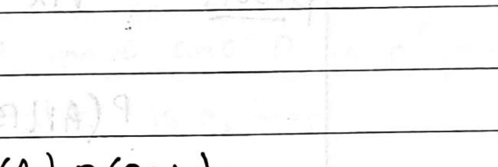
Clearly, $A_1, A_2, \ldots, A_n$ are pairwise disjoint
$\Rightarrow$ So are $B \cap A_1, B \cap A_2, \ldots, B \cap A_n$

$\therefore$ By finite additivity,

$$P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(A_i) \cdot P(B \mid A_i)$$

Example:-
→ 3 chests of drawers with 2 drawers each

$P(G) = 1/3 + 1/6 + 0$



$= 1/2$

* $P(C_2 \mid G) = \dfrac{1/6}{1/2} = \dfrac{1}{3}$ [Bayes' theorem exercise]

BAYES' THEOREM

- Suppose Ω is a ctble sample space and P is a probability on it. $A_1, A_2, \ldots, A_n$ form a partition of Ω st $P(A_i) > 0$ for each i.

- Let B be an event st $P(B) > 0$. Then for each i we have

$$P(A_i|B) = \frac{P(A_i)\, P(B|A_i)}{\sum_{j=1}^{n} P(A_j)\, P(B|A_j)}$$

Remark:- it also has a version for countable partitions of Ω.

proof:- Fix 'i' $\in \{1, \ldots, n\}$

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

$$= \frac{P(A_i) \cdot P(B|A_i)}{\sum_{j=1}^{n} P(A_j) \cdot P(B|A_j)}$$ [from Law of Total prob]

10|10|2023

Eg:- 5% of a population gets a rare disease

false positive :- 1% false negative :- 2%

If test shows positive, what is the prob. that the person actually has the disease

D :- person has disease T^+ :- Test +ve T^- :- Test -ve

$$P(D|T^+) = \frac{P(D)\, P(T^+|D)}{P(D) \cdot P(T^+|D) + P(ND) \cdot P(T^+|ND)} = \frac{5\% \times 98\%}{5\% \times 98\% + 95\% \times 1\%}$$

$$= \frac{\cancel{490}}{\cancel{585}} = \frac{98}{117}$$

Random Variables

10/10/23

→ A variable, whose value depends on chance.

Motivating Example:-

- A coin is tossed thrice.
- Let X = no. of heads out of 3 tosses.

outcome	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
value of X	3	2	2	2	1	1	1	0

value of X	0	1	2	3
probability	1/8	3/8	3/8	1/8

Definition:-

- Suppose Ω is a countable sample space and P is a probability on it. A discrete random variable is a map

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

→ Since Ω is countable, so is Range(X), so we call it a discrete r.v.

→ $P[X = x]$ for any $x \in \mathbb{R}$, $= \{\omega \in \Omega : X(\omega) = x\} \in \mathcal{P}(\Omega)$

$\Rightarrow$ It is enough to compute $P[X = x]$ for each $x \in$ Range(X)

Probability Mass Function:-

- $p_X : \mathbb{R} \to [0,1]$ given by $p_X(x) = P[X = x]$, $x \in \mathbb{R}$

- Clearly if $x \notin$ Range(X), then $p_X(x) = 0$.

Properties of pmf :-

0. $\{x \in \mathbb{R} : p_x(x) > 0\}$ is a countable subset of $\mathbb{R}$ (Range(X))

1. $P_x(x) \geq 0 \quad \forall x \in \mathbb{R}$

2. $\sum_{x \in \mathbb{R}} p_x(x) = 1$

Theorem :- $p : \mathbb{R} \to \mathbb{R}$ is the pmf of some discrete random variable if and only if it satisfies :-

① $p(x) \geq 0 \quad \forall x \in \mathbb{R}$

② $\{x \in \mathbb{R} : p(x) > 0\}$ is a countable subset of $\mathbb{R}$

③ $\sum_{x \in \mathbb{R}} p(x) = 1$

proof :- (only if) part is done

(if part) :-

Define $\Omega = \{x \in \mathbb{R}, p(x) > 0\}$

- for each $x_0 \in \Omega$, Let $P(\{x_0\}) = p(x_0)$

for any $A \subseteq \Omega$ $\quad P(A) = \sum_{x \in A} p(x)$

③ $\Rightarrow P(\Omega) = 1$

P is a prob on Ω

Define a map $X : \Omega \to \mathbb{R}$ by

$X(x) = x \quad \forall x \in \mathbb{R}$

$\forall x_0 \notin \Omega \quad P(X = x_0) = 0$ by ②

$\forall x_0 \in \Omega \quad P(X = x_0) = p(x_0)$

Hence proved.

11/10/23

Exercise:-

- Suppose you have an urn with 'r' red balls, 'b' black balls, 'g' green balls. All balls of same color are identical. Sample uniformly from this urn with replacement. Let N be the minimum no. of balls needed to see balls of all colors. Find pmf of N.

Sol^n :— range(N) = {3, 4, 5, ___, }

Cumulative Distribution Function:-

- The CDF ~~of~~ of a random variable X is a function $F_X: \mathbb{R} \longrightarrow [0,1]$ given by
$$F_X(u) = P[X \le u] \quad u \in \mathbb{R}$$

- Computing CDF from the "3 coins tossed example":-

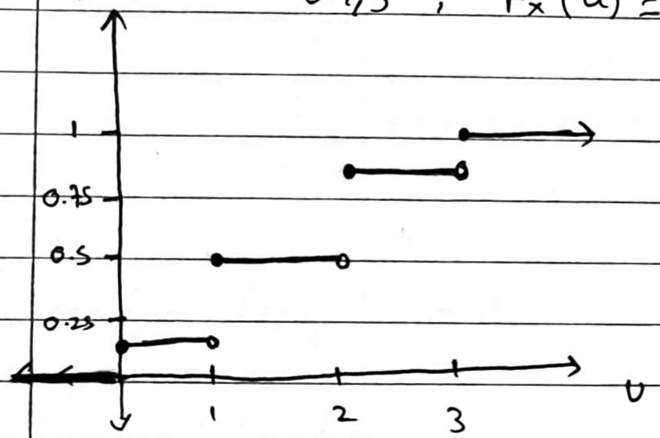
If $u < 0$, $F_X(u) = P[X \le u] = 0$

$1 > u \geq 0$, $F_X(u) = F_X(0) = P[X \le 0] = 1/8$

$2 > u \geq 1$, $F_X(u) = 1/2$

$3 > u \geq 2$, $F_X(u) = 7/8$

$u \geq 3$, $F_X(u) = 1$



Observation from graph of CDF:-

1) $F_X(u)$ is right continuous, i.e, $\forall\ u_0 \in \mathbb{R}$
$$\lim_{y \to u_0^+} F_X(y) = F_X(u_0)$$

2) F_X is non-decreasing

3) It is a step function that jumps exactly at all ~~at Range~~ $u \in$ Range (X).

4) Length of jump at $x = u$ is $P[X = u] = p_X(u)$

5) $\lim_{u \to -\infty} F_X(u) = 0$, $\lim_{u \to \infty} F_X(u) = 1$

Remark:- properties ①② ⑤ ~~is~~ are enjoyed by ~~any~~ CDF of any rv, In fact, they characterise the cdf. i.e., [any $f: \mathbb{R} \to \mathbb{R}$ if it fulfils ①, ②, ⑤, is the cdf of an ~~rv~~.]

