SI427

S2: Room 105

<u>Ref</u>: ① Introduction to Prob theory by Hoel, Port and Stone — classic

② A first course in probability - Sheldon Ross — Basic

Part 1

③ Probability - Alan F. Karr — main for II

④ An intro to prob. and statistics - Rohatgi and Saleh. — Prob for II

Part 2

<u>Syllabus</u>: ① Basic notions:
sample space, Event space (sigma algebra)
probability space, properties, cond prob, Rand. var.

part I

② Discrete random variables, prop of it, dist.
③ cont. random variable
④ expectation, inequalities

---

⑤ mode of convergence (seq of random varb)
- conv. in probab.
- conv in dist
- almost sure conv
- Relation b/w them
- law of large numbers
- central limit theorem

part II

<u>grading</u>:
Midsem - 30%
Endsem - 40%
Quiz 1: 27th Aug
Quiz 2: 15th Oct
Quiz 3 and 4: Surprise
Best 3 out of 4.

<u>Note</u>: <u>No DX grade</u>

31st July : Question: predict fair/unfair coin? - tossing n times

↳ no of $H = m$

chances of head $= \frac{m}{n}$ → we will see how this $\frac{m}{n}$ will approach P.

↓

P

Question : we want to predict the chances of rain at IISTB campus from 8am-10am tomorrow?

Also want to predict amount of rain in the interval?

Basic info : - last 30 day data.
- time series analysis (better prediction)

Notions :

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

Examples :

① toss a coin : Possible outcomes: $\{H, T\}$

② Throwing a dice : Possible outcomes: $\{1, 2, 3, 4, 5, 6\}$

These are called (random) experiments. We know the possible outcomes but can't predict it in a specific trial.

Collection of possible outcomes of an experiment is denoted by Ω. This is called sample space.

Note: sample space is a non-empty set.

Questions : Examples 1: (i) what is the chance that H will appear? $\{H\}$
(ii) what is the chance that H and T will appear? $\{H\} \cap \{T\} = \phi$
(iii) H or T will appear? $\{H\} \cup \{T\} = \Omega$

Examples 2: (i) outcome even? $\{1, 3, 5\}$
(ii) Outcome is even? $\{2, 4, 6\}$
(iii) outcome is 1? $\{1\}$
(iv) outcome is 7? ϕ
(v) outcome is ≥ 1? Ω

these kind of situations are called events.

Events : Observe ① An event is a subset of sample space.

② Def : Event space is collection of all possible events. ← see whether event space will have/have-not all subsets of sample space?

③. If A is an event (subset of Ω) then A^c is an event.
· ϕ is an event
· If A, B are events then $A \cup B$ is an event.

Field / Algebra : let Ω be a non-empty set. A collection of subsets of Ω $\mathcal{A}$ is called field/algebra if the following holds:

(i) $\phi \in \mathcal{A}$ / $\Omega \in \mathcal{A}$

(ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

Example 3 : Tossing a coin until I get H.

(iii) $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

Ex: Suppose $\mathcal{A}$ is a field / Algebra. show that
(i) $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$
(ii) $A_1, A_2, \ldots A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$ → do

$\Omega = \{\omega_1, \omega_2 \ldots \omega_n \ldots\}$ ← sample space
$\omega_n = \underbrace{T.T \cdots T}_{n-1 \text{ times}}.H$

A = get an H at even trial ← event (subset of sample space)
$= \{\omega_2, \omega_4, \ldots\ldots\}$

Example 3 soln: ① what is prob of ω_n
② chances will be sum of chances $\omega_2 + \omega_4 + \cdots\cdots$ → Note: Just because they are disjoint, we add, otherwise we cannot do it.

here $A = \{\omega_2, \omega_4 \ldots\ldots\}$
$= \bigcup_{k=1}^{\infty} \{\omega_{2k}\}$ → This is something extra event spaces have. (Union of countable events)

Defn: Sigma field / Sigma Algebra / σ-field :

Let Ω be a non-empty set. A collection $\mathcal{F}$ of subsets of Ω is called a σ-field if the following hold.

(i) $\phi \in \mathcal{F}$ / $\Omega \in \mathcal{F}$
(ii) $A \in \mathcal{F} \Rightarrow A^c \in \mathcal{F}$
(iii) $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow \bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$

$(\Omega, \mathcal{F}, P)$ → probability space

Ω: Non-empty set / sample space
$\mathcal{F}$: σ-field / event space

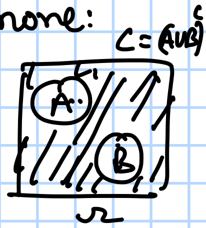
Example : σ-field :

① Ω → non-empty set
$\mathcal{F} = \{\phi, \Omega\}$ → smallest

② $\mathcal{F} = \mathcal{P}(\Omega)$ → largest
↳ power set of Ω

③ Suppose $A \subsetneq \Omega$
$\mathcal{F} = \{\phi, \Omega, A, A^c\}$ → By defination this is enough.

④ $\Omega = \{1, 2, \ldots 6\}$
$A = \{1\}$
let's take one more:
$B = \{2\}$
also as $A, B \in \mathcal{F}$
$A \cup B \in \mathcal{F}$
then
$A^c, B^c \in \mathcal{F}$



$\mathcal{F} = \{\phi, \Omega, A, B, A \cup B, C, A \cup C, B \cup C\}$

check $\mathcal{F}$ is a σ-field, and is smallest σ-field containg A and B.

Example 4 cont: $\mathcal{F} \subsetneq \mathcal{P}(\Omega)$

observations ① A σ-field is a field. → as closed under countable union means it is closed under finite union.

② If a field is finite (# $\mathcal{A}$ is finite) hold: Then it is σ-field.

In particular if Ω is finite then a field $\mathcal{A}$ over Ω is also a σ-field.

Ex: Given an Example of a field which is not a σ-field. → do

Note: ϕ (empty set) corresponds to impossible event.

Note: $(\Omega, \mathcal{F}, P)$ → probability space

expectation from P (prob)

(i) $0 \leq P(A) \leq 1$
(ii) $P(\phi) = 0$
(iii) $P(\Omega) = 1$
(iv) $P(A \cup B) = P(A) + P(B)$ if $A \cap B = \phi$ (Disjoint)
(v) If $A_1, A_2, \ldots$ are disjoint events ($A_i \cap A_j = \phi$ if $i \neq j$) then:

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

think of P as a function from σ-field $\mathcal{F}$ to $[0,1]$

$P: \mathcal{F} \longrightarrow [0,1]$

Ex: Suppose $\mathcal{A}$ is a field / Algebra. show that
(i) $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$
(ii) $A_1, A_2, \ldots A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$

proof: $\mathcal{A}$ is a field so
① $\phi \in \mathcal{A}$, $\Omega \in \mathcal{A}$
② $A \in \mathcal{A}$, $A^c \in \mathcal{A}$
③ if $A, B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

(i) $A, B \in \mathcal{A} \Rightarrow A \cap B \in \mathcal{A}$

as $A \in \mathcal{A}$ and $B \in \mathcal{A}$
$\Rightarrow A^c \in \mathcal{A}$ $\Rightarrow B^c \in \mathcal{A}$

now as $A^c \in \mathcal{A}$ and $B^c \in \mathcal{A}$
$\Rightarrow A^c \cup B^c \in \mathcal{A}$
$\Rightarrow (A \cap B)^c \in \mathcal{A}$
$\Rightarrow A \cap B \in \mathcal{A}$

(ii) $A_1, A_2, \ldots A_n \in \mathcal{A} \Rightarrow \bigcup_{i=1}^{n} A_i \in \mathcal{A}$

as $A_1, A_2 \in \mathcal{A}$
$\Rightarrow A_1 \cup A_2 \in \mathcal{A}$

as $A_1 \cup A_2 \in \mathcal{A}$ and $A_3 \in \mathcal{A}$

$\Rightarrow A_1 \cup A_2 \cup A_3 \in \mathcal{A}$

so for $n=2$ true,
for $n=n$ (suppose true) then

for $\underline{n+1}$: $A_1 \cup A_2 \dots A_n \in \mathcal{A}$
and
$A_{n+1} \in \mathcal{A}$

then $A_1 \cup A_2 \dots A_n \cup A_{n+1} \in \mathcal{A}$
(By induction)

$\rightarrow \mathcal{A} = \{A \subseteq \mathbb{Z} \mid A$ or A^c is finite$\}$

Ex: Given an Example of a field which is not a σ-field.

Here we see that a finite field is always a σ-field so our example should be such that the field is infinite.

Also we want $A_1, A_2, \dots \in \mathcal{A}$ but $\bigcup_{i=1}^{\infty} A_i \notin \mathcal{A}$ } If this occurs then $\mathcal{A}$ is not σ-field

now let $\mathcal{A} = \{A \subseteq \mathbb{Z} \mid A$ is finite or A^c is finite$\}$

Algebra as: ϕ is finite $\phi \in \mathcal{A}$
$\mathbb{Z} \in \mathcal{A}$ (as $\mathbb{Z}^c$ is ϕ)

if A is finite then
$A \in \mathcal{A}$ and $A^c \in \mathcal{A}$

same if $A \in \mathcal{A}$ and $B \in \mathcal{A}$
then
$A \cup B \in \mathcal{A}$

So $\mathcal{A}$ is field.

now let all singleton $\{n\} \in \mathcal{A}$ but
$\bigcup_{n=1}^{\infty} \{n\} = \mathbb{N}$ does not belong to $\mathcal{A}$ as
$\mathbb{N}^c$ (negatives) is also infinite

2nd Aug :

Ω, $\mathcal{F} \leftarrow \sigma$-field (event space)

Ω: non-empty set (sample space)

$P : \mathcal{F} \to [0,1]$

Defn: P (Probability measure) is a map from $\mathcal{F}$ to $[0,1]$ s.t.:

(i) $P(\Omega) = 1$

(ii) If $A_1, A_2, \ldots$ are disjoint events ($A_i \cap A_j = \phi$ if $i \neq j$) then

$$P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$$

Properties of P:

(i) $P(\phi) = 0$

proof: let $A_1 = \Omega$
$A_i = \phi, \ \forall i = 2, 3, \ldots$

as $A_i \cap A_j = \phi$ for any $i \neq j$

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

$$P(\Omega) = P(\Omega) + \sum_{i=2}^{\infty} P(A_i)$$

$$\Rightarrow \sum_{i=2}^{\infty} P(A_i) = 0$$

$$\Rightarrow P(A_i) = 0 \ \forall i = 2, 3, \ldots$$

$$(\text{As } P(A) \geq 0)$$

$$\Rightarrow P(\phi) = 0$$

(ii) $A_1, A_2, \ldots A_n$ are disjoint events then $P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i)$

proof: $A_1, A_2, \ldots A_n$,
$A_{n+1} = \phi, A_{n+2} = \phi, \ldots$

then

$$P(A_1 \cup A_2 \cup \cdots A_n \cup A_{n+1} \cdots) = P(A_1) + P(A_2) + \cdots P(A_n) + 0 + 0 + \cdots$$

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

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

proof: from (ii) and also fact

$$\Omega = A^c \cup A$$
$$A^c \cap A = \phi$$
$$P(A^c \cup A) = P(\Omega) = P(A^c) + P(A)$$
$$1 = P(A^c) + P(A)$$
$$P(A^c) = 1 - P(A)$$

(iv) Suppose $A, B \in \mathcal{F}$ and $A \subseteq B$ then $P(A) \leq P(B)$

Proof: $B = A \cup (B \setminus A)$

now firstly as $A \in \mathcal{F}$
$\Rightarrow A^c \in \mathcal{F}$

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

so, $B \cap A^c \in \mathcal{F}$

now $P(B) = P(A) + P(B \setminus A)$

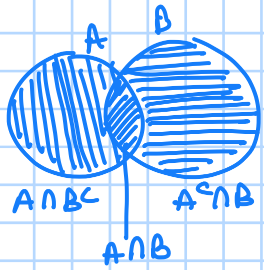
as $P(\text{Any}) \geq 0$

$$P(B) \geq P(A)$$

(v) $A, B \in \mathcal{F}$

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

Proof:



$P(A \cup B)$
$= P(A \cap B^c)$
$+ P(A \cap B)$
$+ P(A^c \cap B)$

also $P(A) = P(A \cap B^c) + P(A \cap B)$

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

(vi) Inclusion - exclusion priciple:

$A_1, A_2, \ldots A_n \in \mathcal{F}$

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k) \cdots + (-1)^{n+1} P\left(\bigcap_{i=1}^{n} A_i\right)$$

proof: let $A_1, A_2 \dots A_k \in \mathcal{F}$

for $k = 2$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$$

(Already proved)

let suppose true for k. then for $k+1$:

$A_1, A_2, \dots A_k, A_{k+1} \in \mathcal{F}$

$$P\left(\bigcup_{i=1}^{k+1} A_i\right) = P\left(\bigcup_{i=1}^{k} A_i \cup A_{k+1}\right)$$

$$= P\left(\bigcup_{i=1}^{k} A_i\right) + P(A_{k+1}) - P\left(\bigcup_{i=1}^{k} A_i \cap A_{k+1}\right)$$

$$= P\left(\bigcup_{i=1}^{k} A_i\right) + P(A_{k+1}) - P\left((A_1 \cap A_{k+1}) \cup (A_2 \cap A_{k+1}) \cdots\cdots\right)$$

$$= \sum_{i=1}^{k} P(A_i) - \sum_{i<j}^{k} P(A_i \cap A_j) + \cdots\cdots + (-1)^{k+1} P(A_1 \cap A_2 \dots A_k) + P(A_{k+1})$$

$$- \left[\sum_{i=1}^{k} P(A_i \cap A_{k+1}) - \sum_{i<j}^{k} P(A_i \cap A_j \cap A_{k+1}) + \cdots\cdots + (-1)^{k+1} P(A_1 \cap A_2 \dots A_{k+1})\right]$$

$$= \sum_{i=1}^{k+1} P(A_i) - \sum_{i<j}^{k+1} P(A_i \cap A_j) \cdots + (-1)^{k+2} P(A_1 \cap A_2 \dots A_{k+1})$$

so by k = true $\Rightarrow$ $k+1$ true

$\therefore$ By induction, true

Example: (i) $\Omega = \{H, T\}$, $\mathcal{F} = \mathcal{P}(\Omega)$

(a) fair coin: $\frac{1}{2} = P(H)$

① By $P(H \cup T)$ method

② $P(E) = \frac{\#\text{fav out}}{\text{total out}}$

only if all outcomes equally likely.

(b) not a fair coin: $P(H) = p = 1 - P(T)$

$\hookrightarrow$ not know/not given

(ii) throwing a dice:

$\Omega = \{1, 2, \dots 6\}$

$\mathcal{F} = \mathcal{P}(\Omega)$

(a) fair dice: $P(\{1\}) = \frac{1}{6}$ (By prev. methods)

Note: $P(\{i\}) = \frac{1}{6}, \ \forall i = 1, 2, \dots 6$

if $E \in \mathcal{F}$ eg. E(outcome is odd) $= \{1, 3, 5\}$

$$P(E) = P(\{1\} \cup \{3\} \cup \{5\})$$

$$= \frac{1}{6} \times 3 = \frac{1}{2}$$

(b) not a fair dice:

$P(\{1\}) = p \rightarrow$ not known/not given

$P(\{i\}) = p_i, \ i = 1, 2, \dots 6$

$0 \le p_i \le 1$

$\sum p_i = 1$

Note: given $E \in \mathcal{F}$, we can calculate $P(E)$ in terms of P_i's

If an outcome is a simple event / element event.

sometimes it is called atoms — μ

Def^n: $E \in \mathcal{F}$ is called an 'atom' if $E \neq \phi$ and there is no proper subset of E in $\mathcal{F}$.

Eg: ① $\Omega = \{1, 2, 3, 4\}$

$$\mathcal{F} = \{\phi, \Omega, \underset{A}{\{1\}}, \underset{A^c}{\{2,3,4\}}\} \rightarrow \text{This is } \sigma\text{-field}$$

$$\text{Atoms} = \underset{A}{\{1\}}, \underset{A^c}{\{2,3,4\}}$$

② $\Omega = \{1,2,3,4\}$ $\quad \mathcal{F} = \mathcal{P}(\Omega)$

$\{1\}, \{2\}, \{3\}, \{4\}$ are atoms.

Note: Now after identifying atoms, and assigning P values to these atoms, we can calculate P of any element in $\mathcal{F}$.

Example: Tossing two coins:

$$\Omega = \{(H,H), (T,T), (H,T), (T,H)\}$$

$$\mathcal{F} = \mathcal{P}(\Omega)$$

(a) coins are fair:

Atoms: $\{(H,H)\}, \{(T,T)\}, \{(H,T)\}, \{(T,H)\}$

$$P(\text{each atom}) = \frac{1}{4}$$

(b) not fair: by intuition if

$P(H) = p$

$P(H,H) = p^2$

$P(H,T) = p(1-p)$

$\vdots$

Example: Tossing a coin until we get H.

$$\Omega = \{\omega_1, \omega_2, \dots\}$$

$$\omega_i = \underbrace{TT\cdots T}_{i-1}H \qquad \mathcal{F} = \mathcal{P}(\Omega)$$

① fair:

$P(\omega_1) = \frac{1}{2}$

$P(\omega_2) = \frac{1}{4}$

$\vdots$

② Unfair:

$P(\omega_1) = p$

$P(\omega_2) = p(1-p)$

$\vdots$

Ex: Suppose there are n men in a party. They throw their hats into the centre of the room, then hats are mixed up, and each men select a hat randomly.