2. Property ③, ④ are enjoyed by cdf of any discrete rv.

Examples and Families of Discrete rv's :-

0) Degenerate RV. Defⁿ :- A discrete rv X is called degenerate at $C\ (\in \mathbb{R})$ if $P[X=c]=1$

Notation $X \equiv C$

$$p_X(x) = \begin{cases} 1, & x=c \\ 0, & x \neq c \end{cases}$$

1) Bernoulli Distribution:-

- Fix $p \in (0,1)$
- A discrete rv X is said to follow bernoulli distribution with parameter p if $\text{Range}(X) = \{0,1\}$ and $P[X=1]=p$, $P[X=0]=1-p \in (0,1)$

Notation $X \sim Ber(p)$

$$p_X(x) = \begin{cases} 1-p, & x=0 \\ p, & x=1 \\ 0, & \text{otherwise} \end{cases}$$

Indicator Random Variables:-

Ω is a ctble sample space

$A \in \mathcal{P}(\Omega)$ is an event

$$I_A : \Omega \to \mathbb{R}, \quad I_A(\omega) = \begin{cases} 0, & \omega \notin A \\ 1, & \omega \in A \end{cases}$$

$P[I_A = 1] = P(A)$

Case 1) :- $P(A) = 0, \ P[I_A = 0] = 1$

Case 2) :- $P(A) = 1, \ P[I_A = 1] = 1$

Case 3) :- $P(A) \in (0,1), \quad I_A \sim \text{ber}[P(A)]$

$$P[I_A = 1] = P(A)$$
$$P[I_A = 0] = 1 - P[A)$$

Bernoulli Trials :-

- Trials with two possible outcomes :- Success(S), failure (F)
- For the next ③ families of d.r.v, we shall use repeated and independent Bernoulli trials.
- $P(S) = p \in (0,1)$ and it stays constant [Assumption]

2. Binomial Distribution :-

- ~~fix~~ $n \in \mathbb{N}$
- Let X = no. of successes out of n independent bernoulli trials.

$$\text{range}(X) = \{0, 2, \ldots, n\}$$

Take $k \in \text{Range}(X) = \{0, 1, 2, \ldots, n\}$

$$P[X = k] = \binom{n}{k} p^k q^{n-k}, \quad k \in \{0, 1, 2, \ldots, n\}$$

- 'X', a drv, follows binomial distribution with parameters (n, p) if x has pmf

$$p_X(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & \text{if } x \in \{0, \ldots, n\} \\ 0 & \end{cases}$$

Notation:- $X \sim Bin(n,p)$

$\rightarrow$ if $n=1$, $X \sim Ber(p)$

$$X \sim Bin(1,p) \iff X \sim Ber(p)$$

13/10/23

Suppose X = No. of Successes out of n Bernoulli Trials

then $X \sim Bin(n,p)$

Define r.v, $X_1, X_2, X_3, \ldots, X_n$ as follows

$$X_1 = I_{(1^{st}\text{ trial is a success})} = \begin{cases} 1, & \text{1st. trial success} \\ 0 & \text{2nd trial success} \end{cases}$$

$\vdots$

X_n

Each $X_i \sim Ber(p)$

$$X = X_1 + X_2 + X_3 + \cdots + X_n = \sum_{i=1}^{n} X_i$$

Each $X_i \in Ber(p)$

Ex. An mcq test has 20 q's, 4 choices - single correct type. A student eliminates 1 incorrect answer from each q, then chooses at random from the rest 3 options. The student will get a scholarship if 18 q's are correctly answered. Find prob that student gets scholarship.

$\_/\_/\_$

(3) Geometric Distribution:-

Suppose X denotes no. of ind Bernoulli trials $(P(S) = p \in (0,1))$ to get first success

$\text{range}(X) = \{1, 2, \ldots\}$

$k \in \mathbb{N} = \text{Range}(X)$

Then $P[X = k] = P[\underbrace{FFF\ldots F}_{(k-1)}S] = pq^{k-1}, \; k \in \mathbb{N}$

$$\Rightarrow \quad p_X(x) = \begin{cases} pq^{x-1}, & \text{if } x \in \mathbb{N} \\ 0, & \text{otherwise} \end{cases}$$

Why is this a valid pmf?

Clearly, $p_X(x) \geq 0 \quad \forall x \in \mathbb{R}$

Also, $\{x \in \mathbb{R} : p_X(x) > 0\} = \mathbb{N}$ which is ctble

Finally, $\sum_{x \in \mathbb{R}} p_X(x) = \sum_{x \in \mathbb{N}} pq^{x-1} = \sum p(1 + q + q^2 + \ldots)$

$$= p \cdot \frac{1}{1-q} \quad [\because |q| < 1]$$

$$= p \cdot 1/p = 1$$

Definition:-

A discrete r.v X is said to follow geometric distribution with parameter $p \in (0,1)$ if its pmf is

$$p_X(x) = \begin{cases} pq^{x-1}, & x \in \mathbb{N} \\ 0, & \text{otherwise} \end{cases}$$

Notation :- $X \sim \text{Geo}(p)$

Why is the $Bin(n,p)$ pmf

$$p_X(x) = \begin{cases} \binom{n}{x} p^x q^{n-x} & \text{if } x \in \{0,1,2,\dots n\} \\ 0 & \text{otherwise} \end{cases}$$

a valid pmf?

→ Clearly $p_X(x) \geq 0 \quad \forall x \in \mathbb{R}$

$\{x \in \mathbb{R}, p_X(x) > 0\}$ is countable

$$\sum_{x \in \mathbb{R}} p_X(x) = \sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = 1$$

Remark :- $Y \sim Geo(p)$ defined on no. of failures before the first success

$Y = X - 1 \Rightarrow$ Range$(Y) = \mathbb{N} \cup \{0\}$

pmf (Y) is

$$p_Y(y) = \begin{cases} pq^y, & y \in \mathbb{N} \cup \{0\} \\ 0 & \text{otherwise} \end{cases}$$

(4) Negative Binomial Distribution

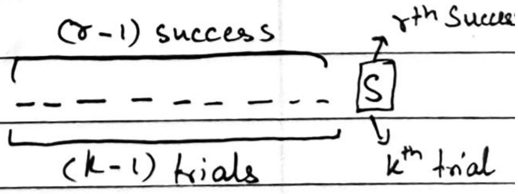
- Fix $r \in \mathbb{N}$, $p \in (0,1)$

Let X be no. of independent trials needed to get the rth success

Range$(X) = \{r, r+1, r+2, \dots\}$

$P(S) = p \in (0,1)$

Take $k \in \{r, r+1, \dots\}$



$(X = k)$ can happen in $\binom{k-1}{r-1}$ ways which are disjoint.

Each way corresponds to an arrangement of $(r-1)$ S's and $(k-r)$ failures.

prob of each such arrangement is $p^r q^{k-r}$

$\therefore$ By finite additivity, $P(X=k) = \binom{k-1}{r-1} p^r q^{k-r}$, $k \geq r$

Definition

a discrete r.v is ~~called~~ said to follow negative binomial distribution with parameter $r (\in \mathbb{N})$ and $p \in (0,1)$, if its pmf is,

$$p_X(x) = \begin{cases} \binom{x-1}{r-1} p^r q^{x-r} & , x \in \{r, r+1, \dots\} \\ 0 & , \text{otherwise} \end{cases}$$

Notation $X \sim NB(r, p)$

Remark :- In some books, the negative binomial r.v counts the no. of failures before the r^{th} success. Then,

$Y = X - r \Rightarrow \text{Range}(Y) = \mathbb{N} \cup \{0\}$

$$p_Y(y) = \begin{cases} \binom{y+r}{y} p^r q^y & , y \in \{0, 1, 2, \dots\} \\ 0 & , \text{otherwise} \end{cases}$$

Bernoulli $\longleftrightarrow$ Binomial

$X \sim Ber(p) \Leftrightarrow X \sim Bin(1, p)$

$S \sim Bin(n, p) \Leftrightarrow S = \sum_{i=1}^{n} x_i$, $X_i \sim Ber(p)$

Geometric

$X \sim Geo(p) \Leftrightarrow X \sim NB(1, p)$

Negative Binomial

If $Z \sim NB(r, p)$ then $Z = \sum_{i=1}^{r} X_i$ where each $X_i \sim Geo(p)$

Suppose $r = 2$, Z takes value 9 with following outcomes,

$$\underbrace{FFFS}_{\substack{x_1 \\ (4)}}\ \underbrace{FFFFS}_{\substack{x_2 \\ (5)}}$$

Clearly $Z = X_1 + X_2$

X_1 = No. of ind. Bernoulli trials for 1^{st} S

X_2 = No. of ind. Bernoulli trials for 2^{nd} S

$$Z = X_1 + X_2,$$

$$X_i \sim Geo(p)$$

$$\Rightarrow X = \sum_{i=1}^{n} X_i \quad \text{where each } X_i \sim Geo(p)$$

→ Sample space on which $X \sim NB(r,p)$ is defined.

$$\Omega = \bigcup_{n \geq r} \Omega_n \quad \text{where } \forall\, n \geq r,$$

$$\Omega_n = \{(o_1, o_2, \ldots, o_n),\ \text{Each } o_i \in \{S,F\},\ o_n = S,\ \sum_{i=1}^{n} I(o_i = S) = r\}$$

$$|\Omega| = \binom{n-1}{r-1}$$

Clearly, $\Omega_r, \Omega_{r+1}, \Omega_{r+2}, \ldots$, are pairwise disjoint

→ Define P on Ω, $\omega \in \Omega$

$\Rightarrow \omega$ belongs to only one of the Ω_i's for $i \geq r$,

$\Rightarrow P(\{\omega\}) = p^r q^{n-r}$

This defines a prob because,

$$\sum_{\omega \in \Omega} P(\{\omega\}) = \sum_{n \geq r} \sum_{\omega \in \Omega} p^r q^{n-r}$$

$$= \sum_{n \geq r} \binom{n-1}{r-1} p^r q^{n-r} = 1$$

$\Rightarrow X : \Omega \to \mathbb{R}$

$$X(\omega) = \begin{cases} r, & \text{if } \omega \in \Omega_r \\ r+1, & \text{if } \omega \in \Omega_{r+1} \\ \vdots & \vdots \end{cases}$$

i.e, $X(\omega) = n \quad \forall\, n \geq r$

One can easily check that,

$X \sim NB(r,p)$ by computing the pmf of X.