(i) Possibility that none of the man gets their own hats?

(ii) Exactly k men get their own hats?

Ω = set of all bijections from n-men to n-hat
= set of all permutations of n-objects

$\mathcal{F} = P(\Omega)$

(i) A_i = event that i^{th} man gets his hat
$P(A_i) = \frac{(n-1)!}{n!}$

A = event that no men get their hat
$= A_1^c \cap A_2^c \cap A_3^c \cdots \cap A_n^c$
$= (A_1 \cup A_2 \cup \cdots A_n)^c$

$P(A) = 1 - P(A_1 \cup A_2 \cup \cdots A_n)$

now, $P(A_i \cap A_j) = \frac{(n-3)!}{n!}$

$P(A_i \cap A_j \cap A_k) = \frac{(n-4)!}{n!}$

then $P(A) = 1 - \left[\sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) \cdots + (-1)^{n+1} P(A_1 \cap A_2 \cdots A_n)\right]$

$= 1 - \left[n \times \frac{1}{n} - \binom{n}{2}\frac{(n-2)!}{n!} + \binom{n}{3}\frac{(n-3)!}{n!} \cdots + (-1)^{n+1}\binom{n}{n}\frac{(n-n)!}{n!}\right]$

note: $\binom{n}{r} = \frac{n!}{(n-r)!(r)!}$

$P(A) = 1 - \left[1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} \cdots + (-1)^{n+1}\frac{1}{n!}\right] \approx e^{-1}$

(ii) now we can select k people by $\binom{n}{k}$ ways. let's fix them,

now $P(1^{st}$ guy gets his hat$) = \frac{1}{n}$

$P(2^{nd}$ guy gets his hat now$) = \frac{1}{n} \times \frac{1}{(n-1)}$

$\vdots$ $P(k^{th}$ guy gets his hat now$) = \frac{1}{n} \cdot \frac{1}{(n-1)} \cdots \frac{1}{(n-(k-1))}$

also so now after fixing them, remaining $n-k$

$P(A_{n-k}) = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} \cdots + (-1)^{n-k}\frac{1}{(n-k)!}$

total $P(E) = P(A_{n-k}) \times \frac{1}{(n)(n-1) \cdots \cancel{(n+1-k)}} \times \frac{\cancel{(n)!}}{\cancel{(n-k)!}(k)!} = \frac{1}{k!} \times P(A_{n-k})$

7th Aug:

matching problem (ii) what is the prob. that exactly k-men select their hat?

for (i) $P(A) = 1 - \left[1 - \frac{1}{2!} + \frac{1}{3!} \cdots + (-1)^{n+1}\frac{1}{n!}\right]$

$$P_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} + \cdots (-1)^n \frac{1}{n!} \approx e^{-1}$$

(ii) $P(1^{st} \text{ gets hat}) = \frac{1}{n}$

$P(1,2,3 \ldots k \text{ get hat}) = \frac{1}{n} \cdot \frac{1}{n-1} \cdot \frac{1}{n-2} \cdots \frac{1}{(n-k+1)}$

P none others $(n-k)$ get hats:

$$P_{n-k} = \frac{1}{2!} - \frac{1}{3!} + \cdots + (-1)^{n-k}\frac{1}{(n-k)!}$$

ways of selecting k from $n = \binom{n}{k}$

So $P_{final} = \binom{n}{k}\left(\frac{1}{n \cdot n-1 \cdots (n-k+1)}\right) P_{n-k}$

$= \frac{\cancel{n!}}{\cancel{(n-k)!}(k)!} \cdot \frac{1}{\cancel{\cdots}} P_{n-k}$

$= \frac{P_{n-k}}{k!}$

or $P_{n-k} = \frac{\#\text{ for permutations that none get own hat} = N}{\text{total no of permutations of } n-k}$

$\Rightarrow N = (P_{n-k})(n-k)!$

Actual problem: n men

out of n-men, exactly k of them get their own hat

P(exactly k out of n men get own hat) $\rightarrow$ 1 way $\times$ N

$= \frac{\binom{n}{k} \times N}{n!}$

1 2 ... k — own man — $\rightarrow$ No one gets hat — $\hookrightarrow$ none get own hat

$= \frac{P_{n-k}}{k!}$

or $\underbrace{\binom{n}{k}}_{\text{select k from n}} \times \underbrace{\frac{1}{(n)(n-1)\cdots(n-k+1)}}_{\text{P that those k gets own hat}} \times \underbrace{P_{n-k}}_{\text{P that n-k does not}}$

Continuity of probability

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\{A_n\}_{n\geq 1}$ be a seq. of increasing events i.e
$A_1 \subseteq A_2 \subseteq A_3 \subseteq \dots$

Then $P(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n)$

proof: Note $\bigcup_{n=1}^{\infty} A_n \in \mathcal{F}$ as $A_n \in \mathcal{F}$ $\forall n\geq 1$

$B_1 = A_1$

$B_2 = A_2 \setminus A_1$

$B_n = A_n \setminus A_{n-1}$ $\forall n \geq 2$

Note: $B_i \in \mathcal{F}$ $\forall i$

$\underbrace{\bigcup_{n=1}^{\infty} B_n = \bigcup_{n=1}^{\infty} A_n \; ; \; \bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{n} B_k}$ → Exercise → Done (Down)

$\Rightarrow P(\bigcup_{n=1}^{\infty} B_n) = P(\bigcup_{n=1}^{\infty} A_n) = \underbrace{\sum_{n=1}^{\infty} P(B_n)}_{\text{finite series as } P(\ ) \in B/w\ 0 \text{ and } 1} = \lim_{n\to\infty} \left(\sum_{k=1}^{n} P(B_k)\right)$

$= \lim_{n\to\infty} \left(\bigcup_{k=1}^{n} B_k\right)$

$= \lim_{n\to\infty} \left(\bigcup_{k=1}^{n} A_k\right)$

$= \lim_{n\to\infty} P(A_n)$

Exe: suppose $\{A_n\}_{n\geq 1}$ be a seq of decreasing events, then
$A_1 \supseteq A_2 \supseteq A_3 \supseteq \dots$ → Done (down)

$P(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} P(A_n)$

Conditional probability:

Let $(\Omega, \mathcal{F}, P)$ be a prob. space.

Let $B \in \mathcal{F}$ and $P(B) > 0$.

Then conditional prob. that an event A occurs given that B has occurred is
$P(A\cap B)/P(B)$

denoted by: $P(A|B) = P(A\cap B)/P(B)$ → as now both A and B occurs $P(A\cap B)$ and as sample space is reduced $P(B)$ ∴ $P(A|B) = \frac{P(A\cap B)}{P(B)}$

dice: even number occurred i.e $\{2, 4, 6\}$

$P(4 \text{ has occurred}) = \frac{1}{6}$

$P(4 \text{ has occurred} | \text{even occurred}) = \frac{1}{3} = \frac{1/6}{3/6} = \frac{1}{3}$

Properties of cond. probability:

Result:

(i) Suppose $0 < P(B) < 1$

then $P(A) = P(A|B)\,P(B) + P(A|B^c)\,P(B^c)$

proof: $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$

$$P(A\cap B) = P(A|B)\,P(B)$$
$$P(A\cap B^c) = P(A|B^c)\,P(B^c)$$

$$A = (A\cap B) \cup (A \cap B^c)$$

$$P(A) = P(A\cap B) + P(A\cap B^c)$$
$$P(A) = P(A|B)\,P(B) + P(A|B^c)\,P(B^c)$$

(ii) Suppose $B_1, B_2, \ldots, B_n$ are partitions of Ω, and $P(B_i) > 0 \;\; \forall i = 1, 2, \ldots n$

LAW OF TOTAL PROB / PARTITION THEORM

$$P(A) = \sum_{i=1}^{n} P(A|B_i)\,P(B_i)$$

Prob. 2 boxes

Box 1 : 2 white 3 red

Box 2 : 3 white 4 red

A Ball is picked from Box 1 at random, and put into Box 2. Then a ball is picked from Box 2. and colour is noted. What is the prob. that it is red.

$P(\text{red}) = P(\text{red} \mid \text{white was picked from I}) \times P(\text{white was picked from I})$
$+ P(\text{red} \mid \text{red was picked from I}) \times P(\text{red was picked from I})$

$$= \left(\frac{4}{8}\right) \times \frac{2}{5} + \left(\frac{5}{8}\right) \times \frac{3}{5}$$

A: Ball chosen from Box 2 is red.
B: Ball chosen from Box 1 is red.

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B)$$

$P(A|B) = \dfrac{5}{8}$

$P(A|B^c) = \dfrac{4}{8}$

$P(B) = \dfrac{3}{5}$

$P(B^c) = \dfrac{2}{5}$

(iii) let $A_1, A_2, \ldots\ldots A_n \in \mathcal{F}$ and $P(\bigcap_{k=1}^{n} A_k) > 0$

as > 0 $A_i \neq \emptyset, \forall i = 1, 2, \ldots n$

then: $P(\bigcap_{k=1}^{n} A_k)$ multiplication theorm

$= P(A_1)\,P(A_2|A_1)\,P(A_3|A_2 \cap A_1) \ldots \ldots P(A_n|A_1 A_2 \ldots A_{n-1})$

$P(E_1 E_2 \ldots E_n)$
$= P(E_1)\,P(E_2|E_1)\,P(E_3|E_1 E_2) \ldots\; P(E_n|E_1 E_2 \ldots E_{n-1})$

proof: $P(A_1)\, P(A_2 | A_1) \cdots$

$= P(A_1) \cancel{\dfrac{P(A_1 \cap A_2)}{P(A_1)} \dfrac{P(A_1 \cap A_2 \cap A_3)}{P(A_1 \cap A_2)} \cdots}$

$= P(\bigcap_{i=1}^{n} A_k)$

(iv) $\omega(\Omega, \mathcal{F}, P)$ be a prob space and $A \in \mathcal{F}$ and $P(A) > 0$

Define: $\tilde{P}(B) = P(B | A) \quad \forall B \in \mathcal{F}$

then $\tilde{P}$ is a prob. measure/map on $\mathcal{F}$.

that means $(\Omega, \mathcal{F}, \tilde{P})$ is a prob. space

&

Non empty set
- Ω
- $\mathcal{F}$ σ-field
- $\tilde{P}$ some conditions

$\tilde{P} : \mathcal{F} \to [0,1]$

$\tilde{P}(B) = P(B|A), \forall B \in \mathcal{F}$

proof: To check $\tilde{P} : \mathcal{F} \to [0,1]$

$$\tilde{P}(B) = \frac{P(B \cap A)}{P(A)} \geq 0$$

as $P(A \cap B) \leq P(A)$

$\Rightarrow 0 \leq \tilde{P}(B) \leq 1$

$\tilde{P}(\phi) = P(\phi)/P(A) = 0$

$\tilde{P}(\Omega) = \dfrac{P(A)}{P(A)} = 1$

let $B_1, B_2, B_3, \ldots \in \mathcal{F}$ and $B_i \cap B_j \neq \phi$ for $i \neq j$

$$\tilde{P}(\bigcup_{n=1}^{\infty} B_n) = \frac{P((\bigcup_{n=1}^{\infty} B_n) \cap A)}{P(A)}$$

$$= \frac{P(\bigcup_{n=1}^{\infty} (B_n \cap A))}{P(A)}$$

$$= \sum_{n=1}^{\infty} \frac{P(B_n \cap A)}{P(A)} \quad (\because \text{countable identity of } P)$$

$$= \sum_{n=1}^{\infty} P(B_n | A) = \sum_{n=1}^{\infty} \tilde{P}(B_n)$$

Exe: To prove: $\bigcup_{k=1}^{n} A_k = \bigcup_{k=1}^{n} B_k$

Proof: now $B_1 = A_1$

$B_2 = A_2 \setminus A_1$

$B_n = A_n \setminus A_{n-1}$

Note: $\forall i \leq n \quad A_i \subseteq A_n$

now, $\bigcup_{i=1}^{2} B_i = B_1 \cup B_2 = A_1 \cup (A_2 \setminus A_1)$

$= A_1 \cup A_2$

$= \bigcup_{i=1}^{2} A_i$

$\therefore$ for $n=2$, true

let for $n=k$ true then:

$$\bigcup_{i=1}^{k} B_i = \bigcup_{i=1}^{k} A_i$$

now for $n=k+1$

$$\left(\bigcup_{i=1}^{n} B_i\right) \cup B_{k+1} = \left(\bigcup_{i=1}^{k} A_i\right) \cup \left(A_{k+1} \setminus A_k\right)$$

$$= A_k \cup (A_{k+1} \setminus A_k)$$

$$= A_{k+1}$$

$$= \bigcup_{i=1}^{k+1} A_i$$

$\therefore$ true $\forall n$.

Exe: Suppose $\{A_n\}_{n \geq 1}$ be a seq of decreasing events, then
$A_1 \supseteq A_2 \supseteq A_3 \supseteq \cdots$

$$P\left(\bigcap_{n=1}^{\infty} A_n\right) = \lim_{n \to \infty} P(A_n)$$

Proof: Here $1 - P\left(\bigcap_{n=1}^{\infty} A_n\right) = P\left(\left(\bigcap_{n=1}^{\infty} A_n\right)^c\right)$

$= P\left(\bigcup_{n=1}^{\infty} A_n^c\right)$

now $A_1^c \subseteq A_2^c \subseteq \cdots$

$\therefore \lim_{n \to \infty} P(A_n^c) = P\left(\bigcup_{n=1}^{\infty} A_n^c\right)$

$\Rightarrow \lim_{n \to \infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right)$

Exercise : Show that $P(A|B) = P(B|A)\dfrac{P(A)}{P(B)}$ whenever $P(A)P(B) \neq 0$

as $P(A)P(B) \neq 0$

$P(A \cap B) \neq 0$

$\therefore \quad \dfrac{P(A\cap B)}{P(A)} = P(B|A)$

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

$\Rightarrow \quad P(B|A)\,P(A) = P(A \cap B) = P(A|B)\,P(B)$

14th Aug :

conditional probability: Suppose $P(B)>0$

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

Better way to write it

$A_1, A_2, \ldots, A_n \in \mathcal{F}$, $A_i \cap A_j = \phi$ for $i \neq j$

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

Then $P(B) = \sum_{i=1}^{n} P(B|A_i)\,P(A_i)$

• $(\Omega, \mathcal{F}, P)$ $B \in \mathcal{F}$ and $P(B)>0$

define $\tilde{P} : \mathcal{F} \to [0,1]$

$$\tilde{P}(A) = P(A|B)$$

Note $\tilde{P}$ is a prob. map on $\mathcal{F}$.

Baye's formula: Let $A_1, A_2, \ldots, A_n \in \mathcal{F}$ and $A_i \cap A_j = \phi$ for $i \neq j$ and $\bigcup_{i=1}^{n} A_i = \Omega$.

Suppose $P(A_i)>0$, $\forall i$

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

Proof: $$P(A_j|B) = \frac{P(A_j \cap B)}{P(B)} = \frac{P(B|A_j)P(A_j)}{\sum_{i=1}^{n} P(B|A_i)\,P(A_i)}$$

Ex: A box contains b black, r red balls, one of the balls drawn at random and note the colours, and put the balls along with c additional balls of the same colour. Now we again draw a ball. Show that the prob that the first drawn ball was black, given that second drawn ball is red.
is $b/b+r+c$

what is the prob. that

$$P(\text{first Black} \mid \text{second red})$$

$$= \frac{P(\text{second red} \mid \text{first Black}) \times P(\text{first Black})}{\ldots}$$

$$= \frac{\left(\frac{r}{r+c+b}\right) \times \left(\frac{b}{r+b}\right)}{P(\text{second red} \mid \text{first Black}) \times P(\text{first Black}) + P(\text{second red} \mid \text{first red}) \times P(\text{first Red})}$$

$$= \frac{\left(\frac{r}{\cancel{r+c+b}}\right) \times \left(\frac{b}{\cancel{r+b}}\right)}{\left(\frac{r}{\cancel{r+c+b}}\right) \times \left(\frac{b}{\cancel{r+b}}\right) + \left(\frac{r+c}{\cancel{r+c+b}}\right) \times \left(\frac{r}{\cancel{r+b}}\right)}$$

P(last is Red)

$$\overbrace{P(B)}^{B} = P(B|A)P(A) + P(B|A^c)P(A^c)$$

A: first draw is red

$$= \frac{(r+c)r + rb}{(r+c+b)(r+b)}$$

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

$$= \frac{\cancel{\gamma} b}{\cancel{\gamma} b + \gamma(r+c)} = \frac{b}{r+c+b}$$

Random variables:

let $(\Omega, \mathcal{F}, P)$ be a prob. space.

A random variable X is a map $\Omega \to \mathbb{R}$ s.t. $\{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F}, \forall x \in \mathbb{R}$

for all / for x in $\mathbb{R}$. fixed

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

$\{\omega \in \Omega \mid X(\omega) \leq x\} \in \mathcal{F} \;\; \forall x \in \mathbb{R}$ (technical condition)

Remark: we use capital letters to denote random variables X, Y, Z etc and small letters to denote its value/realisation

Example: $(\Omega, \mathcal{F}, P)$ $X: \Omega \to \mathbb{R}$

$X(\omega) = 10, \forall \omega \in \Omega$

$\{\omega \mid x(\omega) \leq 1\} = \phi \in \mathcal{F}$ $\Big\}$ Any $x \in \mathbb{R}$ s.t $x < 10$ then $\{\omega \mid X(\omega) \leq x\} = \phi \in \mathcal{F}$

for $x \geq 10$

$\{\omega \mid X(\omega) \leq x\} = \Omega \in \mathcal{F}$

so X is random variable.