__/__/__

5 Uniform Distribution on a finite set

→ Suppose $A \subseteq \mathbb{R}$ is a finite set

Let $A = \{a_1, a_2, \ldots, a_n\} \Rightarrow |A| = n$

Def^n :- A rv X is said to follow uniform distribution on the set A, if X has pmf.

$$P_X(x) = \begin{cases} 1/n & \text{if } x \in A \\ 0, & \text{otherwise} \end{cases}$$

This is clearly a valid pmf

Notation :- $X \sim Unif(A)$

An Important special case :-

$A = \{1, 2, \ldots, n\}$

$\Rightarrow X = Unif(\{1, 2, 3, 4, \ldots, n\})$

↳ Discrete Uniform Distribution.

6 Hypergeometric Distribution :-

- Fix $N, M, n \in \mathbb{N}$ st $n \leq N$ and $0 \leq M \leq N$
- Suppose there are N items out of which $\underbrace{M}_{\text{identical}}$ are of type 1, N-M are of type 2.
- We select 'n' items from this N.

Let X = ~~prob~~ no. of items of Type 1 in the 'n' selected items.

→ Range(X) = ?

$\max(0, n-N+M) \leq X \leq \min(n, M)$

$$P(X = k) = \frac{\binom{n}{k}\binom{N-M}{n-k}}{\binom{N}{n}}$$

$Range(X) = \{k \in \mathbb{Z}, \max\{0, n-N+M\} \leq k \leq \min\{n, M\}\}$

Definition :-

An rv X is said to follow hypergeometric distribution with parameter N, M, n, if it has prob

$$p_X(x) = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}}, & \text{if } x \in \text{Range}(X) \\ 0, & \text{otherwise} \end{cases}$$

Ex :- Check that this is a valid prob

Notation :- $X \sim Hyp(N, M, n)$

Application :- capture - recapture method :-

- Suppose a pond has N fishes

Goal :- estimate N.

Method :-

- Catch M fish and mark them
- Put them back in the pond
- Catch 'n' of them. Suppose X of them are marked.

$$\hat{N} = \frac{nM}{X}$$

→ Proportion of marked fish in the population $= \frac{M}{N}$

If the sample is good :- $\frac{X}{n} \sim \frac{M}{N}$

7. Poisson Distribution :-

- An rv X is said to follow Poisson distribution with parameter $\lambda \in (0, \infty)$ if its pmf is

$$p_X(x) = P(X = x) = \begin{cases} e^{-\lambda} \dfrac{\lambda^x}{x!}, & \text{if } x \in \mathbb{N} \cup \{0\} \\ 0, & \text{otherwise} \end{cases}$$

Q. Where does it arise?

→ Suppose ~~for~~ $X \sim Bin(n,p)$ st n is very large and p is very small such that $np \approx \lambda (\in (0,\infty))$. Then X approximately follows Poisson distribution with parameter λ.

Eg:- No. of surviving bacteria in a ~~p~~ bacterial colony, no. of roadaccidents in a country over a period of time are modeled using Poisson distribution.

MEAN / EXPECTATION / Expected value.

"Avg value" of $X = \sum(\text{value}).(\text{Prob})$
where X is ~~the~~ a discrete random variable.

20/10/2023

Suppose X takes finitely many values say, $x_1, x_2, \ldots, x_n$ with probabilities $p_1, p_2, p_3, \ldots, p_n$. In other words, $Range(X) = \{x_1, \ldots, x_n\}$ with $P[X = x_i] = p_i > 0$. Clearly $\sum p_i = 1$

In this case, X has a finite mean

$$E(X) = \sum_{i=1}^{n} x_i p_i = \sum_{x \in \mathbb{R}} x\, p_X(x)$$

Now, suppose that $Range(X)$ is countably infinite. In this case, we shall look at the sums

$$\sum_{x>0} x\, p_X(x) \quad \text{and} \quad \sum_{x<0} x\, p_X(x)$$

Case 1:-

if $\sum_{x>0} x\, p_X(x) < \infty$, and $\sum_{x<0} (-x)\, p_X(x) < \infty$

P $\qquad$ N

$\Rightarrow X$ has finite mean,

$$E(X) = \sum_{x>0} x\, p_X(x) + \sum_{x<0} x\, p_X(x) = \sum_{x\in\mathbb{R}} x\, p_X(x)$$

Note that,

$$\sum_{x\in\mathbb{R}} |x|\, p_X(x) = \sum_{x>0} x\, p_X(x) + \sum_{x<0} (-x)\, p_X(x) < \infty$$

Conversely if $\sum_{x\in\mathbb{R}} |x|\, p_X(x) < \infty$ then $P < \infty$, $-N < \infty$

$\rightarrow$ The series is ~~abb~~ absolutely convergent

Case 2:-

$P = \infty$, $N < \infty$

$\Rightarrow E(X) = \sum_{x\in\mathbb{R}} x\, p_X(x) = P - N \Rightarrow \underline{\infty}$

~~does~~ doesnot have finite mean.

Case 3:- $P < \infty$, $N = \infty$

$E(x) = P - N = -\infty$

Case 4:- $P = \infty$, $N = \infty$

$\Rightarrow$ In this case $E(x)$ is not defines.

Examples:-

0. $X = c \Rightarrow E(X) = c1 = c$

1. $X \sim Ber(p) \Rightarrow E(x) = 1\cdot p + 0\cdot q = p.$

2. $X \sim \text{Unif}\{x_1, x_2, \ldots, x_n\}$

$$E(X) = \sum_{i=1}^{n} x_i \cdot \frac{1}{n} = \frac{1}{n} \cdot \sum_{i=1}^{n} x_i = \bar{x}$$

3. $X \sim \text{Bin}(n, p)$

$$E(X) = \sum_{x=0}^{n} x\, p_X(x) = \sum_{x=0}^{n} x \cdot \binom{n}{x} p^x q^{n-x}$$

$$= \sum_{x=0}^{n} \frac{n!}{(x-1)!\,(n-x)!} p^x q^{n-x}$$

$$= \sum_{x=1}^{n} np \left[\frac{(n-1)!}{(x-1)!(n-x)!}\right] p^{x-1} q^{n-x}$$

$$= np\,(p+q)^{n-1}$$

$$= \underline{\underline{np}}$$

4. $X \sim \text{Poi}(\lambda)$

In this case, $E(X) = \sum_{x \in \mathbb{N} \cup \{0\}} x\, p_X(x)$

$$= \sum_{x=0}^{\infty} x e^{-\lambda} \frac{\lambda^x}{x!}$$

$$= \sum_{x=1}^{\infty} e^{-\lambda} \frac{\lambda^x}{(x-1)!}$$

$$= e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!}$$

$$= \lambda e^{-\lambda} \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda.$$

5. $X \sim Geo(p)$

- In this case,

$$E(X) = \sum_{x \in \mathbb{N}} x \, p_X(x)$$

$$= \sum_{x \in \mathbb{N}} x \, p q^{n-1}$$

$$= p \sum_{x=1}^{\infty} x q^{x-1}$$

$$= p \cdot (1-q)^{-2} = p \cdot \frac{1}{p^2} = \frac{1}{p}$$

6. $X \sim NB(r, p)$ Show that:- $E(X) = r/p$

7. Suppose x is discrete r.v with pmf

$$p_X(x) = \begin{cases} \frac{6}{\pi^2} \cdot \frac{1}{x^2}, & x \in \mathbb{N} \\ 0, & \text{otherwise} \end{cases}$$

[valid pmf as $\sum \frac{1}{x^2} = \frac{\pi^2}{6}$]

Here, $E(x) = \sum_{x \in \mathbb{R}} x \, p_X(x)$

$$= \sum_{x \in \mathbb{N}} x \cdot \frac{6}{\pi^2} \cdot \frac{1}{x^2}$$

$$= \frac{6}{\pi^2} \sum_{x \in \mathbb{N}} \frac{1}{x} = \infty$$

$\Rightarrow$ X doesnot have a finite mean

Note:- If given $x \in \{1, -2, 3, -4, \dots\}$, $E(x) \neq \frac{6}{\pi^2} \log_e 2$

___/___/___

Discrete Random Vectors or Jointly distributed discrete random Variables

→ We shall start with the bivariate case

→ Suppose X, Y are 2 (bivariate) discrete r.v.s, defined on the same countable sample space Ω with a prob P on it

This means that $X: \Omega \to \mathbb{R}$ and $Y: \Omega \to \mathbb{R}$ are well defined

Therefore, combining these two functions, we think of a function $\Omega \to \mathbb{R}^2$ defined by (X, Y) is called

$$\underset{\sim}{Z}(\omega) \longrightarrow (X(\omega), Y(\omega)) \quad \omega \in \Omega$$

a bivariate discrete random ~~variable~~ vector

Note:- Range of this map has to be countable as Ω is countable
Hence (X, Y) can take countably many values in $\mathbb{R}^2$.

25/10/2023

Jointly Distributed Discrete r.v.s or Discrete Random Vectors (bivariate case)

Two discrete r.v.s x and y defined on the same sample space Ω, which has a prob P on it.

$$x: \Omega \longrightarrow \mathbb{R} \quad , \quad y: \Omega \longrightarrow \mathbb{R}$$

Therefore we get a $\Omega \to \mathbb{R}^2$ defined by $\omega \longrightarrow (x(\omega), Y(\omega))$

This map is a bivariate discrete random vector.

___/___/___

Remarks:

① Since X, Y are defined on the same sample space we can talk about their joint distribution function, joint pmf, we can add them, multiply them etc.

② As in case of discrete r.v.s, we shall forget the map $\omega \longmapsto (X(\omega), Y(\omega))$ and think of (X, Y) as a two-dimensional vector whose value depends on chance.

We shall focus on computing probabilities of X and Y jointly taking various values.

Def^n:- For a bivariate ~~r.v~~ r.v, the joint cdf or joint distribution function is defined as

$$F_{X,Y}(x,y) = P[X \leq x, Y \leq y], \quad (x,y) \in \mathbb{R}^2$$

$$F_{X,Y} = \mathbb{R}^2 \longrightarrow [0,1] \text{ is a map}$$

$$F_X(x) = P[X \leq x] = P(\overbrace{X \leq x, Y < \infty}^{A}) \qquad A_n \nearrow A$$

$$= P[\lim_{y_n \to \infty} \underbrace{(X \leq x, Y \leq y_n)}_{A_n}]$$

$$= \lim_{n \to \infty} P[(X \leq x), (Y \leq y_n)]$$

$\because$ for any seq. $y_n \nearrow \infty$ we get

$$P[X \leq x] = \lim_{n \to \infty} P[(X \leq x), (Y \leq y_n)] \quad \text{as } n \to \infty, \; y_n \to \infty$$

$$\Rightarrow \quad P[X \leq x] = \lim_{y \to \infty} P[(X \leq x, Y \leq y)] \quad \forall x \in \mathbb{R}$$

$$F_X(x) = \lim_{y \to \infty} F_{X,Y}(x,y) \quad \forall x \in \mathbb{R}$$

$$F_Y(y) = \lim_{x \to \infty} F_{X,Y}(x,y) \quad \forall y \in \mathbb{R}$$

Ex:- Show that:-

① $\forall\ a,b \in \mathbb{R}$ $P(X>a, Y>b) = 1 - F_X(a) - F_Y(b) + F_{X,Y}(a,b)$

② $\forall\ a_1, b_1, a_2, b_2, \in \mathbb{R}$ with $a_1 < a_2$ and $b_1 < b_2$

$$P[a_1 < X \le a_2,\ b_1 < Y \le b_2]$$
$$= F_{X,Y}(a_2,b_2) - F_{X,Y}(a_1,b_2) - F_{X,Y}(a_2,b_1) + F_{X,Y}(a_1,b_1)$$

③ Using ② show that:-

$$F_{X,Y}(a_2,b_2) - F_{X,Y}(a_1,b_2) - F_{X,Y}(a_2,b_1) + F_{X,Y}(a_1,b_1) \geq 0$$

whenever $a_1, b_1, a_2, b_2 \in \mathbb{R}$, $a_1 < a_2$, $b_1 < b_2$.

Going back to univariate case:-

CDF of any r.v X is non-decreasing i.e, $a<b \Rightarrow F_X(a) \le F_X(b)$

Proof:- $a<b \Rightarrow P(X \le b) = P(X \le a) + P(a < x \le b)$

$$\Rightarrow F_X(b) - F_X(a) = P(a < x \le b) \geq 0$$
$$\Rightarrow F_X(b) \geq F_X(a)$$

Def^n :- For a bivariate discrete random vector (X,Y), the joint pmf is defined by $p_{X,Y}(x,y) = P[X=x, Y=y]$, $(x,y) \in \mathbb{R}^2$

$$[\,p_{X,Y} : \mathbb{R}^2 \longrightarrow [0,1] \text{ is a map}\,]$$

Clearly, $\text{Range}(X,Y) = \{(x,y) \in \mathbb{R}^2 : p_{X,Y}(x,y) > 0$

$$\subseteq \text{Range}(X) \times \text{Range}(Y)$$

Facts:- ① $\forall\ x \in \mathbb{R}$, $p_X(x) = \sum_{y \in \mathbb{R}} p_{X,Y}(x,y) = \sum_{y : p_{X,Y}(x,y)>0} p_{X,Y}(x,y)$

② $\forall\ y \in \mathbb{R}$ $p_Y(y) = \sum_{x \in \mathbb{R}} p_{X,Y}(x,y) = \sum_{x : p_{X,Y}(x,y)>0} p_{X,Y}(x,y)$

__/__/__

Example:- Polya's Urn Scheme

→ We draw 3 balls in Polya's Urn Scheme

$X =$ No. of red balls drawn in the first 2 draws

$Y =$ No. of black balls drawn in last 2 draws

→ $\text{Range}(X) = \text{Range}(Y) = \{0,1,2\}$

$\text{Range}(X,Y) = \{0,1,2\} \times \{0,1,2\}$

but $P(X=2, Y=2) = 0$, $P(X=0, Y=0) = 0$

Notation :- B_i = i^{th} drawn ball is black

R_i = i^{th} drawn ball is red

B_i^c "

$$
\begin{aligned}
p_{X,Y}(0,0) &= 0 \\
p_{X,Y}(0,1) &= P[B_1 \cap B_2 \cap R_3] = 1/12 \\
p_{X,Y}(0,2) &= 1/4 \\
p_{X,Y}(1,0) &= 1/12 \\
p_{X,Y}(1,1) &= 1/6 \\
p_{X,Y}(1,2) &= 1/12 \\
p_{X,Y}(2,0) &= 1/4 \\
p_{X,Y}(2,1) &= 1/12 \\
p_{X,Y}(2,2) &= 0
\end{aligned}
$$

Y \ X	0	1	2	Marginal pmf of $Y = p_Y$
0	0	1/2	1/4	1/3
1	1/12	1/6	1/12	1/3
2	1/4	1/12	0	1/3
Marginal pmf of $X = p_X$	1/3	1/3	1/3	1

__/__/____

Remarks :-

① p_X, p_Y, are called marginal pmfs because they are written in the margin of the table.

② Ingeneral $\sum_{(x,y) \in Range(x,y)} p_{X,Y}(x,y) = 0$

③ Note that $X, \cancel{y}Y \sim \text{Unif}\{0,1,2\}$

However X, Y are not "independent".

How to compute $P[X \le 1, Y \le 1]$?