Example: $\underbrace{\Omega = \{H, T\} \quad \mathcal{F} = \mathbb{P}(\Omega) \quad P(\{H\}) = p, \; P(\{T\}) = 1-p}_{(\Omega, \mathcal{F}, P)}$

$X: \Omega \to \mathbb{R}$
$X(H) = 1$
$X(T) = 0$ $\Big\}$ Random variable verification

$\{\omega : X(\omega) \leq 1.5\} = \{H, T\} \in \mathcal{F}$

$\{\omega : X(\omega) \leq -1\} = \phi \in \mathcal{F}$

$\{\omega : X(\omega) \leq 0.9\} = \{T\} \in \mathcal{F}$

$$\{\omega \mid X(\omega) \leq x\} = \begin{cases} \phi & ; \text{ if } x < 0 \\ \{T\} & ; \text{ if } 0 \leq x < 1 \\ \{H, T\} & ; \text{ if } x \geq 1 \end{cases} \qquad F(x) = \begin{cases} 0 & ; \; x < 0 \\ 1/2 & ; \; 0 \leq x < 1 \\ 1 & ; \; x \geq 1 \end{cases}$$

Example: let $(\Omega, \mathcal{F}, P)$ be a probability space and $A \in \mathcal{F}$ and $A \neq \phi$

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

$$X(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ -1 & \text{if } \omega \notin A \end{cases}$$

$$\{\omega \mid X(\omega) \leq x\} = \begin{cases} \phi ; & x < -1 \\ A^c ; & -1 \leq x < 1 \\ \Omega ; & x \geq 1 \end{cases}$$

$$F(x) = \begin{cases} 0 ; & x < -1 \\ P(A^c) ; & -1 \leq x < 1 \\ 1 ; & x \geq 1 \end{cases}$$

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

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

$X : \Omega \to \mathbb{R} \qquad X(\omega) = \omega$

$$\{\omega \in \Omega \mid X(\omega) \leq x\} = \begin{cases} \phi ; & x < 1 \\ \{1\} ; & 1 \leq x < 2 \\ \{1,2\} ; & 2 \leq x < 3 \\ \{1,2,3\} ; & 3 \leq x < 4 \\ \{1,2,3,4\} ; & 4 \leq x \end{cases}$$

as $\{1\} \notin \mathcal{F}$

X is not a random variable.

Remark: ① we want to bring abstract outcomes of sample points to a common setup which is familiar to us

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

② often we are interested to study a specific property related to a random experiment outcome. For that we associate an appropriate random variable

Example: Toss a coin n times

$$\Omega = \{(a_1, a_2, \dots a_n) \mid a_i = H \text{ or } T\}$$

$$\mathcal{F} = \mathcal{P}(\Omega)$$

we are interested in the total number of heads.

$$X((a_1, a_2, \dots a_n)) = \sum_{i=1}^{n} 1_{\{H\}}(a_i)$$

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

$$1_{\{H\}}(a_i) = \begin{cases} 1 & \text{if } a_i = H \\ 0 & \text{if } a_i = T \end{cases}$$

③ importance of technical condition: we want to keep track of prob. of situation related to X.

we are interested to calculate $X(\omega) = 100$

↳ we want to keep track of these

To calculate $P(\{\omega \mid X(\omega) \leq 10\})$ we need technical condition.

(cumulative)
Distribution function : Let $(\Omega, \mathcal{F}, P)$ be a prob. space and $X: \Omega \to \mathbb{R}$ be a random variable.

A Df: $F: \mathbb{R} \to [0,1]$ is called the dist function of X.

$$F(x) = P(X \leq x)$$
$$= P(\underbrace{\{\omega \mid X(\omega) \leq x\}}_{\in \mathcal{F}})$$

Exercise: write down dist fns of random variable we discussed

Example: $(\Omega, \mathcal{F}, P)$

$X: \Omega \to \mathbb{R}$
$X(\omega) = 10 \quad \forall \omega \in \Omega$

$$\{\omega \in \Omega \mid X(\omega) \leq x\} = \begin{cases} \phi ; & x < 10 \\ \Omega ; & 10 \leq x \end{cases}$$

$$F(x) = P\left(\{\omega \in \Omega \mid X(\omega) \leq x\}\right)$$
$$= P\left(\begin{cases} \phi ; & x < 10 \\ \Omega ; & 10 \leq x \end{cases}\right)$$
$$= \begin{cases} 0 & ; x < 10 \\ 1 & ; 10 \leq x \end{cases}$$

Property: ① If $x < y$, then $F(x) \leq F(y)$

proof: $F(y) = P(\{\omega \mid X(\omega) \leq y\})$
$F(x) = P(\{\omega \mid X(\omega) \leq x\})$

$\{\omega \mid X(\omega) \leq x\} \subseteq \{\omega \mid X(\omega) \leq y\}$

for $A \subseteq B$
$P(A) \leq P(B)$

$\Rightarrow P(\{\omega \mid X(\omega) \leq x\}) \leq P(\{\omega \mid X(\omega) \leq y\})$

$\Rightarrow F(x) \leq F(y)$

② $P(a < x \leq b) = P(\{\omega \mid a < X(\omega) \leq b\}) = F(b) - F(a)$

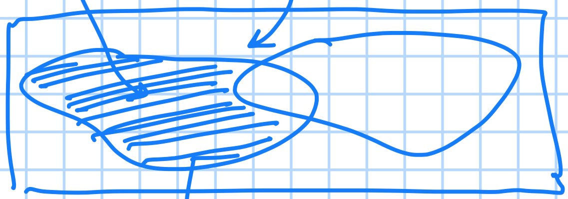
$\{\omega \mid a < X(\omega) \leq b\} = \{\omega \mid X(\omega) \leq b\} \cap \underbrace{\{\omega \mid X(\omega) > a\}}_{\text{in } \mathcal{F}}$

as it's comp. in $\mathcal{F}$

here $P(a < x \leq b) = P(\{\omega \mid a < X(\omega) \leq b\})$

$= P(\{\omega \mid x(\omega) \leq b\} \cap \{\omega \mid x(\omega) > a\})$

$= P(\{\omega \mid x(\omega) \leq b\} \cap \{\omega \mid x(\omega) \leq a\}^c)$



B A $P(B - A) = P(B \cap A^c)$

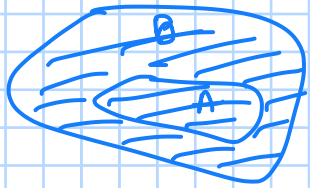
$P(\{\omega \mid x(\omega) \leq b\} - \{\omega \mid x(\omega) \leq a\})$

$= P(\{\omega \mid x(\omega) \leq b\}) - P(\{\omega \mid x(\omega) \leq a\})$

$F(b) = P(X \leq b)$

$= P(\{\omega \mid x(\omega) \leq b\})$ but as $a < b$

$F(a) = P(\{\omega \mid x(\omega) \leq a\})$



$F(b) - F(a)$

21st Aug :

Random variables :

A f^n $X: \Omega \to \mathbb{R}$ is called random variable if $\{\omega : X(\omega) \leq x\} \in \mathcal{F}, \forall x \in \mathbb{R}$

Equivalently, X is a random variable

$$\{\omega \mid X(\omega) > x\} \in \mathcal{F}, \text{ for all } x \in \mathbb{R}$$

Remark : ① It is clear from the def^n that whether a f^n $X: \Omega \to \mathbb{R}$ is a random variable or not depends on choice of σ-field $\mathcal{F}$ on Ω.

② Constant f^n on Ω is a random variable, it does not depend on choice of $\mathcal{F}$.

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

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

Define $X: \Omega \to \mathbb{R}$
$X(\omega) = \omega + 1$

X is not a random variable as:
for $x = 3$

$$\{\omega \in \Omega \mid X(\omega) \leq x\}$$

$$\{\omega \in \Omega \mid \omega + 1 \leq 3\}$$

$$= \{\omega \in \Omega \mid \omega \leq 2\} = \{1, 2\} \notin \mathcal{F}$$

$\therefore$ Not a random variable

Example : non const f^n $= Y$ s.t $Y: \Omega \to \mathbb{R}$ is a random variable for

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

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

$$Y(\omega) = \begin{cases} 1 & ; \omega = 1 \\ 2 & ; \omega = 2, 3, 4 \end{cases}$$

$$\{\omega \in \Omega \mid Y(\omega) \leq x\} = \begin{cases} \phi & ; x \in (-\infty, 1) \\ \{1\} & ; x \in [1, 2) \\ \underbrace{\{1,2,3,4\}}_{=\Omega} & ; x \in [2, \infty) \end{cases}$$

$\therefore Y$ is a random variable.

Note: This is true for all $Y(\omega) = c_1$ for $\omega = 1$
$Y(\omega) = c_2$ for $\omega = 2, 3, 4$
for $c_1 \neq c_2$

Exercise: Prove for c_1, c_2 above this errors, check that there is another function is only one.
↳ non-const.
→ see down

Exercise: Let X be a random variable on $(\Omega, \mathcal{F}, P)$. Suppose X takes values in $\{a_1, a_2, \ldots a_k\}$.

Show that $\{\omega \mid X(\omega) = a_i\} = \{X = a_i\} \in \mathcal{F}$

wlog assume $a_1 < a_2 < \cdots a_k$

$$\{\omega \mid X(\omega) = a_1\} = \{\omega \mid X(\omega) \leq a_1\} \in \mathcal{F}$$

$$\{\omega \mid X(\omega) = a_2\} = \{\omega \mid X(\omega) \leq a_2\} \setminus \{\omega \mid X(\omega) \leq a_1\} \in \mathcal{F}$$

$$\{X = a_i\} = \{X \leq a_i\} \setminus \{X \leq a_{i-1}\} \in \mathcal{F}$$

so, $\{X = a_i\} \in \mathcal{F}$ for $i = 1, 2, \ldots k$

observe: $\bigcup_{i=1}^{k} \{X = a_i\} = \Omega$

$$\{X = a_i\} \cap \{X = a_j\} = \phi \quad \text{for } i \neq j$$

Exercise: Suppose X is a function on Ω and X takes value in $\{a_1, \ldots a_k\}$. Suppose $\{\omega \mid X(\omega) = a_i\} \in \mathcal{F}$ for $i = 1, 2, \ldots k$. Is X a random variable? (converse of above)
↓ wrt to given $\mathcal{F}$.

wlog $a_1 < a_2 < \cdots a_k$

$$\{\omega \mid X(\omega) \leq x\} = \phi \quad \text{if } x < a_1$$

$$\{\omega \mid X(\omega) \leq x\} = \{\omega \mid X(\omega) = a_1\} \quad \text{if } x \in [a_1, a_2)$$

now $\{\omega \mid X(\omega) \leq x\}$ where $x \in [a_1, a_{i-1})$

$$= \bigcup_{j=1}^{i-1} \{\omega \mid X(\omega) = a_j\} \in \mathcal{F} \quad \text{as } \{\omega \mid X(\omega) = a_j\} \in \mathcal{F}$$

$\therefore X$ is a random variable.

Remark: last two eq. tells us the following:
Let X be a function on Ω which takes values in $\{a_1, \ldots a_k\}$ then X is a random variable $\Leftrightarrow \{X = a_i\} \in \mathcal{F}$ for $i = 1, 2, \ldots k$

Exercise: Let $X : \Omega \to \mathbb{R}$ be a function and X takes values in $\{a_1, \ldots a_k\}$. Find the smallest σ-field $\mathcal{F}$ on Ω with respect to which X is a random variable.

Let $\mathcal{G}$ be the smallest σ-field $\{X = a_i\} \in \mathcal{G}$ $\forall i = 1, 2, \ldots k$

Let $B_i = \{X = a_i\} = \{X \leq a_i\} \setminus \{X \leq a_{i-1}\}$

observe $B_1, B_2, \ldots B_k$ forms a partition of Ω.

$$\mathcal{G} = \left\{ \bigcup_{j=J} B_j : J \subseteq \{1,2,\dots K\} \right\}$$

Exercise: $\Omega = \{-2,-1,0,1,2\}$
$X(\Omega) = \omega^2$
Find the smallest σ-field with respect to which X is a r.v.

$$\mathcal{G} = \{\phi, \Omega, \{0\}, \{-1,1\}, \{-2,2\} \dots \}$$

let $B_1 = \{0\}$
$B_2 = \{-1,1\}$
$B_3 = \{-2,2\}$

$$\mathcal{G} = \left\{ \bigcup_{k \in K} B_k \;\middle|\; K \subseteq \{1,2,3\} \right\}$$ as 3, total $2^3 = 8$

$$= \{ \phi, \Omega, \{0\}, \{-1,1\}, \{-2,2\}, \{0,-1,1\}, \{0,-2,2\}, \{-1,-2,2,-1\}\}$$

Infer: ① A f^n $X : \Omega \to \mathbb{R}$ will be a random variable or not depends on the choice of the $\mathcal{F}$.

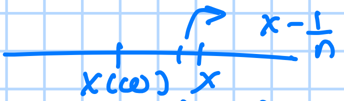
② If X takes finite values then
X is a random variable $\Leftrightarrow \{X = a_i\} \in \mathcal{F}$, $i = 1,2,\dots k$

③ If X takes finitely many values, then we know the smallest σ-field w.r.t which X is a random variable.

④ const f^n is always a random variable.

⑤ If $\mathcal{F} = P(\Omega)$, then any f^n on Ω is a random variable.

exercise: let X be a r.v on $(\Omega, \mathcal{F})$. Show that $\{\omega \mid X(\omega) < x\} \in \mathcal{F}$ for all $x \in \mathbb{R}$.



$\exists\, n \in \mathbb{N}$ s.t. Archemedian Property
$X(\omega) \leq x - \frac{1}{n}$

Claim: $$\{\omega \mid X(\omega) < x\} = \bigcup_{n=1}^{\infty} \left\{\omega \mid X(\omega) \leq x - \frac{1}{n}\right\}$$

as $\left\{\omega \mid X(\omega) \leq x - \frac{1}{n}\right\} \in \mathcal{F}$

$$\Rightarrow \bigcup_{n=1}^{\infty} \left\{\omega \mid X(\omega) \leq x - \frac{1}{n}\right\} \in \mathcal{F}$$
as $\mathcal{F}$ is σ-field

$$\Rightarrow \{\omega \mid X(\omega) < x\} \in \mathcal{F}$$

proof of claim: if $y \in \{\omega \mid X(\omega) < x\}$ then
$X(y) < x$, and $x - X(y) > 0$
$\exists\, n_0 \in \mathbb{N}$ s.t
$\frac{1}{n_0} < x - X(y)$

then $y \in \left\{\omega \mid X(\omega) \leq x - \frac{1}{n_0}\right\}$

$$\therefore \left\{ \omega \,\middle|\, X(\omega) < x \right\} \subseteq \bigcup_{i=1}^{\infty} \left\{ \omega \,\middle|\, X(\omega) \le x - \frac{1}{i} \right\}$$

now if $y \in \bigcup_{i=1}^{\infty} \left\{ \omega \,\middle|\, X(\omega) \le x - \frac{1}{i} \right\}$

$$\Rightarrow y \in \bigcup_{i=1}^{\infty} \left\{ \omega \,\middle|\, X(\omega) < x \right\}$$

(trivial)

$$\therefore \left\{ \omega \,\middle|\, X(\omega) < x \right\} = \bigcup_{n=1}^{\infty} \left\{ \omega \,\middle|\, X(\omega) \le x - \frac{1}{n} \right\}$$

Distribution function: $F : \mathbb{R} \to [0,1]$

$F(x) = P(X \le x)$

properties:

(i) If $x < y$, then $F(x) \le F(y)$

(ii) $P(x < X \le y) = F(y) - F(x)$

(iii) F is right-continuous

$$\lim_{h \downarrow 0} F(x+h) = F(x)$$

(iv) left limit of F exist

$$\lim_{h \downarrow 0} F(x-h) \text{ exist}$$

To prove: $\lim_{h \downarrow 0} F(x+h) = F(x)$, for any $x \in \mathbb{R}$

proof: let $\{h_n\}_{n \ge 1}$ be a seq of positive real numbers such that $h_n \downarrow 0$.

$$\lim_{n \to \infty} F(x+h_n) = \lim_{n \to \infty} P(\underbrace{X \le x + h_n}_{A_n \supseteq A_{n+1}})$$

$$h_1 \ge h_2 \ge h_3 \cdots$$

$$x + h_1 \ge x + h_2 \ge \cdots$$

$$\{X \le x + h_n\} \subseteq \{X \le x + h_{n-1}\}$$

by definition: $= \lim_{n \to \infty} P(A_n) = P\left(\bigcap_{n=1}^{\infty} A_n\right)$

claim: $\bigcap_{n=1}^{\infty} A_n = \bigcap_{n=1}^{\infty} \left\{ \omega \,\middle|\, X(\omega) \le x + h_n \right\}$

$$= \left\{ \omega \,\middle|\, X(\omega) \le x \right\}$$

$$\therefore \lim_{n \to \infty} P\left(\bigcap_{n=1}^{\infty} A_n\right) = P(A) = P\left(\left\{ \omega \,\middle|\, X(\omega) \le x \right\}\right)$$

proof of claim: let $y \in \left\{ \omega \,\middle|\, X(\omega) \le x \right\}$ then

$$X(y) \le x \Rightarrow X(y) \le x + h_n, \forall n$$

$\therefore y \in \bigcap_{n=1}^{\infty} \{\omega \mid x(\omega) \leq x + h_n\}$

$\therefore \{\omega \mid x(\omega) \leq x\} \subseteq \bigcap_{n=1}^{\infty} \{\omega \mid x(\omega) \leq x + h_n\}$

now if $y \in \bigcap_{n=1}^{\infty} \{\omega \mid X(\omega) \leq x + h_n\}$ and as $\{h_n\}$ is a seq decreasing to 0. $x(y) \leq x + h_n, \forall n$

$\Rightarrow X(y) < x$ (from prev)

$\Rightarrow X(y) \leq x$

$\therefore \bigcap_{n=1}^{\infty} \{\omega \mid X(\omega) \leq x + h_n\} \subseteq \{\omega \mid X(\omega) \leq x\}$

$\therefore \bigcap_{n=1}^{\infty} \{\omega \mid X(\omega) \leq x + h_n\} = \{\omega \mid X(\omega) \leq x\}$

Exercise: Prove for c_1, c_2 above this occurs, check that there is another function is only one.
↳ non-const.

as $\mathcal{F}$ is finite, $\bigcup_{i=1}^{n} \{X = x_i\} = \Omega$

and

$\{X = x_i\} \cap \{X = x_j\} = \phi \quad i \neq j$

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

only for two sets.