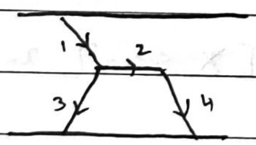
$$P(X \le 1, Y \le 1) = \sum_{x \le 1, y \le 1} p_{X,Y}(x,y)$$

$$= p_{X,Y}(0,1) + p_{X,Y}(1,0) + p_{X,Y}(1,1)$$

$$= 1/3$$

27/10/23

Drainage Network Model :-



Suppose :-

→ Paths behave independently

→ $P(\text{path } i \text{ is open}) = p \quad \forall\, i = 1,2,3,4$

X = No. of open paths

$$Y = \begin{cases} 1 & \text{, if water reaches the soft rock} \\ 0 & \text{, otherwise} \end{cases}$$

$X \backslash Y$	0	1	Marginal pmf P_X of X
0	q^4	0	q^4
1	$4pq^3$	0	$4pq^3$
2	$5p^2q^2$	p^2q^2	$6p^2q^2$
3	p^3q	$3p^3q$	$4p^3q$
4	0	p^4	p^4
Marginal pmf P_Y of Y	$1-p^2-p^3+p^4$	$p^2+p^3-p^4$	1

Remark:- If X takes a higher value, Y is more likely to take a higher value, similarly for lower values. X, Y are "not independent"

Properties of univariate CDF :-

- Suppose X is a r.v and $F(x) = F_X(x) = P(X \leq x)$, $x \in \mathbb{R}$ is its cdf
- Then :- (i) F is non-decreasing :-

Take $x_1 < x_2 \Rightarrow [X \leq x_1] \subseteq [X \leq x_2]$

$P[X \leq x_1] \leq P[X \leq x_2]$

$\Rightarrow F(x_1) \leq F(x_2)$. This proves 1

(ii) Take $x_0 \in \mathbb{R}$

To show :- F is right continuous $\Rightarrow \lim_{x \to x_0^+} F(x) = F(x_0) = \lim_{x \to x_0} F(x)$

Enough to show $\lim_{n\to\infty} F(x_n) = F(x_0)$ for any seq $x_n \searrow x_0$

Define $A_n = [X \leq x_n]$ $\forall n \geq 1$ and $A = [X \leq x_0]$ x_n is non increasing

Ex :- Check $A_n \searrow A$ i.e., $A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots \supseteq A_n$ (follows from $x_1 \geq x_2 \geq \cdots \geq x_0$)

and $\bigcap_{n \in \mathbb{N}} A_n = A$

If further, X is a discrete r.v, then F is a step function with jump discontinuities from the left side priecisely at the points $x \in$ Range(X) and the size of the jump at such an $x = P[X = x]$

$A_n \searrow A \Rightarrow P(A_n) \searrow P(A)$ [Continuity of prob from above

$F(x_n) = P(X \le x_n) \searrow P(X \le x_0) = F(x_0) \Rightarrow$ (ii) holds

(iii) $\lim_{x \to \infty} F(x) = 1$

Enough to show:-

$\lim_{n \to \infty} F(x_n) = 1 \quad \forall \; x_n \nearrow \infty$

Here $x_n \nearrow \infty$ means $x_1 \le x_2 \le x_3 \ldots\ldots$

and $\lim_{n \to \infty} x_n = \infty$ ie,

$\forall \; M > 0 \; \exists \; N = N_{(M)} \in \mathbb{N}$ st

$n \geq N \Rightarrow x_n > M$

Let $A_n = [X \le x_n]$, $n \in \mathbb{N}$ clearly $x_1 \le x_2 \le \ldots\ldots$

$\Rightarrow A_1 \subseteq A_2 \subseteq \ldots\ldots$

Check:- $x_n \nearrow \infty \Rightarrow A_n \nearrow \Omega$ i.e. $\bigcup_{n \in \mathbb{N}} A_n = \Omega$

PTO

$A_n \nearrow A \Rightarrow P(A_n) \nearrow P(A)$ (Continuity of prob from below)

In this case, $A = \Omega \Rightarrow P(A_n) \nearrow P(\Omega) = 1$

$\Rightarrow F(x_n) = P[X \leq x_n] \nearrow 1$

Thus proved (iii)

Remark:- We never used the fact that X is discrete.

Exercise:- Prove (iv). $[\lim_{x \to -\infty} F(x) = 0]$

Exercise:- Suppose X is an rv and $F(x) = F_X(x) = P[X \leq x], x \in \mathbb{R}$ is its cdf then show that $\forall\ x_0 \in \mathbb{R}$

$$\lim_{x \to x_0^-} F(x) = P(X < x_0)$$

Hint:- Enough to show $\lim_{n \to \infty} F(x_n) = P(X < x_0)$

for any sequence $x_n < x_0$ st $x_n \nearrow x_0$

In the discrete case:-

Case 1:- If $x_0 \notin \text{Range}(X)$ then $P(X = x_0) = 0$

$$\Rightarrow \underset{\underset{F(x_0)}{\parallel}}{P(X \leq x_0)} = P(X < x_0) \Rightarrow \lim_{x \to x_0^-} F(x) = F(x_0) = \lim_{x \to x_0^+} F(x)$$

If $x_0 \notin \text{Range}(X)$ then cdf F of X is actually continuous at x_0

Case 2:- If $x_0 \in \text{Range}(X) \Rightarrow P(X = x_0) = p_X(x_0) > 0$

In this case:- $\lim_{x \to x_0^-} F(x) = P(X < x_0) < P(X \leq x_0) = F(x_0)$

$$\left[\because P(X \leq x_0) = P(X < x_0) + \underbrace{P(X = x_0)}_{>0} > P(X < x_0)\right]$$

→ In particular, we have shown that F is not left. cont. at any $x_0 \in Range(X)$

- F is right cont. at $x_0 \Rightarrow$ it has jump discontinuity from left at each $x_0 \in Range(X)$.

$$\text{Size of jump} = F(x_0) - \lim_{x \to x_0^-} F(x) = P(X \leq x_0) - P(X < x_0)$$
$$= P[X = x_0]$$
$$= p_X(x_0) \geq 0$$

Remarks:-

① Properties i – iv canbe shown to hold for any rv X.

②. If a function $F: \mathbb{R} \longrightarrow [0,1]$ satisfies (i),(iv) then it is the cdf of some rv X.

Back to random vectors:-

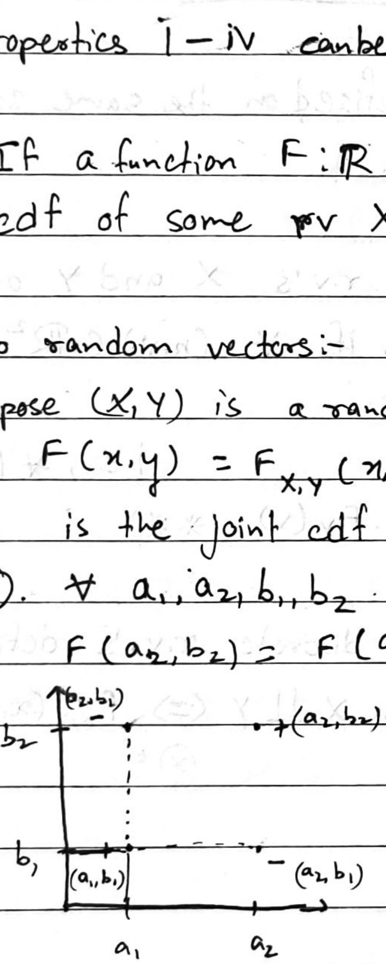
→ Suppose (X, Y) is a random vector and,

$$F(x,y) = F_{X,Y}(x,y) = P(X \leq x, Y \leq y), \quad (x,y) \in \mathbb{R}^2$$

is the joint cdf of X, Y, then F satisfies the following properties

①. $\forall\ a_1, a_2, b_1, b_2 \in \mathbb{R}$ st $a_1 < a_2$, $b_1 < b_2$, we have,

$$F(a_2, b_2) - F(a_1, b_2) - F(a_2, b_1) + F(a_1, b_1) \geq 0$$



ii) $F(x,y)$ is right continuous in both coordinates, i.e., if $x^{(n)} \searrow x$, $y^{(n)} \searrow y$, then $F(x^{(n)}, y^{(n)}) \searrow F(x,y)$ as $n \to \infty$

iii) $\longrightarrow \lim_{\substack{x \to \infty \\ y \to \infty}} F(x,y) = 1$ ie $\forall\ \varepsilon > 0\ \exists\ M > 0$ st

$$x > M, y > M \Rightarrow |F(x,y) - 1| < \varepsilon$$

$\Rightarrow \lim_{n \to \infty} (x^{(n)}, y^{(n)}) = 1$, whenever $x_n \nearrow \infty$, $y_n \nearrow \infty$

___/___/___

(iv) for each $y \in \mathbb{R}$ $\lim_{x \to -\infty} F(x,y) = 0$

For each $x \in \mathbb{R}$ $\lim_{y \to -\infty} F(x,y) = 0$

Exercise:- Verify (i) → (iv) for $F(x,y)$

Remark:- (It can be shown that $F : \mathbb{R}^2 \to [0,1]$ satisfying (i) → (iv) is the joint ~~cdf~~ of some random vector (X,Y).