$\therefore$ only this.

s.t $\{X = x_i\} = \{1\}$
(wlog)
$\{X = x_j\} = \{2,3,4\}$

Quiz-1: upto random variables
problem set upto problem 4, set 2.
3 questions, 10 marks.

23rd Aug :

Distribution f^n: let X be a r.v defined on a prob. space $(\Omega, \mathcal{F}, P)$. Distribution f^n of X is defined as $F(x) = P(X \leq x)$, $x \in \mathbb{R}$

(i) F is right cont.

$$\lim_{h \downarrow 0} F(x+h) = F(x)$$

It is enough to prove that if $h_n \downarrow 0$ as $n \to \infty$, then $\lim_{n\to\infty} F(x+h_n) = F(x)$

so,

$$\lim_{n\to\infty} F(x+h_n) = \lim_{n\to\infty} P(\{\omega \mid X(\omega) \leq x + h_n\})$$

since $h_n \downarrow 0$,

$A_n = \{\omega \mid X(\omega) \leq x + h_n\}$ is a decreasing seq of events

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

now $\lim_{n\to\infty} P(A_n) = P(\bigcap_{n=1}^{\infty} A_n)$ ($\because$ cont. of probability)

claim: $\bigcap_{n=1}^{\infty} A_n = \{\omega \mid X(\omega) \leq x\} = A$

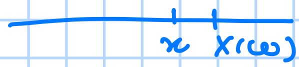
$A \subseteq A_k \ \forall k$

$\Rightarrow A \subseteq \bigcap_{n=1}^{\infty} A_n$

To prove: $\bigcap_{n=1}^{\infty} A_n \subseteq A$ (for the above claim)

proof: suppose not true, then $\exists\, \omega \in \bigcap_{n=1}^{\infty} A_n$ s.t

$X(\omega) > x$



then we can find an $N \in \mathbb{N}$ s.t
$X(\omega) > x + h_N$ (this is possible as $h_N \downarrow 0$)

this contradicts the fact that $\omega \in A_N$.

$\therefore \bigcap_{n=1}^{\infty} A_n \subseteq A$

now with this $P(\bigcap_{n=1}^{\infty} A_n) = P(A) = F(x)$

(ii) left limit of F exist $\lim_{h\downarrow 0} F(x-h)$ exist

proof:

enough to prove that if $h_n \downarrow 0$ then $\lim_{n\to\infty} F(x-h_n)$ exist.

$$\lim_{n\to\infty} F(x-h_n) = \lim_{n\to\infty} P(X \leq x - h_n) = P(\bigcup_{n=1}^{\infty} A_n)$$

as $h_n \downarrow 0$

$h_1 > h_2 > h_3 \cdots$

$A_n = \{\omega \mid X(\omega) \leq x - h_n\}$, then $A_n \uparrow$.

now, $\bigcup_{n=1}^{\infty} A_n = \{\omega \mid X(\omega) < x\}$ $\quad x - h_1 < x - h_2 < \cdots\cdots$

$A_1 \subseteq A_2 \subseteq A_3 \cdots\cdots$

to prove this, let $\omega' \in \{\omega \mid X(\omega) \leq x - h_n\}$

then as $x > x - h_n$ (because $h_n > 0$ $\forall n \in \mathbb{N}$)

$$\omega' \in \{\omega \mid X(\omega) \le x - h_n < x\}$$
$$\Rightarrow \quad \omega' \in \{\omega \mid X(\omega) < x\}$$
$$\therefore \quad \text{if } \omega' \in \bigcup_{n=1}^{\infty} A_n \Rightarrow \omega' \in \{\omega \mid X(\omega) < x\}$$

now, if $\omega' \in \left\{\omega \mid X(\omega) < x\right\}$ then $X(\omega') < x$

$x - h_n$

$X(\omega')$ x

$$x - X(\omega') > 0$$
$$\exists\ h_n \in \mathbb{N} \text{ s.t}$$
$$x - X(\omega') \ge h_n$$
$$\Rightarrow X(\omega') \le x - h_n$$
$$\therefore \quad \omega' \in \bigcup_{n=1}^{\infty} A_n$$
$$\therefore \bigcup_{n=1}^{\infty} A_n = \left\{\omega \mid X(\omega) < x\right\}$$

(iii) suppose you know the distribution of a r.v X

Right limit, left limit

$$P(X=x) = P(X \le x) - P(X < x)$$
$$= F(x) - \lim_{y \uparrow x} F(y)$$

Discrete random variable:

A random variable X is called discrete r.v if it takes values in a (atmost) coutable subsets of $\mathbb{R}$.

or

$P(X \in C) = 1$

prob that X is in a coutable set

Probability mass function: (p.m.f)

$$P : \mathbb{R} \to [0,1]$$
$$\text{s.t} \quad P(x) = P(X=x)$$

suppose $C = \{x_1, x_2, \ldots\}$

$P(x) = 0$ if $x \notin C$

$P(x) \ge 0$ if $x \in C$

$$\sum_{i=1}^{\infty} P(x_i) = \sum_{i=1}^{\infty} P(X = x_i)$$
$$= P\left(\bigcup_{i=1}^{\infty} \{X = x_i\}\right)$$
$$= P(X \in C) = P(\Omega) = 1$$

Exercise: Suppose $F : \mathbb{R} \to [0,1]$ and C is a coutable Subset of $\mathbb{R}$ and

$$\left.\begin{array}{l} f(x) = 0 \quad \text{if } x \notin C \\ f(x) \ge 0 \quad \text{if } x \in C \\ \text{and} \quad \sum_{x \in C} f(x) = 1 \end{array}\right\} (*)$$

Is f a p.m.f ?

$\omega \leftarrow$ coutable as

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

and $P : \mathbb{R} \to [0,1] \leftarrow$ coutable

$\therefore$ domain also coutable.

Let $\Omega = \mathbb{N}$

$\mathcal{F} = P(\mathbb{N})$

define:

let $P(\{1\}) = f(x_1)$ ← defination of P. $(P: \mathcal{F} \to [0,1])$

$P(\{i\}) = f(x_i)$

$$P(\Omega) = P(\mathbb{N}) = \sum P(\{i\})$$
$$= \sum f(x_0)$$
$$= 1$$

$(\Omega, \mathcal{F}, P) \to$ prob space done

$$P(B) = \sum_{i \in B} f(\{i\}) \quad B \subseteq \mathbb{N}$$

now, a random variable: $X : \underset{\mathbb{N}}{\Omega} \to \mathbb{R}$ for $n \in \mathbb{N}$

$$X(n) = x_n$$

$$p(x) = P(X = x)$$

let $x \notin C$

$$P(x) = P(X = x) = P(\phi) = 0$$

$x \in C$

$$p(x) = P(X = x) = P(\{\omega \mid X(\omega) = x_i\})$$
$$= f(x_i)$$

Conclusion of exercise:

If a function f which satisfies ⊛ is a probability mass fn of a random variable X.

Example: ① Bernoulli (P).

X is called Bernoulli (P),

$P(X=1) = P$

$P(X=0) = 1-P$

where $0 < P < 1$.

eg: $T \to 0 = X$

$H \to 1 = X$

② Binomial (n, P):

$$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$k = 0, 1, 2, \ldots n$

where $0 < P < 1$

eg: Tossing a coin n-times : $P(\{H\}) = p$

$P(\text{no of } H = k \text{ out of } n) = \binom{n}{k} (p)^k (1-p)^{n-k}$ ← P that tails will be n-k times

↑ no. of ways of choosing k from n

← P that heads will be k times

if $P(\{H\}) = \frac{1}{2}$ then $P(\text{no of heads} = k) = \binom{n}{k} \frac{1}{2^n}$

$$= \binom{n}{k} \frac{1}{2^k} \frac{1}{2^{n-k}}$$
$$= \binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k}$$

③ Poisson (λ) :

$$P(X=k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k=0,1,\ldots\ldots \quad \text{and } \lambda>0$$

eg:

$$\binom{n}{k} p^k (1-p)^k \qquad n\to\infty \quad p\to 0 \quad np\to\lambda>0$$

$$\lim_{\substack{n\to\infty \\ p\to 0 \\ np\to\lambda}} \binom{n}{k} p^k (1-p)^{n-k} = L$$

then $\log L = \log\binom{n}{k} + \log p^k + \log(1-p)^{n-k}$

$$\log L = \log\left(\frac{(n)(n-1)\cdots(n-k+1)}{k!}\right) + k\log p + n\log(1-p)$$

$$= \log\frac{1}{k!} + k\log(np) + n\log(1-p)$$

$$= \log\frac{\lambda^k}{k!} + n\log(1-p)$$

now $(1-p)^n = (1-p)^{\frac{\lambda}{p}} = \left((1-p)^{\frac{1}{p}}\right)^{\lambda} = e^{-\lambda}$

as $(1+p)^{\frac{1}{p}} = 1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots\cdots$

$= e$

$$= \log\frac{\lambda^k}{k!} + \log e^{-\lambda}$$

$$\log L = \log\frac{\lambda^k e^{-\lambda}}{k!}$$

$$L = \frac{\lambda^k e^{-\lambda}}{k!}$$

conclusion : If n is large and p is small, then Binomial prob. is close to poisson prob.

$X \sim \text{Binomial}(n, p)$
$Y \sim \text{Poi}(\lambda)$
n is large, p is small and $np=\lambda$

then $P(X=k) \approx P(Y=k)$

28^{th} Aug :

Discrete random variables

① Bernoulli (P)
② Binomial (n, P)
③ Poisson (λ)

Observation

If $n \to \infty$
$np \to \lambda$ ($\lambda > 0$)
then $\binom{n}{k} p^k (1-p)^{n-k} \to \frac{e^{-\lambda} \lambda^k}{k!}$

If n is large enough, p is small so that np moderate, then Binomial prob. is close to Poi(λ) where $\lambda = np$

Exercise: Suppose 1000 letters on a page of a book. The probability that a letter is miss printed is P. What is the probability that there are 100 missing points.

A letter is missprinted or not does not depend on the letters.

$$P(X=k) = \binom{n}{k} p^k (1-p)^k = \binom{1000}{100} (p)^{100} (1-p)^{900}$$

$$\sim e^{-1000p} \frac{(1000p)^{100}}{100!} = \frac{e^{-\lambda} \lambda^{100}}{(100)!} \quad \text{where } \lambda = 1000p$$

Exercise: Suppose we dont know the number of letters on a page. We want to find prob. that there are k many misprints. (λ given)

This follow from $(e^{-\lambda}) \frac{\lambda^k}{k!}$

Examples : (Poisson) : (i) No of misprints on a given page

Poisson dist is used
(ii) No of wrong telephone number dialed in a day.
(iii) No of customers entering in a store on a given day
(iv) No of particles dislodged in a fixed time period from a radioactive material.

One more random variable : ④ geometric : X values in $\{1, 2, 3, \ldots\}$

$P(X=k) = (1-p)^{k-1} p$ — Prob. mass function

p lies b/w 0 and 1

Example : (geometric) We toss a coin until we get an H.

Possible number of tossing $= \{1, 2, \ldots\}$

$P(k\text{-tosses are required to get H}) = (1-p)^{k-1} (p)^k$

In genral if we identify an event/outcome of an experiment as sucess then the number of trials required to get the 1st sucess follows geometric distribution.

In coin tossing we identify appearance of H as a sucess.

⑤ Negative binomial : X takes values in $\{r, r+1, r+2, \ldots\}$ where r is $\in \mathbb{N}$ is fixed.

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

$$P(X=k) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \text{ where } k = r, r+1, \ldots$$

Example : (Neg Binomial)

Tossing a coin until we get r-th Head.

$\{r, r+1, \ldots\} \to$ as atleast r times

$$P(k \text{ tosses req}) = \binom{k-1}{r-1} p^r (1-p)^{k-r} \qquad P(H) = p$$

$k-r$ tails

r many heads

out of first $k-1$, $r-1$ should be head, k^{th} — Head

In general min. number of trials required to get r-th success (where r is a fixed natural number) follows negative binomial (p, r) where p is the success prob.

Exercise - check problems/examples from Ross's book to get idea when and how these distributions are used to calculate probabilities.
do

Random vectors :

Let $(\Omega, \mathcal{F}, P)$ be a prob space. A map $X : \Omega \to \mathbb{R}^d$ where $d \in \mathbb{N}$ is called a random variable for

$$\{\omega \mid X(\omega) \le x\} \in \mathcal{F} \quad , \text{for all } x \in \mathbb{R}^d$$

Suppose $X = (X_1, X_2, \ldots X_d)$
$x = (x_1, x_2, \ldots x_d)$

$\{X \le x\}$ means $\{X_i \le x_i \;, \text{for } 1 \le i \le d\}$

Note : we will discuss for $d = 2$, to keep notion basic.

Observations - ① Suppose $X : \Omega \to \mathbb{R}^2$ is a random vector.
$X = (X_1, X_2)$
then both X_1, X_2 are random variables.

Proof :
let $x_1 \in \mathbb{R}$

$$\{\omega \mid X_1(\omega) \le x_1\} \in \mathcal{F}$$

$$\{\omega \mid X_1(\omega) \le x_1\} = \bigcup_{n=1}^{\infty} \{\omega \mid (X_1(\omega), X_2(\omega)) \le (x_1, n)\}$$

$$= \bigcup_{n=1}^{\infty} \underbrace{\{\omega \mid (X_1, X_2)(\omega) \le (x_1, n)\}}_{\substack{\text{each one is in } \mathcal{F} \\ \text{as } X \text{ is a random variable}}} \in \mathcal{F}$$

similarly X_2 is a random variable.

② Suppose X, Y are two random variables defined on $(\Omega, \mathcal{F}, P)$ then (X, Y) is a random vector. Trivial

proof :
$(X, Y) : \Omega \to \mathbb{R}^2$

$$\{\omega \mid (X, Y) \le (x, y)\} = \underbrace{\{\omega \mid X(\omega) \le x\}}_{\in \mathcal{F}} \cap \underbrace{\{\omega \mid Y(\omega) \le x\}}_{\in \mathcal{F}}$$

$$\Rightarrow \{\omega \mid (X, Y) \le (x, y)\} \in \mathcal{F}$$

Conclusion: If $X = (X_1, X_2, \ldots X_d)$ is a random variable iff each component is a random variable.

Def^n : Distribution f^n of a random vector $X = (X_1, X_2) : \Omega \to \mathbb{R}^2$ is defined as $F(x_1, x_2) = P((X_1, X_2) \leq (x_1, x_2))$

Discrete random vector (of dim 2):

let $X : \Omega \to \mathbb{R}^2$ be a random vector.
It is called discrete random vector if it takes values on atmost countable set S of $\mathbb{R}^2$.

Observe : ① If $X = (X_1, X_2)$ is a discrete random vector, then X_1, X_2 are discrete random variables.

② suppose X_1, X_2 are two discrete random variables defined on $(\Omega, \mathcal{F}, P)$, is $X = (X_1, X_2)$ a discrete random vector → yes

Proof :
Suppose X_1, X_2 takes values in countable sets S_1, S_2 respectively then $X = (X_1, X_2)$ takes values in $S_1 \times S_2$

$$S_1 \times S_2 = \{(x, y) \mid x \in S_1, y \in S_2\} \text{ is a countable set.}$$

So X is a discrete random vector.

conclusion - $X = (X_1, X_2, \ldots X_d)$ is a discrete random variable iff each comp. is a disc random variable.

Suppose (X, Y) is a discrete random vector of dim 2.

Suppose (X, Y) takes values in a countable set $S = \{(x_i, y_j) \mid i, j \in \mathbb{N}\}$

Probability mass f^n of (X, Y) :

$$p(x, y) = P(X = x, Y = y)$$

Observe : ① $p(x, y) = 0$, if $x \neq x_i$ or $y \neq y_j$
② $p(x_i, y_j) \geq 0$ for $(x_i, y_j) \in S$
$\sum_{i, j \in \mathbb{N}} p(x_i, y_j) = 1$

30th Aug :

Discrete random vector

Suppose $X=(X_1, X_2)$ is a discrete random vector taking values in

$$S=\{(x_i, y_j) ; i,j \in \mathbb{N}\}$$

prob. mass fn of X is given by

$$P: \mathbb{R}^2 \to [0,1]$$
$$P(x,y)=P(X_1=x, X_2=y)$$

Note : $P(x,y)=0$ if $x \neq x_i$ or $y \neq y_i$
$P(x_i, y_i) \geq 0$ for $i,j \in \mathbb{N}$

Sometimes, the p.m.f of X is called joint p.m.f.