Independence of two random variables:-

Suppose X and Y are r.v. defined on the same sample $X : \Omega \to \mathbb{R}$, $Y : \Omega \to \mathbb{R}$ with a prob P on it.

Defn:- We say that the r.v's X and Y are independent (and write $X \perp\!\!\!\perp Y$) if $\forall (u,v) \in \mathbb{R}^2$, $P(X \leq u, Y \leq v) = P(X \leq u) \cdot P(Y \leq v)$

i.e., $\forall (u,v) \in \mathbb{R}^2$

$F_{X,Y}(u,v) = F_X(u) \cdot F_Y(v)$ (**)

Theorem:- Let X, Y be two discrete r.v.'s defined on the same sample space then $X \perp\!\!\!\perp Y \iff P_{X,Y}(x,y) = P_X(x) P_Y(y)$ (*) $\forall (x,y) \in \mathbb{R}^2$.

proof:-

→ If part:-

Suppose (*) holds $\forall (x,y) \in \mathbb{R}^2$

Then $\text{Range}(X,Y) = \{(x,y) \in \mathbb{R}^2,\ p_{X,Y}(x,y) > 0\}$
$= \{(x,y) \in \mathbb{R}^2,\ p_X(x) \cdot p_Y(y) > 0\}$
$= \{(x,y) \in \mathbb{R}^2,\ p_X(x) > 0,\ p_Y(y) > 0\}$

31 / 10 / 2023

$= \{x : p_X(x) > 0\} \times \{y : p_Y(y) > 0\}$

$= Range(X) \times Range(Y)$

To show that $(**)$ holds,

$$F_{X,Y}(u,v) = P(X \le u, Y \le v)$$

$$= \sum_{\substack{(x,y) \in Range(X,Y) \\ x \le u, y \le v}} p_{X,Y}(x,y)$$

$$= \sum_{\substack{x \in Range(X), x \le u \\ y \in Range(Y), y \le v}} p_X(x)\, p_Y(y)$$

$$= \sum p_X(x) \cdot \sum p_Y(y)$$

$$= P(X \le u) \cdot P(Y \le v)$$

$$= F_X(u) \cdot F_Y(v)$$

$\Rightarrow (**)$ holds $\Rightarrow X \perp\!\!\!\perp Y$

Only if part:-

$\rightarrow$ Given $X \perp\!\!\!\perp Y \Rightarrow (**)$ holds

To show:- $(*)$ holds

Take $(x,y) \in \mathbb{R}^2$ To show:- $p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$

i.e., to show,

$$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$

Idea:- Approximate the point (x,y) using boxes, i.e., chosen two dimensional rectangles.

define events:-

$A_n = [x - \frac{1}{n} \le X \le x,\ y - 1/n \le Y \le y]$

clearly $A_1 \supseteq A_2 \supseteq A_3 \cdots\cdots$

s.t. $\cap A_n = [X=x, Y=y] = A$

$\Rightarrow A_n \searrow A$

$\Rightarrow P(A_n) \searrow P(A)$

$P(A_n) = P[(X,Y) \in (x-1/n, x] \times (y-1/n, y]]$

$= F_{X,Y}(x,y) - F_{X,Y}(x, y-1/n) - F_{X,Y}(x-1/n, y) + F_{X,Y}(x-1/n, y-1/n)$

$= F_X(x).F_Y(y) - F_X(x).F_Y(y-1/n) - F_X(x-1/n).F_Y(y) + F_X(x-1/n).F_Y(y-1/n)$

$= [F_X(x) - F_X(x-1/n)][F_Y(y) - F_Y(y-1/n)]$

$= P(x-1/n < X \le x) \cdot P(y-1/n < Y \le y)$

$\searrow P(X=x) \cdot P(Y=y)$

$[x-1/n < X \le x] \searrow [X=x]$

$[y-1/n < Y \le y] \searrow [Y=y]$

$\Rightarrow$ We have $P(A_n) \searrow P(A)$

$\Rightarrow P(A) = P(X=x) \cdot P(Y=y)$

$\Rightarrow P(X=x, Y=y) = P(X=x) \cdot P(Y=y)$.

Remarks:-

→ Corollary:-

(X,Y) is a d.r.v, then $X \perp\!\!\!\perp Y$ iff $P[X \in A, Y \in B] = P(X \in A) \cdot P(X \in B)$

$\forall\ A, B \subseteq \mathbb{R}$

Proof :- exercise.

__/__/__

① → The above corollary says that

② → ~~At~~ In both the previously discussed examples the r.vs X and Y not independent

Ex :- verify.

③ → If $X \perp\!\!\!\perp Y$, then $Range(X,Y) = Range(X) \times Range(Y)$

(follows from proof of 'if' part. Reverse does not hold.

'K' - dimensional discrete random vectors and Independence ($k \geq 2$)

$X_1, X_2, \ldots, X_k$ are k d.r.va's defined on the same sample space Ω with a prob P on it. This means that each $X_i : \Omega \to \mathbb{R}$ is a map. Combining these k maps, we get a map $\Omega \to \mathbb{R}^k$ defined by

$$\omega \longmapsto (X_1(\omega), X_2(\omega), \ldots, X_k(\omega)) = X$$

X :- is a discrete random vector of dimention k.

Joint CDF of 'k' dimensional r.ve :-

$$F_X = F_{X_1, X_2, \ldots, X_k} \quad \mathbb{R}^k \longrightarrow [0,1]$$

$$F_X(x) = F_{X_1, X_2, \ldots, X_k}(x_1, x_2, \ldots, x_k) = P(X \leq x_1, X \leq x_2, \ldots, X \leq x_k)$$

→ Marginal CDF's compiled taking limits of other variables tending to infinity.

Eg :- $$F_{X_1, X_2, X_3}(x_1, x_2, x_3) = \lim_{x_4 \to \infty} F_{X_1, X_2, X_3, X_4}(x_1, x_2, x_3, x_4)$$

__/__/__

Defⁿ:-

The joint pmf of $\underset{\sim}{X}$ is a map $\mathbb{R}^n \to [0,1]$

$$p_{X_1,X_2,\dots,X_n}(x_1,x_2,\dots,x_n) = P(X_1=x_1, X_2=x_2, \dots, X_n=x_n)$$

$\forall \; \underset{\sim}{x} \in \mathbb{R}^n$

Also, $\text{Range}(\underset{\sim}{X}) = \text{Range}(X_1,X_2,\dots,X_n)$

$$= \{\underset{\sim}{x} \in \mathbb{R}^n, \; p_X(x) > 0\}$$

$$\subseteq \text{Range}(X_1) \times \text{Range}(X_2) \dots \times \text{Range}(X_n)$$

As In the bivariate case,

$$P[\underset{\sim}{X} \in B] = \sum_{x \in B \cap \text{Range}(x)} p_X(x)$$

Joint cdf from joint pmf:-

$\forall \; \underset{\sim}{u} = (u_1,u_2,\dots,u_n) \in \mathbb{R}^n$

$$F_{X_1,X_2,\dots,X_k}(u_1,u_2,\dots,u_n) = P[X_1 \le u_1, \dots, X_n \le u_n]$$

$$= \sum_{x_1 \le u_1} \sum_{x_2 \le u_2} \dots \sum_{x_n \le u_n} (x_1,x_2,\dots,x_n)$$

→ Independence of 'k' random variables

→ $X_1,\dots,X_k$ are k rvs on same Ω with prob P defined on it.

Defⁿ:- $X_i \; (1 \le i \le k)$ are independent if $\forall \; \underset{\sim}{x} \in \mathbb{R}^n$,

$$P(X_1 \le x_1, X_2 \le x_2, \dots, X_n \le x_n] = \prod_{i=1}^{k} P[X_1 < x_n]$$

i.e., $F_{\underset{\sim}{X}}(\underset{\sim}{x}) = \prod_{i=1}^{n} F_{X_i}(x_i)$

Roughly speaking, this means that rvs $X_1, --, X_n$ do not influence each other.

<u>Theorem</u>:- Suppose $\underset{\sim}{x}$ is a drve. Then $X_1, ---, X_n$ are independent iff $p_x(x) = p_{X_1}(x_1) \cdot p_{X_2}(x_2) \cdots\cdots p_{X_n}(x_n)$
$\forall (x_1, x_2, \cdots x_n) \in \mathbb{R}^n$

proof:- Exe:-

<u>Corollary</u>:-

$\underset{\sim}{X}$ is a drve, then the X_i's are independent iff $P[X_1 \in A_1, X_2 \in A_2, ----, X_n = A_n] = \prod_{i=1}^{n} P[X_i \in A_i]$ $\forall A_i \subseteq \mathbb{R}$

<u>Remarks</u>:-

① If $X_1, ----, X_n$ are independent, then they donot influence each other in the sense that $[X_1 \in A_1], [X_2 \in A_2], ---, [X_n \in A_n]$ are independent $\forall A_1, --, A_n \in \mathbb{R}$

② If $X_1, --, X_n$ are independent then $Range(\underset{\sim}{X}) = Range(X_1) \times Range(X_2)$ $--- \times (Range\ X_n)$

③ If $X_1, X_2, ---, X_n$ are independent, then so are the rvs belonging to any subset of $\{X_1, X_2, --, X_n\}$.

$\Rightarrow X_1 \perp\!\!\!\perp X_3 \to$ Exe.

④ Suppose $X_1, ---, X_4$ are independent d.r.vs, then the following hold.

i) X_1^2, e^{X_2}, $\log(1+|X_3|)$, $\sin X_4$ are independent

ii) $(X_1^2 + X_3^2) \perp\!\!\!\perp (X_2^3 + X_4^7)$

iii) $(X_1 + X_2 + X_4) \perp\!\!\!\perp X_3$

(iv) $X_1+X_2,\ e^{X_3},\ X_4^2$ are independent

1/11/2023

Question

- How to compute expectation of a function of a random variable

Motivating Example:-

Suppose $X \sim \text{Unif}\{-1,0,1\}$

$\Rightarrow P(X=0) = P(X=-1) = P(X=+1) = 1/3$

Define $Y = X^2$

Goal:- To find $E(Y) = E(X^2)$

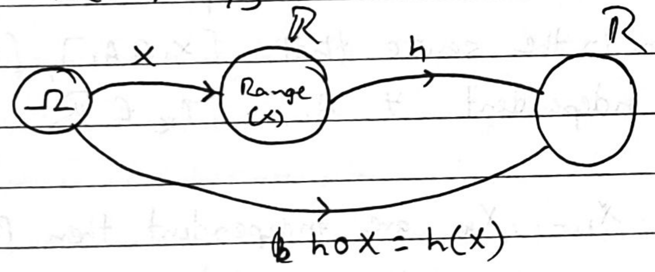
$\text{Range}(X) = \{-1,0,1\}$

$\text{Range}(Y) = \{0,1\}$

$\left.\begin{array}{l} P(Y=0) = P(X=0) = 1/3 \\ P(Y=1) = 1 - P(Y=0) = 2/3 \end{array}\right\} Y \sim \text{Ber}(2/3)$

$\Rightarrow E(Y) = E(X^2) = 2/3$

Setup:-



Question

How to Compute $E[h(X)]$

- How to figure out if $h(X)$ has finite mean
- If yes, how to compute?

Approach 1:-

Find out the pmf of $Y = h(X)$ (Too tedious)

Approach 2:-

Do it directly using the pmf of X (more practical)

Theorem :-

- Suppose X is a d.r.va with pmf p_X and $h : \text{Range}(X) \to \mathbb{R}$ is a function. Then $Y = h(X) = h \circ X$ is another d.r.va (on same sample space as X) Y has finite mean, provided

$$\sum_{x \in \text{Range}(X)} |h(x)|\, p_X(x) < \infty \quad , \quad E(Y) = \sum_{x \in \text{Rang}(X)} h(x) \cdot p_X(x)$$

Application of Theorem on the Motivating Example :-

$\text{Range}(X) = \{-1, 0, 1\}$

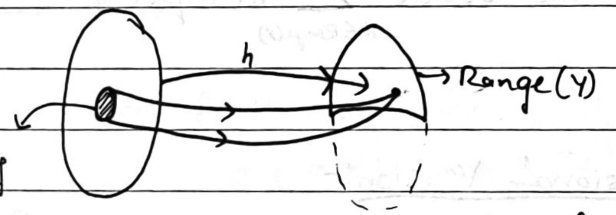
$h(x) = x^2$

Y has finite mean because $\text{Range}(Y)$ is a finite set

Theorem :- $E(X^2) = \sum_{x \in \text{Range}(X)} x^2 p_X(x)$

$$= \left[(-1)^2 \times \frac{1}{3}\right] + \left[0^2 \times \frac{1}{3}\right] + \left(1^2 \times \frac{1}{3}\right) = \frac{2}{3}$$

proof :-

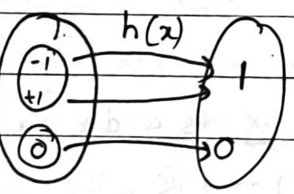


for each $y \in \text{Range}(Y)$ $\quad A_y = \{x \in \text{Range}(X) : h(x) = y\}$

$$= h^{-1}(\{y\})$$

Clearly, $\text{Range}(X) = \bigcup_{y \in \text{Range}(Y)} A_y$ pairwise disjoint union

prev. example :-



$A_1 = \{-1, 1\}$, $A_0 = \{0\}$

Back to proof :-

$$\sum_{y \in \text{Range}(Y)} |y|\, P[Y = y] < \infty$$

$$= \sum_{y \in Range(Y)} |y| \, P[X \in A_j]$$

$$= \sum_{y \in Range(Y)} |y| \sum_{x \in A_y} p_X(x)$$

$$= \sum_{y \in Range(Y)} \sum_{x \in A_y} |h(x)| \, p_X(x)$$

Therefore Y has finite mean provided

$$\sum_{x \in Range(X)} |h(x)| \, p_X(x) = \sum_{y \in Range(Y)} |y| \, P[Y=y] < \infty$$

Suppose $\sum_{x \in Range(X)} |h(x)| \, p_X(x) < \infty$

Then, $E(Y) = \sum_{y \in Range(Y)} y \, P[Y=y] = \sum_{x \in Range(X)} h(x) \, p_X(x)$

by the same calculation as before with the help of absolutely summability of the series $\sum_{x \in Range(X)} h(x) \, p_X(x)$

A Multi-dimensional Version:-

Setup:-

$$\begin{tikzcd} \Omega \text{ (ctble)} \arrow[r, "\underset{\sim}{X} = (X_1, \ldots, X_n)"] \arrow[rr, bend right, "Y = h(\underset{\sim}{X}) = h \circ \underset{\sim}{X}"'] & Range(\underset{\sim}{X}) \subseteq \mathbb{R}^k \arrow[r, "h"] & \mathbb{R} \end{tikzcd}$$

Theorem:- Suppose that $\underset{\sim}{X}$ is a d.r.v with joint pmf $p_X(\underset{\sim}{x})$ and $h : Range(\underset{\sim}{X}) \longrightarrow \mathbb{R}$. Then $Y := h(\underset{\sim}{x})$ is a discrete r.v with finite mean provided

$$\sum_{x \in Range(\underset{\sim}{x})} |h(\underset{\sim}{x})| \, p_X(\underset{\sim}{x}) < \infty$$ and in this case

$$E(Y) = \sum_{x \in Range(X)} h(\underset{\sim}{x}) \, p_X(\underset{\sim}{x})$$ (proof:- Exercise)

___/___/___

① **Theorem:-** Suppose $X_1, \ldots, X_n$ are d.r.va.s defined on same sample space st each X_i has finite mean. Then for all $\alpha_1, \alpha_2, \ldots, \alpha_k \in \mathbb{R}$, the linear combination $\alpha_1 X_1 + \alpha_2 X_2 + \ldots + \alpha_n X_n$ has finite mean and,

$$E\left(\sum_{i=1}^{k} \alpha_i X_i\right) = \sum_{i=1}^{k} \alpha_i E(X_i)$$

Remarks:-

① Note that the space V of all real valued random variables defined on the same sample space forms a vector space over $\mathbb{R}$

$\Rightarrow$ $S = \{X \in V : X \text{ has finite mean}\}$

is a subspace of V and expectation $E(X)$ is a linear map from $S \to \mathbb{R}$

② The linearity of expectation is very useful.

Proof of Theorem ①:-

Steps:-

(1) If X has finite mean then so does αX for any $\alpha \in \mathbb{R}$

$$E(\alpha X) = \alpha E(X)$$

(2) If X_1, X_2 have finite mean then so does $X_1 + X_2$

$$E(X_1 + X_2) = E(X_1) + E(X_2)$$

(3) ② $\Rightarrow$ additivity of expectation for k r.v.s.