Exercise : Can we get the p.m.f of X_1 from the p.m.f of X?
let's denote p.m.f of X_1 by P_{X_1}

$$P_{X_1}(x) = P(X_1=x) = P(X_1=x, X_2 \in \mathbb{R})$$
$$= P(\{\omega \mid X_1(\omega)=x\})$$
$$= P(\{\omega \mid X_1(\omega)=x\} \cap \underbrace{\{\omega \mid X_2(\omega) \in \mathbb{R}\}}_{= \Omega})$$
$$= P(A \cap \Omega) = P(A)$$
$$= P(X_1=x, X_2=y_j, \forall j \in \mathbb{N})$$
$$= P((X_1, X_2)=(x, y_j), j \in \mathbb{N})$$
$$= P\Big(\bigcup_{j \in \mathbb{N}} \{(X_1, X_2)=(x, y_j)\}\Big)$$
$$= \sum_{j \in \mathbb{N}} P(\{(X_1, X_2)=(x, y_j)\})$$
$$= \sum_{j \in \mathbb{N}} P(x, y_j)$$

Note :
$\{\omega \mid (X_1, X_2)(\omega)=(x, y_j)\}$ where $j \in \mathbb{N}$ are disjoint events.
$\{\omega \mid (X_1, X_2)(\omega)=(x, y_1)\}$
$\cap \{\omega \mid (X_1, X_2)(\omega)=(x, y_2)\} = \phi$
as $y_1 \neq y_2$
(countable disjoint)

If $x \neq x_i$, then $P_{X_1}(x) = \sum_{j \in \mathbb{N}} P(x, y_j) = 0$

$$P_{X_1}(x_i) = \sum_{j \in \mathbb{N}} P(x_i, y_j)$$

Note : $P_{X_2}(y) = \sum_{i \in \mathbb{N}} P(x_i, y)$

Similarly $\sum_{i \in \mathbb{N}} P_{X_1}(x_i) = \sum_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} P(x_i, y_j) = 1$

Exercise : Suppose X_1, X_2 are discrete random variables defined on same probability space $(\Omega, \mathcal{F}, P)$. Then (X_1, X_2) are random vector. Suppose P_{X_1}, P_{X_2} are p.m.fs of X_1 and X_2 respectively, we want to find the pmf of (X_1, X_2), can we find it?

$$P(x, y) = P(\{\omega \mid (X_1, X_2)(\omega)=(x,y)\})$$
$$= P((X_1, X_2)=(x,y))$$
$$= P(\underbrace{\{\omega \mid X_1(\omega)=x\}}_{A} \cap \underbrace{\{\omega \mid X_2(\omega)=y\}}_{B})$$

here $P(A \cap B) = P(A) + P(B) - P(A \cup B)$
or $= P(A \mid B)\, P(B)$

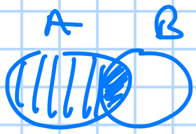
independent prob
$P(A \mid B) = P(A)$
$\frac{P(A \cap B)}{P(B)} = P(A)$ $\Rightarrow P(A \cap B) = P(A)P(B)$
} Both cases same ans

If both are independent then $= P(\{\omega \mid X_1(\omega) = x\})\, P(\{\omega \mid X_2(\omega) = y\})$

Independence: Two events (A, B) are independent if $P(A \cap B) = P(A)P(B)$

Exercise: The following are equivalent (TFAE)

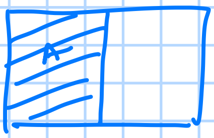
(i) A ind of B
(ii) A ind of B^c
(iii) B ind of A^c
(iv) A^c ind of B^c



Now, $P(A \cap B) + P(A \cap B^c) = P(A)$
$\Rightarrow P(A)P(B) + P(A \cap B^c) = P(A)$
$\Rightarrow P(A)P(B^c) = P(A \cap B^c)$
$\Rightarrow$ A and B^c are ind

A common mistake: A and B are independent if $A \cap B = \phi$

Exercise: $\Omega = \{1,2,3,4\}$ $\mathcal{F} = \mathcal{P}(\Omega)$ $P(\{i\}) = \frac{1}{4}$ for $i = 1,2,3,4$. Let $A = \{1,2\}$
List $B \in \mathcal{F}$ s.t A and B are ind.



$P(A) = \frac{1}{2}$ $P(A)P(B) = P(A \cap B) = \frac{1}{4}$
$P(B) = \frac{1}{2}$

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

Defn: Let $A_1, A_2, \ldots A_n \in \mathcal{F}$. We say $A_1, A_2, \ldots A_n$ are independent if
$$P\Big(\bigcap_{i \in I} A_i\Big) = \prod_{i \in I} P(A_i) \quad \forall I \subseteq \{1, 2, \ldots, n\}$$

Defn: $A_1, A_2, \ldots A_n$ are called pair wise independent if $P(A_i \cap A_j) = P(A_i) \cap P(A_j)$ for $i \neq j$

Remark: Independence of $A_1, A_2, \ldots A_n$ implies pairwise independence but converse is NOT TRUE.

Exercise: Suppose A, B, C are independent events. Show that
done down
(i) $A, B \cap C$ are ind
(ii) $A, B \cup C$ are ind
(iii) $A, B \setminus C$

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

Exercise: Find events A, B, C s.t they are pairwise independent but not independent
done down

Defn: Let $C \in \mathcal{F}$ and $P(C) > 0$. Events A, B are called conditionally independent given C if
$$P(A \cap B \mid C) = P(A \mid C)\, P(B \mid C)$$
There is a natural generalisation of this concept for a collection of events $A_1, A_2, \ldots A_n$

Now:

Suppose (X_1, X_2) is a discrete random vector.

$$P((X_1,X_2)=(x,y)) = P(\{X_1=x\}\cap\{X_2=y\})$$

$$= P(\{X_1=x\})\,P(\{X_2=y\}) \quad \forall\, x,y\in\mathbb{R}$$

as discrete random vector see defn - why this will happen.

this $= P(x_i, y_j)$

$= P_{X_1}(x_i)\,P_{X_2}(y_j) = P_{X_1}(x_i)\,P_{X_2}(y_j)$

Defn:

Suppose X_1, X_2 are discrete random variables. We say X_1, X_2 are independent if

$$P((X_1,X_2)=(x,y)) = P(X_1=x)P(X=y) \quad \forall\, x,y\in\mathbb{R}$$

conclusion: If X_1, X_2 are independent then we can find joint p.m.f from marginal p.m.f P_{X_1}, P_{X_2}.

for $X \to$ no of heads out of n

$Y \xrightarrow[n]{n}$ no of tails out of n } dependent

if we draw $N \sim Poi(\lambda)$

$X =$ No of H } from N

$Y =$ No of T

sometimes notion of independence is not consistent (Here no of coin toss is randomised)

Example:

Suppose X_1, X_2 are discrete random variables and independent. $g: \mathbb{R}\to\mathbb{R}$, $h: \mathbb{R}\to\mathbb{R}$

$Y_1 = g(X_1)$

$Y_1(\omega) = g(X_1(\omega))$, $\omega\in\Omega$

$Y_2 = h(X_2)$

Note: Range of Y_1, Y_2 are countable so Y_1, Y_2 are discrete random variable.

✓ Prove Y_1 and Y_2 are independent. (① random var ② dis ③ ind)

done see down

Exercise: Suppose A, B, C are independent events. Show that

(i) $A, B\cap C$ are ind

(ii) $A, B\cup C$ are ind

(iii) $A, B\setminus C$

(i) as $P(A\cap B\cap C) = P(A)P(B)P(C)$

$= P(A)\,P(B\cap C)$

(ii) as $P(B\cup C) = P(B) + P(C) - P(B)P(C)$

$P(A\cap(B\cup C)) = P((A\cap B)\cup P(A\cap C))$

$= P(A\cap B) + P(A\cap C) - P(A\cap B\cap C)$

$= P(A)P(B) + P(A)P(C) - P(A)P(B)P(C)$

$= P(A)P(B\cup C)$

(iii) $B\setminus C = B\cap C^c$ as $P(B)$ and $P(C^c)$ are ind

A, B, C^c ind

Exercise: Find events A, B, C s.t they are pairwise independent but not independent

let $A = \{1,2\}$
$B = \{2,3\}$
$C = \{3,4\}$
then $P(A \cap B) = \frac{1}{4} = P(A)P(B)$
$P(A \cap B \cap C) = \phi \neq P(A)P(B)P(C)$

Example:
Suppose X_1, X_2 are discrete random variables and independent. $g: \mathbb{R} \to \mathbb{R}$
$h: \mathbb{R} \to \mathbb{R}$

$Y_1 = g(X_1)$
doubt check ← $Y_1(\omega) = g(X_1(\omega))$, $\omega \in \Omega$
$Y_2 = h(X_2)$

prove Y_1 and Y_2 are independent. (① random var
② dis
③ ind)

$g: \mathbb{R} \to \mathbb{R}$ $\quad Y_1(\omega) = g_1(X_1(\omega))$ $\quad \omega \in \Omega$
$h: \mathbb{R} \to \mathbb{R}$ $\quad Y_2(\omega) = g_2(X_2(\omega))$

as X_1, X_2 are discrete random ind. variables

$$P(\underbrace{\{\omega \mid X_1(\omega) \leq x_1\}}_{A} \cap \underbrace{\{\omega \mid X_2(\omega) \leq x_2\}}_{B}) = P(A \cap B) = P(A)P(B)$$

as $\{\omega \mid X_1(\omega) \leq x\} \in \mathcal{F}$
$\forall x \in \mathbb{R}$ (as X_1 is a random variable)

as $g: \mathbb{R} \to \mathbb{R}$

$X_1(\omega)$ will have some values
(as it is discrete)
then $g(X_1(\omega))$ will also have discrete values.
$\therefore Y_1 = g(X_1(\omega))$ is a discrete random variable.

now as X_1, X_2 are ind

$$P(\{\omega \mid X_1(\omega) \leq x_1\} \cap \{\omega \mid X_2(\omega) \leq x_2\})$$
$$= P(\{\omega \mid X_1(\omega) \leq x_1\}) P(\{\omega \mid X_2(\omega) \leq x_2\})$$

as $X_1(\omega)$ are discrete values
for $g_1(X_1(\omega)) \leq x_1$ the same will happen,
$\therefore$ independent

4th sept :

Independence of discrete random variables:

let (X,Y) be a discrete random vector with prob. mass function p.
we say X,Y are independent if

$$P(X=x, Y=y) = P(X=x)\,P(Y=y), \quad \forall x,y \in \mathbb{R}$$
$$(*)$$

If (X,Y) takes values in $S=\{(x_i,y_j) \mid i\in\mathbb{N}, j\in\mathbb{N}\}$ then this $(*)$ condition implies $P(X=x_i, Y=y_j) = P(X=x_i)P(Y=y_j)$

$$\Rightarrow P(x_i, y_j) = P_X(x_i)P_Y(y_j), \quad \forall i,j\in\mathbb{N}$$

X,Y ind $\Rightarrow$ joint pmf = product of marginal pmf

See how converse is true

Note: converse is also true as $(*)$ holds trivially if $x\neq x_i$ or $y \neq y_j$

Conclusion: Two discrete random variables are independent iff joint p.m.f is product of marginal pmf.

Exercise: Suppose X,Y are discrete ind random variables then

$$P(X\in A, Y\in B) = P(X\in A)P(X\in B)$$
for $A,B \subseteq \mathbb{R}$

suppose X takes values in $\{x_1, \dots\}$
Y takes values in $\{y_1, \dots\}$

$$P(X\in A, Y\in B) = P\Big(\bigcup_{\substack{i,\, x_i\in A\\ j,\, y_j\in B}} \{X=x_i, Y=y_j\}\Big)$$

countable union of disjoint events

$$= \sum_{j:\, y_j\in B}\sum_{i:\, x_i\in A} P(X=x_i, Y=y_j)$$
$$= \sum\sum P(X=x_i)P(Y=y_j)$$
$$= P(X\in A)P(Y\in B)$$

Exercise: Suppose X,Y are discrete independent variables. $g,h:\mathbb{R}\to\mathbb{R}$
Let $Z=g(X)$, $V=h(Y)$

Show that Z,V are independent discrete random variables.

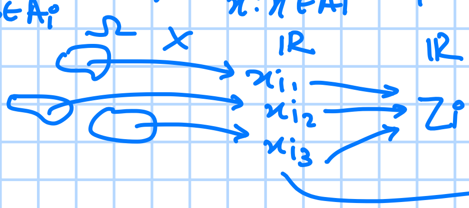
$$\begin{tikzcd} \Omega \arrow[r, "X"] \arrow[rr, bend right, "Z"'] & \mathbb{R} \arrow[r, "g"] & \mathbb{R} \end{tikzcd}$$

Z takes values on a atmost countable set say on $\{z_1, z_2, \dots\}$

$$\{Z=z_i\} = \{\omega \mid Z(\omega)=z_i\}$$

Suppose X takes values in $S=\{x_1, x_2, \dots\}$
let $A_i = \{x\in S \mid g(x)=z_i\}$

$$\{Z=z_i\} = \bigcup_{x:\, x\in A_i}\{X=x\} = \bigcup_{x:\, x\in A_i}\{\omega \mid X(\omega)=x\} \in \mathcal{F}$$



$\therefore \{Z=z_i\}\in\mathcal{F}$

By same argument, V is a discrete random variable.

$P(Z=z_i, V=v_j) = P(X \in A_i, Y \in B_j)$ where $B_j = \{y \in T \mid h(y)=v_j\}$ and Y takes values in $T=\{y_1, y_2, \dots\}$

and A_i as defined before.

Since X, Y are independent
$$P(X \in A_i, Y \in B_j) = P(X \in A_i)\, P(Y \in B_j) = P(Z=z_i)\, P(V=v_j)$$

Hence, Z and V are independent.

Def^n: We say a collection of discrete random variables $X_1, X_2, \dots X_n$ are independent if $P(X_{i_1}=x_{i_1}, X_{i_2}=x_{i_2}, \dots X_{i_k}=x_{i_k})$
$$= P(X_{i_1}=x_{i_1})\, P(X_{i_2}=x_{i_2}) \cdots P(X_{i_k}=x_{i_k})$$
for any $\{i_1, i_2, \dots i_k\} \subseteq \{1,2,\dots n\}$ and $x_{i_1}, \dots x_{i_k} \in \mathbb{R}$

Expectation / Expected value / mean value:

Suppose X is a discrete random variable taking value in $\{x_1, \dots\}$, Expectation of X is denoted as
$$E(X) = \sum_{i=1}^{\infty} x_i P(X=x_i) = \sum_{i=1}^{\infty} x_i P(x_i)$$ if $\sum_i x_i P(x_i)$ is absolutely summable, mean $\sum_i |x_i| P(x_i) < \infty$

We have $|x_i|$ as if not absolutely summable then it leads to $\sum x_i P(X=x_i)$ not being unique. (sometimes)

Exercise: Find expectation of the standard discrete random variables which we have discussed in class. → done down

Remarks: Let X be a discrete random variable and $f: \mathbb{R} \to \mathbb{R}$.
Let $Y = f(X)$ → is a discrete random variable → see proof right after this
then $E(Y)$ exist iff $\sum_{i=1}^{\infty} |f(x_i)| P(x_i) < \infty$, find $E(Y)$

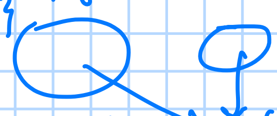
Terminology:
If $\sum_i |x_i| P(x_i) < \infty$ then we say that $E(X)$ exist.

$E(Y)$ exist then $\underbrace{\sum y_j P(Y=y_j)}_{\text{this is finite}} < \infty$

$$\sum_j |y_j| P(Y=y_j) < \infty$$

X takes values in $S = \{x_1, x_2, \dots\}$ $\quad \cup A_j = \{x_1, x_2 \dots\}$
$$A_j = \{x \in S \mid f(x) = y_j\}$$

Aj Aj+1



$P(Y=y_j) = \sum_{x : x \in A_j} P(X=x)$, then $\quad$ Note $A_j \cap A_{j+1} = \phi$

$$\sum_j |y_j| P(Y=y_j) = \sum_j |y_j| \sum_{x: x \in A_j} P(X=x) = \sum_j |y_j| \sum_{x : x \in A_j} P(x) = \sum_j \sum_{x \in A_j} |f(x)|\, P(x)$$

$$= \sum_{i=1}^{\infty} |f(x_i)| P(x_i)$$

$E(Y) = \sum_{i=1}^{\infty} f(x_i) P(x_i) \rightarrow$ this result is very useful

(Also Note, $E(Y)$ exist if $\sum |f(x_i)| P(x_i) < \infty$)

Usefulness: To calculate $E(Y)$ we don't have to calculate pmf of Y. We can calculte $E(Y)$ from the p.m.f of X.

Exercise: Suppose $X: \Omega \rightarrow \mathbb{R}^d$ be a discrete random vector with p.m.f P. Let $g: \mathbb{R}^2 \rightarrow \mathbb{R}$.

① Show that $Y = g(X)$ is a discrete random variable.

② Show that $E(Y)$ exist iff $\sum_{\bar{x} \in S} |f(\bar{x})| P(\bar{x}) < \infty$

where X takes values in S. (Atmost countable)

Exercise: Find expectation of the standard discrete random variables which we have discussed in class.