Combining (1), ③ we get each X_i has finite mean

$\Rightarrow$ Each $\alpha_i X_i$ has finite mean

$\Rightarrow$ $\sum_{i=1}^{k} \alpha_i X_i$ has finite mean

Also, $E\left(\sum_{i=1}^{k} \alpha_i X_i\right) \overset{②}{=} \sum_{i=1}^{k} E(\alpha_i X_i) \overset{①}{=} \sum_{i=1}^{k} \alpha_i E(X_i)$

Proof of ①:- Take $\alpha \in \mathbb{R}$ and a d.r.va X with finite mean

Define a map

$$h : \text{Range}(X) \longrightarrow \mathbb{R}$$

by $h(x) = \alpha x$

Then $h(X) = \alpha X$. ~~By~~ It has finite by the theorem proved in class

because $\sum_{x \in Range(X)} |h(x)|\, p_X(x) = \sum_{x \in Range(X)} |\alpha x|\, p_X(x)$

$= |\alpha| \sum_{x \in Range(X)} |x|\, p_X(x) < \infty$

Therefore, $E(\alpha X) = E(h(X)) = \sum_{x \in Range(X)} h(x)\, p_X(x)$

$= \sum_{x \in Range(X)} \alpha x\, p_X(x)$

$= \alpha \sum_{x \in Range(X)} x\, p_X(x)$

$= \alpha E(X)$

proof of 2 :-

Suppose we have a double sequence $\{a_{ij}\}_{\substack{i \in \mathbb{N} \\ j \in \mathbb{N}}}$ of Real no-s.

Q:- When is $\sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} a_{ij} = \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} a_{ij}$? $(*)$

ans :- Not always

Fubini's theorem (Discrete version)

i) If $a_{ij} \geq 0$, then $(*)$ holds

ii) If $\sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} |a_{ij}| = \sum_{j \in \mathbb{N}} \sum_{i \in \mathbb{N}} |a_{ij}| < \infty$ then also $(*)$ holds.

_/_/_

Proof (2)

- Suppose x_1, x_2 have a finite mean, $\because$ both are discrete r.v defined on same Ω, (X_1, X_2) is a random vector. Let p_{X_1,X_2} be the joint pmf of X_1, X_2

Define $h: \text{Range}(X_1, X_2) \longrightarrow \mathbb{R}$ by
$h(x_1, x_2) = x_1 + x_2$, $(x_1, x_2) \in \text{Range}(X_1 X_2)$

To check that $X_1 + X_2 = h(X_1, X_2)$ has finite mean, we need to show

$$\sum_{(x_1,x_2)\in \text{Range}(X_1,X_2)} |h(x_1,x_2)|\, p_{X_1,X_2}(x_1,x_2) = \sum_{(x_1,x_2)} |x_1+x_2|\, p_{x_1,x_2}(x_1,x_2) < \infty$$

Note that $\text{Range}(X_1, X_2) \subseteq \text{Range}(X_1) \times \text{Range}(X_2)$
for each $(x_1, x_2) \in \text{Range}(X_1) \times \text{Range}(X_2)$ but
$(x_1, x_2) \notin \text{Range}(X_1, X_2)$, $p_{(X_1,X_2)}(x_1, x_2) = 0$

Hence, $$\sum_{(x_1,x_2)\in \text{Range}(X_1,X_2)} |x_1+x_2|\, p_{X_1,X_2}(x_1,x_2) = \sum_{x_1\in \text{Range}(X_1)} \sum_{x_2 \in \text{Range}(X_2)} |x_1+x_2|\, p_{X_1,X_2}(x_1,x_2)$$

$$\leq \sum_{x_1}\sum_{x_2} (|x_1| + |x_2|)\, p_{(X_1,X_2)}(x_1,x_2)$$

$$= \sum_{x_1}\sum_{x_2} |x_1|\, p_{(X_1,X_2)}(x_1,x_2) + \sum_{x_1}\sum_{x_2} |x_2|\, p_{(X_1,X_2)}(x_1,x_2)$$

$$= \mathrm{I} + \mathrm{II}$$

$\hookrightarrow$ Prove both are finite

$$\mathrm{I} = \sum_{x_1}\sum_{x_2} |x_1|\, p_{(X_1,X_2)}(x_1,x_2)$$

$$= \sum_{x_1\in \text{Range}(X_1)} |x_1| \sum_{x_2 \in \text{Range}(X_2)} p_{(X_1,X_2)}(x_1,x_2)$$

$$= \sum_{x_1\in \text{Range}(X_1)} |x_1| \cdot p_{X_1} < \infty$$ as X_1 has finite mean.

Similarly for II, we get it is a finite

$\because$ I + II $< \infty$

we get $\sum |x_1+x_2|\, p_{(X_1,X_2)}(x_1,x_2) < \infty$

$\Rightarrow X_1 + X_2$ has finite mean.

Therefore, $E(X_1+X_2) = \sum (x_1+x_2)\, p_{X_1,X_2}(x_1,x_2)$

$$= \sum_{x_1}\sum_{x_2} x_1\, p_{X_1,X_2}(x_1,x_2) + \sum_{x_1}\sum_{x_2} x_2\, p_{X_1,X_2}(x_1,x_2)$$

$$= \text{I} + \text{II}$$

$$= E(X_1) + E(X_2)$$

Remark:- We used the theorem stated in prev. class with $h(x_1,x_2) = x_1+x_2$ & both parts fubini's Theorem.

Corollary:- for any dr.ve. $\underset{\sim}{X} = (X_1, X_2, \ldots, X_k)$, if each X_i has finite mean ~~too~~, then so does $X_1+X_2+\cdots+X_k$ and $E(X_1+X_2+\cdots+X_k) = E(X_1)+\cdots+E(X_k)$. Addivity of expectation.

* Remark:- holds for dependent r.v aswell.

Applications:-

① $X \sim Bin(n,p)$

Then $X = X_1+X_2+\cdots+X_n$, each $X_i \sim Ber(p)$

Therefore, $E(X) = E(X_1)+\cdots+E(X_n) = np$

② $X \sim NB(r,p)$

$X = X_1+\cdots+X_r$, $X_i \sim Geo(p)$

$\sum E(X_i) = \frac{r}{p}$

③ $X \sim Hyp(N, M, n)$

N items $\to$ m of type ①

$\quad\to$ N-m of type ②

Choose n items w/o replacement

X = no. of I type ~~samples~~ items chosen in sample.

To compute $E(X)$

Let r.v.s $X_1, X_2, \ldots, X_n$ be defined as follows for each $i = 1, \ldots, n$

$$X_i = \begin{cases} 1, & X_i \text{ is of type ①} \\ 0, & X_i \text{ is of type ②} \end{cases}$$

$$\Rightarrow \quad X = X_1 + X_2 + \cdots + X_n$$

$$E(X_i) = \frac{M}{N} \quad \forall\ i = 1, \ldots, n$$

$$E(X) = n \times \frac{m}{N}$$

$\to$ Go back to Maxwell - Boltzmann Statistics

- $r \in \mathbb{N}$ distinguishable particles and $n \geq 2$ distinguishable cells
- Unlimited capacity of cells
- Arrange particles randomly in cells.

Define an r.v X as follows:-

X = no. of empty cells.

$E(X) \to$ difficult using pmf of X

$\Rightarrow$ using additivity of ~~expression~~ expectation:-

define $$X_i = \begin{cases} 1, & i^{th} \text{ cell empty} \\ 0, & i^{th} \text{ cell non-empty} \end{cases}$$

Clearly $X = X_1 + X_2 + \cdots + X_n$

$$E(X) = E(X_1) + \cdots + E(X_n)$$

$i = \{1, \ldots, n\}$ $E(X_i) = P(i^{th} \text{ cell is empty}) = \frac{(n-1)^r}{n^r} = \left(1 - \frac{1}{n}\right)^r$

$$E(X) = n \cdot \left(1 - \frac{1}{n}\right)^r$$

$$= \frac{(n-1)^r}{n^{r-1}}$$

Ex:-

- Compute expected no. of empty cells in the Bose-Einstien setup
- Recall the pole and flag problem, $r \in \mathbb{N}$ distinguishable flags, $n \geq 2$ " poles

  Arrange the flags at random on the poles

  Each pole has unlimited capacity

  Let X be the no. of empty poles. Compute $E(X)$

- In the sleepy secretary problem, compute the expected no. of letters sent to the right addresses.