① Bernoli:

$P, \; 1-P$

$P(X = x_1) = P$

$P(X = x_2) = 1 - P$

then

$E(X) = x_1 P + x_2 (1-P)$

for $x_2 = 0$, $x_1 = 1$

$\Rightarrow E(X) = P$

② Binomial:

$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$ also for $k = 0, 1, \ldots, n$ true

$$\therefore E(X) = \sum_{k=1}^{n} k \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \frac{(n)!}{(n-k)!(k)!} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} \cancel{k} \left(\frac{n}{\cancel{k}}\right) \binom{n-1}{k-1} p^k (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{n-k}$$

$$= np \sum_{k=1}^{n} \binom{n-1}{k-1} p^{k-1} (1-p)^{(n-1)-(k-1)}$$

$$= np \left[\cancel{p} + (1 - \cancel{p})\right]^{n-1} = np$$

③ Poision:

$$E(X) = \sum_{1}^{\infty} \frac{e^{-\lambda} \lambda^k}{(k)!} \cdot k = np = \lambda$$

as $n \rightarrow \infty$

④ geometric:

$$E(X) = \sum_{k=1}^{\infty} (1-p)^{k-1} (p) \, k$$

$$(1-P) E(X) = \sum_{k=1}^{\infty} (1-p)^k (p) \, k$$

$$P \, E(X) = \left[\frac{1-P}{1}\right] \cdot 1 \Rightarrow E(X) = \frac{1-P}{P}$$

⑤ inverse:

$$E(X) = \sum_{k=r}^{\infty} \binom{k-1}{r-1} p^r (1-p)^{k-r} \, k$$

$$= \sum_{k=r}^{\infty} \frac{(k-1)!\, k}{(r-1)!} \cdot \frac{r}{r(k-r)!} p^r (1-p)^{k-r}$$

$$= \sum_{k=r}^{\infty} \frac{k!}{(r)!\,(k-r)!} \cdot r\, p^r (1-p)^{k-r}$$

$$= \sum_{k=r}^{\infty} r \binom{k}{r} p^r (1-p)^{k-r}$$

$$= \frac{r}{p} \sum_{k=r}^{\infty} \binom{k}{r} p^{r+1} (1-p)^{(k+1)-(r+1)}$$

$$= r p^r \sum_{k=r}^{\infty} \binom{k}{r} (1-p)^{(k+1)-(r+1)}$$

doubt

Exercise: suppose $X: \Omega \to \mathbb{R}^d$ be a discrete random vector with p.m.f P.
let $g: \mathbb{R}^2 \to \mathbb{R}$.

① Show that $Y = g(X)$ is a discrete random variable.
② Show that $E(Y)$ exist iff $\sum_{\bar{x} \in S} |f(\bar{x})| P(\bar{x}) < \infty$
where X takes values in S. (Atmost countable)

① $Y = g(X)$ is a discrete random variable, firstly what we do is that
as X takes values in $S = \{x_1, x_2, \ldots\}$
$g(X)$ takes values in $\{g_1, g_2, \ldots\}$

now, let $A_i = \{x \in S \mid g(x) = g_i\}$

then $P(Y = g_i) = P(X \in A_i) = \sum_{x_i \in A} P(X = x_i) \in \mathcal{F}$

$\therefore$ $Y = g(X)$ is a discrete variable

② now $E(Y)$ exist then

$$\Leftrightarrow \sum_{1}^{\infty} |g_i| P(g_i) < \infty$$

$$\Leftrightarrow \sum_{1}^{\infty} |g_i| P(X \in A_i) < \infty$$

$$\Leftrightarrow \sum_{1}^{\infty} |g_i| \sum_{x : x \in A_i} P(x) < \infty$$

$$\Leftrightarrow \sum_{1}^{\infty} |g(x_i)| P(x_i) < \infty$$

6th Sept:

Note: X is a discrete random vector, $X: \Omega \to \mathbb{R}^2$.

$f: \mathbb{R}^2 \to \mathbb{R}$ and X takes values on a countable set S.

$Y = f(X)$ is a discrete random variable

$E(Y)$ exist $\Leftrightarrow \sum |f(x_i)| P(x_i) < \infty$

And $E(Y) = \sum f(x_i) P(x_i)$

Proof: last class

Properties of expectation of discrete random variables:

Suppose X and Y are discrete random variables, and $E(X)$, $E(Y)$ exist.

(i) $E(cX)$ exist and $E(cX) = cE(X)$ where $c \in \mathbb{R}$

observe: $cX = Z$ is a random variable

$E(X)$ exist, so $\sum |x_i| P(x_i) < \infty$

$\therefore \sum |c x_i| P(x_i) < \infty$

$\therefore E(cX)$ exist

also $E(cX) = \sum c x_i P(x_i) = c \sum x_i P(x_i) = cE(X)$

(ii) $E(X+Y)$ exist and $E(X+Y) = E(X) + E(Y)$

Here $Z = X + Y$

$= f(X, Y)$ $\quad (X,Y): \Omega \to \mathbb{R}^2$

$= X + Y$ $\quad$ Random vector

$f(x,y) = x + y$

now we have to check

if $\sum\sum |f(x,y)| P(x,y) < \infty$

(X,Y) takes values in

$\{(x_i, y_j) \mid i, j \in \mathbb{N}\}$

$E(X,Y)$ exist iff $\sum\sum |f(x_i, y_j)| P(x_i, y_j) < \infty$

where $f(x_i, y_j) = x_i + y_j$

$\sum\sum |x_i + y_i| P(x_i + y_i)$

$\leq \sum\sum |x_i| P(x_i + y_i) + \sum\sum |y_i| P(x_i + y_i)$

$= \sum |x_i| P(x_i) + \sum |y_i| P(y_i)$

$< \infty$ $\therefore E(X,Y)$ exist

Similarly $E(X+Y) = \sum\sum (x_i + y_i) P(x_i, y_i)$

$= \sum (x_i) P(x_i) + \sum (y_i) P(y_i)$

$= E(X) + E(Y)$

(iii) $E(cX + dY)$ exist and $E(cX+dY) = cE(X) + dE(Y)$ where $c, d \in \mathbb{R}$

proof of this follows from first two

(iv) Suppose $X \leq Y$, then $E(X) \leq E(Y) \rightarrow$ $X - Y$ is a random variable

$E(X-Y) = E(X) - E(Y)$

$X - Y \leq \;\Rightarrow$

$E(X) - E(Y) \leq 0$

(v) Suppose $P(X \leq Y) = 1$, then $E(X) \leq E(Y)$

$P(\{\omega \mid X(\omega) \leq Y(\omega)\})$

now if $X \leq Y$ then we know $\quad (X \leq Y \Rightarrow P(X \leq Y))$

$P(X \leq Y) = P(\Omega) = 1$

now $\{\omega \mid X(\omega) \leq Y(\omega)\}$

Note: $P(X \leq Y)$

$\hookrightarrow$ can have ωs smaller than Ω still giving $P(\{\omega \dots\}) = 1$

example for above condition ($P(X \leq Y)=1 \not\Rightarrow X \leq Y$)

$\Omega = \{1,2,\dots n\}$
$\mathcal{F} = P(\Omega)$
$P(\{i\}) = \frac{1}{n-1}, \; i=1,2,\dots n-1$
$P(\{n\}) = 0$

this is a probability space,
$P(\{1,2,\dots n-1\})=1$
$P(\Omega)=1$
but $\Omega \neq \{1,2,\dots n-1\}$

$X: \Omega \to \mathbb{R}$
$X(\omega)=\omega$
$Y: \Omega \to \mathbb{R}$
$Y(\omega) = \begin{cases} \omega+1 & \text{if } \omega=1,2,\dots n-1 \\ 0 & \text{if } \omega = n \end{cases}$

we $P(X \leq Y)=1$
but $X \not\leq Y$

this was a counter example to
$P(X \leq Y) \Rightarrow X \leq Y$

$Z = Y-X$
$P(Z \geq 0)=1$
$E(Z) = \sum_{z} z_i P(z_i)$

if $z_i < 0$ then
$P(z_i)=0$
similarly $z_i \geq 0$
then $P(z_i)=1$

$\therefore E(Z) \geq 0$
$\Rightarrow E(Y-X) \geq 0$
$\Rightarrow E(Y) \geq E(X)$

(vi) $|E(X)| \leq E(|X|)$
as X is a random variable,
$|X|$ is also a random
variable
and as $X \leq |X|$
$\Rightarrow E(X) \leq E(|X|)$
and $-X \leq |X|$
$\Rightarrow E(-X) \leq E(|X|) \Rightarrow -E(X) \leq E(|X|)$
$\therefore |E(X)| \leq E(|X|)$

Exercise: suppose X and Y are independent discrete random variables and $E(X)$, $E(Y)$ exist,
show that $E(XY)$ exist and find $E(XY)$

Note, we don't have to find p.m.f XY, we can calculate $E(XY)$ using p.m.f of (X,Y)

$X \sim \{x_1, x_2, \dots\}$
$Y \sim \{y_1, \dots\dots\}$

$E(XY)$ exist $\Leftrightarrow \sum\sum |x_i y_j| P(x_i, y_j)$
joint pmf (X,Y)

$\sum\sum |x_i y_j| P(x_i, y_j) = \sum\sum |x_i||y_j| P_X(x_i) P_Y(y_j)$ ($\because$ independence of X,Y)
$= \left(\sum |x_i| P_X(x_i)\right)\left(\sum |y_i| P_Y(y_i)\right)$
$\quad\quad < \infty \quad\quad\quad < \infty$
$\therefore E(XY) < \infty$

and
$E(XY) = E(X)E(Y)$

Common mistake -
$E(XY) = E(X)E(Y)$
$\Rightarrow X, Y$ are independent

Exercise: Suppose $P(|X| \le M) = 1$, then show that $E(X)$ exist and $|E(X)| \le M$
→ done

Def^n: k-th moment of X is defined as $E(X^k) = \sum_i x_i^k P(x_i)$ provided $\sum_i |x_i|^k P(x_i) < \infty$
$(k \in \mathbb{N})$

k-th central moment:
$$E((X-\mu)^k) = \sum_i (x_i - \mu)^k p(x_i)$$
provided $\sum_i |(x_i-\mu)^k| p(x_i) < \infty$
where $\mu = E(X)$ exist

Def^n: Variance:
$k=2$
$E(X-\mu)^2 = Var(X)$ → variance of X

Random var. X → draw value / observe ⇒ close to mean value or $E(X)$

$E(X)$ — what we expect

$E(X-\mu)^2$ — expectation of X around its mean

10 $E(X)=10$ → many values are 10 / close to 10
$Var\ X = 0.001$

if $E(X) = 10$, $Var\ X = 0$
$P(X=10) = 1$
or X is almost a constant random variable

$k=3$, $k=4$ } Have specific names for $E((X-\mu)^k)$
$(k \ge 2)$ (as $k=1$, $E(X) = \mu$)

Result: k-th moment exist iff k-th central moment exist.
→ done

Result: $k < r$, $k, r \in \mathbb{N}$, X is discrete r-v. suppose r-th moment of X exist
Then k-th moment of X exist.

$$\sum |x_i|^k P(x_i) = \sum_{|x_i| \le 1} |x_i|^k P(x_i) + \sum_{|x_i|>1} |x_i|^k P(x_i)$$

$$\le \sum_{|x_i| \le 1} P(x_i) + \sum_{|x_i|>1} |x_i|^r P(x_i) \le 1 + \sum_{|x_i|>1} |x_i|^r P(x_i) < \infty$$
as $r > k$ and $|x_i| > 1$

Exercise: Suppose X, Y are discrete r.vs and k^{th} moment of X, Y exist. Then show that k-th moment of $(X+Y)$ exist. → done

Exercise: Suppose $P(|X| \le M) = 1$, then show that $E(X)$ exist and $|E(X)| \le M$
as $P(|X| \le M)$ $\xleftarrow[-M \quad M]{}\!\!\rightarrow$ X takes values in $[-M, M]$,
then $\sum |x_i| P(x_i) \le \sum M P(x_i) = M < \infty$ ∴ $E(X)$ exist
now as $E(X)$ exist, $|E(X)| \le E(|X|) = \sum |x_i| P(x_i) \le M$ (just proved)

∴ $|E(X)| \le M$

Result: k-th moment exist iff k-th central moment exist.

($\Rightarrow$) using the fact that $r < k$ and k^{th} moment exist
$\Rightarrow$ r^{th} moment exist

here $\sum |x_i|^k p(x_i) < \infty$
also $\sum |x_i|^r p(x_i) < \infty$
$\forall\, r = 1, 2, \ldots k$

now, $\sum |x_i - \mu|^k p(x_i)$

$$\leq \sum \left(|x_i|^k + |\mu|^k + \binom{k}{1} |x_i|^{k-1} |\mu|^1 \cdots \right) p(x_i)$$

as cons $\times$ r^{th} moment exist

$$\Rightarrow \sum |x_i - \mu|^k p(x_i) < \infty$$

($\Leftarrow$) $\sum |x_i - \mu|^k p(x_i)$ exist, then
$\Rightarrow \sum |x_i - \mu|^r p(x_i)$ also exist $\forall\, r = 1, 2, \ldots k$

now, similar to above $\sum |x_i|^k p(x_i)$

$$= \sum |(x_i - \mu) + (\mu)|^k p(x_i)$$
$$\leq \sum |x_i - \mu|^k p(x_i) + \underbrace{\cdots\cdots}_{r^{th} \text{ centre moment} \times \text{some const (exist)}}$$

$$\therefore \sum |x_i|^k p(x_i) < \infty$$

Exercise: Suppose X, Y are discrete r.vs and k^{th} moment of X, Y exist. Then show that k-th moment of $(X+Y)$ exist.

given $\sum_x |x_i|^k p(x_i) < \infty$
and $\sum_y |y_j|^k p(y_j) < \infty$

now let $f(x_i, y_i) = (x_i + y_i)^k$

then $\sum \sum |x_i + y_i|^k P(x_i, y_i)$

$$\leq \sum \sum |x_i|^k p(x_i, y_i) + \sum \sum |y_i|^k p(x_i, y_i)$$
$$= \sum |x_i|^k P_x(x_i) + \sum |y_i|^k p(y_i)$$

$< \infty$ $\therefore$ k^{th} moment of $E(X+Y)$ exist.

10th sept :

k-th order moment

Exist if $\sum |x_i|^k p(x_i) < \infty$.

$$E(X^k) = \sum (x_i^k)\, p(x_i)$$

k-th order central moment

$$E[(X-\mu)^k] = \sum (x_i-\mu)^k\, p(x_i)$$
provided $\sum |x_i-\mu|^k p(x_i) < \infty$

Note : k-th moment exist $\Leftrightarrow$ kth central moment exist (Discrete random variables)

Result : Suppose X and Y have k-th order moment, Then (X+Y) has k-th order moment.

$$E|X|^k = \sum |x_i|^k p(x_i) < \infty$$
$$E|Y|^k = \sum |y_i|^k p(y_i) < \infty$$

if $Z = f(X,Y)$

expected value of Z exist iff

$$\sum |f(x_i,y_j)|\, P(x_i,y_j) < \infty$$

to show:

$$\sum\sum |x_i+y_j|^k\, P(x_i,y_j) < \infty$$

as $|x_i+y_j|^k \leq (|x_i| + |y_j|)^k$
$\leq (2\max\{|x_i|,|y_j|\})^k = 2^k(\max\{|x_i|,|y_j|\})^k$
$= 2^k(\max\{|x_i|^k,|y_j|^k\})$
$\leq 2^k\left(|x_i|^k + |y_j|^k\right)$

$\therefore \sum\sum |x_i+y_j|^k P(x_i,y_j) \leq \sum\sum 2^k(|x_i|^k + |y_j|^k) P(x_i,y_j)$
$\leq 2^k \sum |x_i|^k p_x(x_i) + 2^k \sum |y_j|^k p_y(y_j)$
$< \infty$

Hence, k-th moment of (X+Y) exist.

Def^n : $Var(X) = E(X-\mu)^2$ Non-negative random varible
where $\mu = E(X)$ $\therefore var(X) \geq 0$ (common mistake)

var(X) measures its distribution of X around its mean value.

Def^n : Covariance of X and Y :

$$cov(X,Y) = E[(X-E(X))(Y-E(Y))]$$
$$= E[(X-\mu_X)(Y-\mu_Y)] = E(XY) - E(X)E(Y)$$
where $\mu_X = E(X)$
$\mu_Y = E(Y)$

some kind of measuremet of relationship b/w X and Y.

Exercise : Find cov(X,Y) where X, Y are independent.

$cov(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$
$= E[(XY - \mu_Y X - \mu_X Y + \mu_X\mu_Y)]$
$= E(XY) - \mu_Y E(X) - \cancel{\mu_X E}(Y) + \cancel{\mu_X}\mu_Y$
$= E(XY) - \mu_Y \mu_X$

and also $E(XY) = \sum\sum (x_i y_j) P(x_i y_j)$
$= \sum (x_i) P(x_i) \sum y_j P(y_j)$
$= E(X)E(Y)$

$\therefore cov(X,Y) = E(X)E(Y) - E(X)E(Y)$
$= 0$

Conclusion: If X, Y are independent then $Cov(X,Y) = 0$

Exercise: Suppose $cov(X,Y) = 0$, can we conclude that X and Y are independent?

No as $E(XY) = E(X)E(Y)$
as notion of independence gives huge information, this does not.

Defn: Correlation coefficient

$$cor(X,Y) = \frac{Cov(X,Y)}{\sqrt{var X}\sqrt{var Y}}$$ provided $var(X)$, $var(Y) \geq 0$

$$cor(X,Y) = \frac{E((X-\mu_x)(Y-\mu_y))}{\sqrt{var X}\ \sqrt{var Y}}$$

if $E((X-\mu x)(Y-\mu y)) > 0$
then for $X > \mu x$
$\Rightarrow Y > \mu y$
High probability

Sign of covariance of X, Y gives us some idea of the distribution of (X,Y) around their expected.

But value of covariance doesn't say much.

Note: we will see that correlation gives us much more information regarding the relation b/w X and Y.

Note: X and Y are independent, then $cor(X,Y) = 0$

Cauchy Schwarz inequality:

Suppose X and Y are two random variables with finite second moment.
$(E|X|^2 < \infty,\ E|Y|^2 < \infty)$

then $[E(XY)]^2 \leq E(X^2)E(Y^2)$

and equality holds if $P(X=0)=1$ or
$P(Y=aX)=1$ for some constant $a \neq 0$

with this inequality, $|cov(X,Y)| \leq \sqrt{var X}\sqrt{var Y}$

$Z = (X - \mu x)$
$W = (Y - \mu y)$

then $(E(ZW))^2 \leq E(Z^2)E(W^2)$
$\Rightarrow (cov(XY))^2 \leq var(X)\, var(Y)$
$\Rightarrow |cov(XY)| \leq \sqrt{var(X)}\ \sqrt{var(Y)}$

therefore $-1 \leq cor(X,Y) \leq 1$

If $cor(X,Y) = \pm 1$ then X, Y are linearly related
as, $cor(X,Y) = \pm 1$
$\Rightarrow |cov(X,Y)| = \sqrt{var X}\ \sqrt{var Y}$
also note as equality for $P(X=0)=1$
or $P(X=aY)=1$

here as for $P(X=0)=1$
$\Rightarrow var\ X = 0$
but as $var\ X \neq 0$

we have $P(X=aY)=1$
$\therefore$ X and Y are linearly related

24th sept:

Observation: If X and Y are independent, $Cov(X,Y) = cov(X,Y) = 0$
$\rightarrow (E(XY) - E(X)E(Y))$

But converse not true.
$cov(X,Y) = 0 \not\Rightarrow X, Y$ are independent.

eg:
suppose X, Y are uniformly distributed on $\{-n, -n+1, \ldots, 0, 1, 2, \ldots n\} = S$
$$P(X=i) = \frac{1}{2n+1}, \quad i \in S$$
suppose Y is uniformly distributed on $\{1, 2, \ldots k\}$
$$P(Y=j) = \frac{1}{k}, \quad j \in \{1, 2, \ldots k\}$$

Assume X, Y are independent
Define
$$Z = X^2 + Y$$
$$E(XZ) = E(X^3 + XY)$$
$$= E(X^3) + E(XY)$$
$$= 0$$
$$E(X)E(Z) = 0$$

so $Cov(X, Z) = 0$

$$P(x, z) = \begin{cases} \frac{1}{k} \cdot \frac{1}{2n+1}, & z = x^2 + y \quad x \in \{-n, \ldots n\}, y \in \{1, 2, \ldots k\} \\ 0, & \text{else} \end{cases}$$

Exe: calculate the p.m.f of Z and see that joint pmf of $XZ \neq$ marginal product of X, Z. ← done above

Cauchy-Schwarz inequality:
Let X and Y be (discrete) random variables, where $E(X^2)$ and $E(Y^2)$ exist with mean 0. Then
$$[E(XY)]^2 \leq E(X^2)E(Y^2)$$
Equality holds iff $P(X=0) = 1$ or $P(Y = aX) = 1$, $a \in \mathbb{R}$

proof: suppose $P(X=0) = 1$
then $E(X) = 0$, $P(XY=0) \geq P(X=0) = 1$, $E(X^2) = 0$
thus $E(XY) = 0$
$$(E(XY))^2 = E(X^2)E(Y^2) = 0$$
suppose $P(Y = aX) = 1$
$$\Rightarrow Y(\omega) = aX(\omega)$$
$$E(XY) = E(aX^2)$$
$$= aE(X^2)$$
$$(E(XY)^2) = a^2 (E(X^2)^2)$$
$$= a^2 E(X^2) E(X^2)$$

Now, if $E(X^2) > 0$
define $g(\lambda) = E[(Y - \lambda X)^2]$
$$g(\lambda) \geq 0$$
since $(Y - \lambda X)^2 \geq 0$
$$g(\lambda) = E(Y^2 - 2\lambda XY + \lambda^2 X^2)$$
$$= E(Y^2) - 2\lambda E(XY) + \lambda^2 E(X^2)$$
$$= E(X^2)\left[\lambda^2 - 2\lambda \frac{E(XY)}{E(X^2)} + \frac{(E(XY))^2}{(E(X^2))^2}\right] + E(Y^2) - \frac{(E(XY))^2}{E(X^2)}$$
$$= E(X^2)\left[\lambda - \frac{E(XY)}{E(X^2)}\right]^2 + E(Y^2) - \frac{(E(XY))^2}{E(X^2)}$$

$g(\lambda)$ attains minima at $\lambda = \frac{E(XY)}{E(X^2)}$

$$g(\lambda_0) \geq 0$$
$$E(Y^2) - \frac{[E(XY)]^2}{E(X^2)} \geq 0$$
$$\Rightarrow E(X^2)E(Y^2) \geq (E(XY))^2$$
if equality holds, then
$$g(\lambda_0) = 0$$

$$E(Y - \lambda_0 X)^2 = 0$$
$$P(Y - \lambda_0 x = 0) = 1$$

Sol: ① $|cor(X,Y)| \leq 1$
② $|cor(X,Y)| = 1$
iff $\exists\, a \neq 0, b$ s.t
$P(Y = aX + b) = 1$

if $cor(X,Y) = 1$ then $a > 0$ → dere down
if $cor(X,Y) = -1$ then $a < 0$

Sol: calculate the p.m.f of Z and see that joint pmf of XZ ≠ marginal product of X, Z.

$$P(x,z) = P(X = x, Z = z) = P(X = x,\ X^2 + Y = Z)$$
as $X = x$
$$= P(X = x,\ Y = Z - x^2)$$
$$P_x(X = x)\, P_Z(Z = X^2 + Y = z)$$

clearly $P(X = x)\, P(Z = z)$
$\neq P(X = x,\ Y = z - x^2)$
depend on x

so Not independent.

Sol: ① $|cor(X,Y)| \leq 1$
② $|cor(X,Y)| = 1$
iff $\exists\, a \neq 0, b$ s.t
$P(Y = aX + b) = 1$

if $cor(X,Y) = 1$ then $a > 0$
if $cor(X,Y) = -1$ then $a < 0$

Now as $[E(XY)]^2 \leq E(X^2)E(Y^2)$

for $Z = X - \mu_x$
$W = Y - \mu_y$

$$[E(ZW)]^2 \leq var(X)\, var(Y)$$
$$[cov(X,Y)]^2 \leq var(X)\, var(Y$$

$$\Rightarrow \frac{|cov(X,Y)|}{\sqrt{var(X)}\sqrt{var(Y)}} \leq 1$$

$$\Rightarrow |cor(X,Y)| \leq 1$$

if $|cor(X,Y)|=1$ so they are lin. dep ($X \neq 0$ as var X in deno.)

$\Rightarrow \quad W = \alpha Z$

$\Rightarrow \quad Y = \alpha X + \beta$

now $\quad cor(X,Y)=1$

$\Rightarrow \quad E(X-\mu_x)(Y-\mu_y) \geq 0$

$E(X-\mu_x)(\alpha)(X-\mu_x) \geq 0$

$\Rightarrow \alpha E(X-\mu_x)^2 \geq 0$

$\Rightarrow \alpha > 0 \quad$ (as $E(M^2) \geq 0$)

similary, $\alpha<0$ a similar case occurs.

25th sept:

conditional probab. mass function:

Let X, Y be discrete random variable with joint pmf $P(x,y)$
conditional prob of Y, given $X=x$ is defined as:

$$P_{Y|X}(y|x) = P(Y=y \mid X=x) = \frac{P(x,y)}{P_X(x)} \quad \text{as} \quad P_X(x) = P(X=x)$$

provided $P_X(x) \neq 0$

Note: If X, Y are independent then conditional PMF of Y is same as P.M.F of Y.

Similarly, conditional pmf of X given $Y=y$ is defined as $P_{X|Y}(x|y) = \frac{P(x,y)}{P_Y(y)}$

provided $P_Y(y) \neq 0$

conditional distribution function:

of Y given $X=x$ is defined as

$$F_{Y|X}(y|x) = \sum_{y_i \leq x} P_{Y|X}(y_i|x)$$

Conditional Expectation:

Y given $X=x$

$$E[Y \mid X=x] = \sum_y y\, P_{Y|X}(y|X)$$

denote $\Psi(x) = E[Y \mid X=x]$

we can visualise at function in x.
(function of small x, $P(x) > 0 \rightarrow$ domain)

$\Psi(X) = E[Y|X] : \Omega \longrightarrow \mathbb{R}$ (This can be calculated using $\Psi(x) = E(Y \mid X=x)$)

Random variable as : $\Omega \longrightarrow \mathbb{R}$

Expectation of $\Psi(X)$:

Note: $g: \mathbb{R} \rightarrow \mathbb{R}$

$$E[g(X)] = \sum_{x_i} g(x_i)\, P_X(x_i)$$

we $\Psi(X)$ as $X: \Omega \longrightarrow \mathbb{R}$

$\Psi(X): \Omega \longrightarrow \mathbb{R}$

where $\Psi: \mathbb{R} \rightarrow \mathbb{R}$

$$
\begin{aligned}
E(\Psi(X)) &= \sum_x \Psi(x) P_X(x) \\
&= \sum_x \sum_y y\, P_{y|x}(y|x)\, P_x(x) \\
&= \sum_x \sum_y y \frac{P(x,y)}{\cancel{P_X(x)}} \cancel{P_x(x)} \\
&= \sum_x \sum_y y\, P(x,y) \\
&= \sum_y y\, P(y) = E(Y)
\end{aligned}
$$

Result: $E(\Psi(X)) = E(Y)$ or $E\big(E(Y|X)\big) = E(Y)$

usefulness of result - $E[E[Y|X]] = E[Y]$
so if we know $E[Y|X]$, we can find $E[Y]$.

Eme: A factory producing N no of machines in a day where N is Poisson distributed with parameter λ. A machine can be defective with prob. P independent of any other machines. Find the expected no of defective machine produced in a day.

let Y be no. of defective machines produced in a day.
if $N=n$, then distribution of Y can be known
$Y = \sum_{i=1}^{N} X_i$, $X_i = 1$ if defective, $X_i = 0$ otherwise

$N=n$, $Y = \sum_{i=1}^{n} X_i$ $\{X_i\}$ are independent $Ber(P)$

and $\sum$ bernoui = binomial, coin example, H/T ← Binomial, for n coins
X = no of heads follows binomial (n, P),
so Y = binomial (N), if $N=n$, then
$Y \sim Bin(n, P)$

given $N=n$, the conditional distribution of Y is binomial (n, P). conditional expectation of $E[Y|N=n] = nP$

$$E[Y|N=n] = E[Z]$$

Z follows $bin(n,P)$
so $E[Z] = nP$

now, $E[Y/N] = Np$

$$E[E[Y/N]] = E[Np] = pE[N] = p\lambda$$

Eme: Find $E[N|Y]$
for this $E[N|Y=y]$, then we can calculate.

$$\text{Pmf } P_{N|Y}(n,y) = \frac{P(n,y)}{P_Y(Y=y)}$$

$$\text{Now } E(N|Y=y) = \sum n_i P_{N|Y}(n_i, y)$$

and so we are done.

$P_{N|Y}(n|y)$, as $Y=y$ → No. of defective $m = y$
$n \geq y$
so $P_{N|Y}(n|y) = 0$ for $n < y$

for $n \geq y$

$$P_{N|Y}(n|y) = \frac{P(N=n, Y=y)}{P(Y=y)}$$

$$= P(Y=y|N=n)\frac{P(N=n)}{P(Y=y)}$$

$$= \binom{n}{y} p^y (1-p)^{n-y} \times \left(\frac{e^{-\lambda}\lambda^n}{n!}\right) \times \frac{1}{P(Y=y)}$$

$$P_{N|Y}(n|y) = \binom{n}{p} p^y(1-p)^{n-y} \times \left(\frac{e^{-\lambda}\lambda^n}{n!}\right) \times \frac{1}{\sum_{k=y} P(Y=y|N=k)P(N=k)}$$

$$= \frac{\binom{n}{p} p^y(1-p)^{n-y} \times \left(\frac{e^{-\lambda}\lambda^y}{n!}\right)}{\sum_{k=y}^{\infty} \binom{k}{y} p^y(1-p)^{k-y}\frac{e^{-\lambda}\lambda^k}{k!}} = \frac{(1-p)^{n-y}}{1-(1-p)\lambda}\frac{\lambda^{n-y}}{(n-y)!}$$

from this, $E[N|Y=y] = \frac{y}{(1-(1-p)\lambda)(1-(y+1)(1-p)\lambda)}$

(for $g=1$, we get prev result, we can also prove with definition)

Theorm: $E[\Psi(X)g(X)] = E[Yg(X)]$

$\Psi(X) = E[Y|X]$

Proof: using from next question, $E[Yg(X)|X] = g(X)E[Y|X]$

$$E[Yg(X)|X] = g(X)\Psi(X)$$

$$E[E[Yg(X)|X]] = E[Yg(X)] = E[\Psi(X)g(X)]$$

Eme: show that expected value $E[Yg(X)|X] = g(X)E[Y|X]$

Here $E[Yg(X)|X=x] = \sum_y yg(x)P_{Y|X}(y|x)$

this is random

$$= g(x)\sum y P_{Y|X}(y|x)$$

$$= g(x)E[Y|X=x]$$

$$E[Yg(X)|X] = g(X)E[Y|X]$$

Conditional var of Y given X=x:

$$\mathrm{var}(Y|X=x) = \sum_y (y - E(Y|X=x))^2 P_{Y/X}(y|x)$$

Note: $\mathrm{var}(Y) = E(Y^2) - (E(Y))^2$

here $\mathrm{var}(Y|X=x) = E(Y^2|X=x) - [E(Y|X=x)]^2$

Theorm: $\mathrm{var}(Y) = E[\mathrm{var}(Y|X)] + \mathrm{var}[E(Y|X)]$

proof:

Note: $\mathrm{var}(Y|X) = E(Y^2|X) - (E(Y|X))^2$

$$E[\mathrm{var}(Y|X)] = E(E(Y^2|X)) - E[(E(Y|X))^2]$$

$$= E(Y^2) - E(E(Y|X))^2$$

$$\mathrm{var}[E(Y|X)] = E[(E(Y|X))^2] - [\underbrace{E(E(Y|X))}_{E(Y)}]^2$$

so $E(\mathrm{var}(Y/X)) + \mathrm{var}(E(Y/X)) = E(Y^2) - E(Y)^2$

$= \mathrm{var}(Y)$

Que: Find var(Y) where Y is the number of defective machine.

Here

$$var(Y) = E[var(Y|N)] + var(E(Y|N))$$

$$E(Y|N) = Np \sim Poisson(\lambda) \times P$$

$$var(\underbrace{Y|N=n}_{Binomial(n,P)}) = (nP)(1-P)$$

$$\Rightarrow var(Y|N) = NP(1-P) \sim Poisson(\lambda) \times P \times (1-P)$$

$$E(var(Y/N)) = \lambda P(1-P)$$

$$var(NP) = \lambda P$$

$$\sim Poisson(\lambda) \times P$$

so $var(Y) = \lambda P - \lambda P^2 + \lambda P$

$$= 2\lambda P - \lambda P^2$$

$$var(Y) = \lambda P(2-P)$$

27th sept :

Markov inequality :

(Note : True for any random variable)

Suppose X is a non-negative random variable and $E(X)$ exist. Then for $a>0$,

$$P(X \geq a) \leq \frac{E(X)}{a}$$ (Here we use it for discrete)

example : prob of rain in bombay crossing 100 mL (X = mls of rain)

Here X is non-negative, $E(X) \rightarrow$ expectation of mls of rain for all x_i.

proof :

$$A = \{\omega \mid X(\omega) \geq a\} = \{X \geq a\}$$

$Y = 1_A \rightarrow$ indicator of A.

$$Y(\omega) = \begin{cases} 1 & \text{if } \omega \in A \\ 0 & \text{otherwise} \end{cases}$$

then as $X(\omega) \geq a$

$$X(\omega) \geq Y(\omega) a$$

$$X \geq aY$$

$$\Rightarrow E(aY) \leq E(X)$$

$$\Rightarrow E(Y) \leq \frac{E(X)}{a}$$

$$\Rightarrow \sum y_i P(y_i) \leq \frac{E(X)}{a}$$

$$\Rightarrow P(Y=1) \leq \frac{E(X)}{a}$$

$$\Rightarrow P(A) \leq \frac{E(X)}{a}$$

$$\Rightarrow P(X \geq a) \leq \frac{E(X)}{a}$$

chebyshev's inequality

Suppose X is a (discrete) random variable with mean μ and variance σ^2. Then for $a>0$,

$$P(|X-\mu| \geq a) \leq \frac{\sigma^2}{a^2}$$

proof : Note,

$$P(|X-\mu| \geq a) = P((X-\mu)^2 \geq a^2)$$

$$\leq \frac{E((X-\mu)^2)}{a^2} = \frac{\sigma^2}{a^2}$$

so, $P(|X-\mu| \geq a) \leq \dfrac{\sigma^2}{a^2}$

Note, $P(|X| \geq a) \leq \frac{E(X^2)}{a^2}$ (This is also true

Ex- X = No of items produced in a factory in a day
suppose $E(X) = 50$
(a) give bound on $P(X \geq 75)$ → done above
(b) $\text{var}X = 25$, get bound on $P(40 < X < 60)$

one side chebyshev's inequality:
let X be a random variable with mean 0, variance σ^2, then for $a > 0$,

$$P(X \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

proof: $P(X \geq a) \leq P(X^2 \geq a^2) \leq \frac{1}{a^2} E(X^2) = \frac{\sigma^2}{a^2}$ (we have to do something extra)

let $t > 0$

$$P(X \geq a) = P(X + t \geq a + t)$$
$$\leq P((X+t)^2 \geq (a+t)^2)$$
$$\leq \frac{1}{(a+t)^2} E((X+t)^2) \quad \text{(by markov ineq.)}$$
$$= \frac{\sigma^2 + t^2}{(a+t)^2} \quad (E(X) = 0)$$

now $f(t) = \frac{\sigma^2 + t^2}{(a+t)^2}$

Exe: find $t_0 > 0$ where $f(t)$ attains its minimum, then plug it in $f(t)$ to find $P(X \geq a)$ bound.

$$\left(\frac{df}{dt} = \frac{2t}{\cancel{(a+t)^2}} + (\sigma^2 + t^2)(-\cancel{2})\left(\frac{1}{(a+t)^3}\right) = 0\right)$$

$$t - \frac{\sigma^2 + t^2}{a + t} = 0$$

$$at + \cancel{t^2} = \sigma^2 + \cancel{t^2}$$

$$t_0 = \frac{\sigma^2}{a}$$

$t_0 = \frac{\sigma^2}{a}$

$$f(t_0) = \frac{\sigma^2 + \frac{\sigma^4}{a^2}}{a^2 + \frac{\sigma^4}{a^2} + \frac{\sigma^2}{\cancel{a}} \cdot \cancel{a} \cdot 2} = \frac{\sigma^2\left(\frac{a^2}{a^2} + \frac{\sigma^2}{a^2}\right)}{(\quad)}$$

$$= \frac{\frac{\sigma^2 (\cancel{a^2 + \sigma^2})}{a^2}}{\frac{\cancel{(a^2+\sigma^2)}^2}{\cancel{a^2}}}$$

$$= \frac{\sigma^2}{a^2 + \sigma^2}$$

cor: If $E(X) = \mu$, $\text{var}(X) = \sigma^2$, then for $a > 0$,

$$P(X - \mu \geq a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

and, $$P(X - \mu \leq -a) \leq \frac{\sigma^2}{\sigma^2 + a^2}$$

proof as $Y = X - \mu$
$\mu_Y = 0$, we use the same identity here

$$\text{var}Y = E(Y^2) - E(Y)^2$$
$$= E(X-\mu)^2 - (E(\cancel{X - \mu}))^2 \quad 0$$
$$= \text{var}(X)$$

Continious random variables:

Recall:

$(\Omega, \mathcal{F}, P)$ $\quad X : \Omega \longrightarrow \mathbb{R}$

$\{\omega \mid X(\omega) \leq x\} \in \mathcal{F}, \forall x \in \mathbb{R}$

Discrete random variables, X is discrete means that it range is atmost coutable.

$$\{x_1, x_2, \ldots\ldots\}$$

Dist fn: $\quad F(x) = P(X \leq x)$, $x \in \mathbb{R}$

$$F : \mathbb{R} \rightarrow [0,1]$$

Prop:

1. $F \uparrow$
2. F is right cont
3. F : left limit exist
4. $\lim_{x \to \infty} f(x) = 1$

   $\lim_{x \to -\infty} f(x) = 0$

Note: $P(X = x) = F(x) - \lim_{y \uparrow x} F(y)$

If $P(X = x) > 0$, then F is not continious at x

If X discret with range $\{x_1, x_2, \ldots\ldots\}$ and if $P(X = x_i) > 0$, then F is not cont. at x_i.

For discrete r.v there will be atleast finitely many x_i, s.t $P(X = x_i) > 0$, since $\sum_{i=1}^{\infty} P(X = x_i) = 1$

Therefore, the distribution f^n of a discrete r.v is discontinious at least at finitely many points.

Defn: (Continious random Variable)

A random variable X is called continious r.v if its dist. f^n is continious at all $x \in \mathbb{R}$. Equivalently, X is cont random v if $P(X = x) = 0$ $\forall x \in \mathbb{R}$.

Note: If X is cont r.v then range of X cannot be coutable. Then Ω cannot be coutable.

To define a cont. rv we have to construct a prob. space $(\Omega, \mathcal{F}, P)$ s.t Ω is uncoutable.

$$\Omega = \mathbb{R}, (a,b), [0,1], (0,1)$$

if $\mathcal{F} = \mathcal{P}(\mathbb{R})$ $\rightarrow$ then the issue is we cannot have coutable sum of P property

If we take $\mathcal{F}$ as $\mathcal{P}(\mathbb{R})$, it is not posible to define P on $\mathcal{P}(\mathbb{R})$ as $P : \mathcal{F} \longrightarrow [0,1]$

$P(\mathbb{R})$ is too big to define P on it.

Borel σ-field:

French mathematician borel observed it and identified that we can work with a smaller σ-field on $\mathbb{R}$.

Defn: (Borel σ-field)
A sigma field generated by open intervals of $\mathbb{R}$.

Notation: $\mathcal{B}(\mathbb{R}) = \sigma\{(a,b) \mid -\infty<a<b<\infty;\ a,b \in \mathbb{R}\}$

($\mathcal{C}$ is a collection of subsets of Ω,
σ-field by $\mathcal{C}$ is intersections of all σ-fields containg $\mathcal{C}$
$\sigma(\mathcal{C}) = \bigcap_{\mathcal{C}\in\mathcal{F}} \mathcal{F}$ where $\mathcal{F}$ is a sigma field)

let's see the sets sitting inside $\mathcal{B}(\mathbb{R})$

Defn: (Borel set) An element in $\mathcal{B}(\mathbb{R})$ is called Borel set

Note:
① $(a,b) \in \mathcal{B}(\mathbb{R})$

② $(a,\infty) \in \mathcal{B}(\mathbb{R})$
as $(a,\infty) = \bigcup_{i=1}^{\infty}(a, a+i) \leftarrow$ countable union $\in \mathcal{F}$ as (properties of σ-field used)

③ $(-\infty, a) \in \mathcal{B}(\mathbb{R})$

④ $[a,b] \in \mathcal{B}(\mathbb{R})$
as $[a,b] = \bigcap_{n=1}^{\infty}\left(a-\frac{1}{n}, b+\frac{1}{n}\right) \in \mathcal{F}$

⑤ $\{a\} = [a,a] \in \mathcal{B}(\mathbb{R})$, similar idea
$\{a\} = \bigcap_{n=1}^{\infty}\left(a-\frac{1}{n}, a+\frac{1}{n}\right)$

⑥ finite subset of $\mathbb{R}$ $\in \mathcal{B}(\mathbb{R})$

⑦ Countable subsets $\in \mathcal{B}(\mathbb{R})$ ($\mathbb{N}, \mathbb{Q}, \mathbb{Z}$, etc $\in \mathcal{B}(\mathbb{R})$ and $\mathbb{Q}^c \in \mathcal{B}(\mathbb{R})$ $= \mathbb{R}\setminus\mathbb{Q}$)

⑧ $[a,b) \in \mathcal{B}(\mathbb{R})$
$(a,b] \in \mathcal{B}(\mathbb{R})$

Exe: suppose $\mathcal{C}_1 \subset \mathcal{C}_2$ where $\mathcal{C}_1$ and $\mathcal{C}_2$ are colletion of subsets of Ω, then $\sigma(\mathcal{C}_1) \leq \sigma(\mathcal{C}_2)$

$\sigma(\mathcal{C}_1) = \bigcap_{\mathcal{C}_1\in\mathcal{F}} \mathcal{F}$
$\sigma(\mathcal{C}_2) = \bigcap_{\mathcal{C}_2\in\mathcal{F}} \mathcal{F}$
$\rightarrow$ as $\mathcal{C}_1 \subseteq \mathcal{C}_2$
$\mathcal{C}_1 \in \mathcal{F}$
$\Rightarrow \mathcal{C}_2 \in \mathcal{F}$

$\underbrace{\{\mathcal{F} \mid \mathcal{F} \text{ is } \sigma\text{-field and } \mathcal{C}_2 \subset \mathcal{F}\}}_{T_1} \subseteq \overbrace{\{\mathcal{F} \mid \mathcal{F} \text{ is a } \sigma\text{-field and } \mathcal{C}_1 \in \mathcal{F}\}}^{T_2}$

$\sigma(\mathcal{C}_1) = \bigcap_{\mathcal{F}\in T_2} \mathcal{F} \subseteq \bigcap_{\mathcal{F}\in T_1} \mathcal{F} = \sigma(\mathcal{C}_2)$

Ex- X = No of items produced in a factory in a day
suppose $E(X) = 50$
(a) give bound on $P(X \geq 75)$
(b) $\operatorname{var} X = 25$, get bound on $P(40 < X < 60)$

Now $P(X \geq a) \leq \frac{E(X)}{a}$

then for $E(X) = 50$
$a = 75$

$$P(X \geq 75) \leq \frac{\cancel{50}\,2}{\cancel{75}\,3} = \frac{2}{3}$$

for $P(40 < X < 60) = P(-10 < X - 50 < 10)$
$= P(|X - 50| < 10)$

opp of $P(|X - 50| \geq 10) \leq \frac{\sigma^2}{a^2}$
$= \frac{25}{10.10}$
$= \frac{1}{4}$

$1 - P(|X - 50| < 10) \leq \frac{1}{4}$

$\frac{3}{4} \leq P(|X - 50| < 10)$

4th oct:

Borel sigma fields:

$$\mathcal{B}(\mathbb{R}) = \sigma\{(a,b) \mid -\infty < a < b < \infty\}$$ (Borel sigma fields)

Any set in $\mathcal{B}(\mathbb{R})$ is called borel set

we saw that
1. (a,b)
2. $[a,b)$
3. $(a,b]$
4. $[a,b]$
5. $(-\infty, a]$
6. $\mathbb{Q}$
7. $\mathbb{Z}$
8. $\mathbb{N}$
9. $\mathbb{R}\setminus\mathbb{Q}$

all in Borel σ field of $\mathbb{R}$ or $\mathcal{B}(\mathbb{R})$

Exe: Show that cantor sets is a borel set. → done down

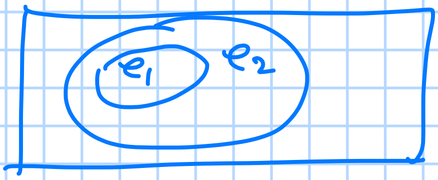
Prop: If $\mathcal{E}_1 \subseteq \mathcal{E}_2$, where $\mathcal{E}_1$ and $\mathcal{E}_2$ are collection of subsets of Ω then $\sigma(\mathcal{E}_1) \subseteq \sigma(\mathcal{E}_2)$

proof:

$$A = \{\mathcal{F} : \mathcal{E}_1 \subset \mathcal{F} \text{ and } \mathcal{F} \text{ is a } \sigma\text{-field}\}$$

$$B = \{\mathcal{F} : \mathcal{E}_2 \subset \mathcal{F} \text{ and } \mathcal{F} \text{ is a } \sigma\text{-field}\}$$

then $B \subseteq A$



$$\sigma(\mathcal{E}_1) = \bigcap_{\mathcal{F}\in A} \mathcal{F}$$

$$\sigma(\mathcal{E}_2) = \bigcap_{\mathcal{F}\in B} \mathcal{F}$$

as $B \subseteq A$

$$\bigcap_{\mathcal{F}\in A} \mathcal{F} \subseteq \bigcap_{\mathcal{F}\in B} \mathcal{F}$$

$$\sigma(\mathcal{E}_1) \subseteq \sigma(\mathcal{E}_2)$$

Theorm:
$$\begin{aligned}\mathcal{B}(\mathbb{R}) &= \sigma\{(a,b] : -\infty < a \leq b < \infty\} \\ &= \sigma\{[a,b) : -\infty < a \leq b < \infty\} \\ &= \sigma\{[a,b] : -\infty < a \leq b < \infty\} \\ &= \sigma\{(-\infty, a] ; a \in \mathbb{R}\}\end{aligned}$$

proof:

$$\mathcal{B}(\mathbb{R}) = \sigma\{(a,b) \mid -\infty < a < b < \infty\}$$

Now there $(a,b] \in \mathcal{B}(\mathbb{R})$

also

$$(a,b] = \bigcap_{n=1}^{\infty}\left(a, b+\frac{1}{n}\right) \in \mathcal{B}(\mathbb{R})$$

so $(a,b] \in \mathcal{B}(\mathbb{R})$

$$\{(a,b] \mid -\infty < a \leq b < \infty\} \subseteq \mathcal{B}(\mathbb{R})$$

$\sigma\{[a,b] \mid -\infty < a \leq b < \infty\} \subseteq \sigma(\mathcal{B}(\mathbb{R})) = \mathcal{B}(\mathbb{R})$

as $\mathcal{B}(\mathbb{R})$ is a σ-field
$\sigma(\mathcal{B}(\mathbb{R})) = \mathcal{B}(\mathbb{R})$

$\underbrace{\sigma\{[a,b]}_{=\mathcal{F}_1} \mid -\infty < a \leq b < \infty\} \subseteq \mathcal{B}(\mathbb{R})$

To show: $(a,b) \in \mathcal{F}_1$

as
$\mathcal{F}_1 = \sigma(\{[a,b] \mid -\infty < a \leq b < \infty\})$

$(a,b) = \bigcup_{n=1}^{\infty} [a, b-\frac{1}{n}] \in \mathcal{F}_1$

so $(a,b) \in \mathcal{F}_1$

now $\{(a,b) \mid -\infty < a < b < \infty\} \subseteq \mathcal{F}_1$

$\Rightarrow \sigma\{(a,b) \mid -\infty < a < b < \infty\} \subseteq \sigma(\mathcal{F}_1)$

$\Rightarrow \mathcal{B}(\mathbb{R}) \subseteq \mathcal{F}_1$

$\therefore \mathcal{B}(\mathbb{R}) \subseteq \mathcal{F}_1$
and $\mathcal{B}(\mathbb{R}) \supseteq \mathcal{F}_1$
$\Rightarrow \mathcal{B}(\mathbb{R}) = \mathcal{F}_1$

proof done down

Theorm: $\mathcal{B}(\mathbb{R}) = \sigma\{\text{open sets of } \mathbb{R}\}$

(open sets: $\forall x \in S$, $\exists r > 0$ s.t $B(x,r) \subseteq S$ or $(x-r, x+r) \subseteq S$)

proof:

As $(a,b) \in \sigma\{\text{open sets of } \mathbb{R}\}$

$\{(a,b) \mid -\infty < a < b < \infty\} \subseteq \sigma\{\text{open sets of } \mathbb{R}\}$

$\Rightarrow \sigma\{(a,b) \mid -\infty < a < b < \infty\} \subseteq \sigma\{\text{open sets of } \mathbb{R}\}$

$\Rightarrow \mathcal{B}(\mathbb{R}) \subseteq \sigma\{\text{open sets of } \mathbb{R}\}$

now as every open set in $\mathbb{R}$ can be written as union of countable disjoint intervals.

$\underset{\text{open set}}{U} = \bigcup_{n=1}^{\infty} I_n$ → open intervals
and
$I_j \cap I_k = \phi$ for $j \neq k$

$U \in \mathcal{B}(\mathbb{R})$ as each $I_n \in \mathcal{B}(\mathbb{R})$

$\sigma(U) \subseteq \sigma(\mathcal{B}(\mathbb{R}))$

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

Exe: show that $X : (\Omega, \mathcal{F}) \longrightarrow \mathbb{R}$ is a random variable iff

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

proof: X is a r.v $\Rightarrow \{\omega \mid X(\omega) \leq x\} \in \mathcal{F} \ \forall x \in \mathbb{R}$

($\Leftarrow$) $X : \Omega \to \mathbb{R}$ and

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

To show: $\{\omega \mid X(\omega) \leq x\} \in \mathcal{F}$

$\forall x \in \mathbb{R}$

$X^{-1}(-\infty, x] = \{\omega \mid X(\omega) \leq x\}$

as $(-\infty, x] \in \sigma(\{(a,b) \mid -\infty < a < b < \infty\})$
$= \mathcal{B}(\mathbb{R})$

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

putting $B = (-\infty, x]$ we get

$\forall x \in \mathbb{R}$

$X^{-1}(-\infty, x] \in \mathcal{F}$

or $\{\omega \mid X(\omega) \leq x\} \in \mathcal{F} \quad \forall x \in \mathbb{R}$

$\therefore X$ is a random variable

($\Rightarrow$) X is a random variable. That is

$X^{-1}(-\infty, x] \in \mathcal{F} \quad \forall x \in \mathbb{R}$

$\Psi = \{A \mid X^{-1}(A) \in \mathcal{F}\}$

prove Ψ is a σ-field

(Ψ is a σ-field because it satisfies 3 properties)

as X is a random variable

$(-\infty, a] \in \Psi$

for any a in $\mathbb{R}$.

$\Rightarrow \{(-\infty, a] \mid a \in \mathbb{R}\} \subseteq \Psi$

$\Rightarrow \sigma\{(-\infty, a] \mid a \in \mathbb{R}\} \subseteq \sigma(\Psi) = \Psi$

$\Rightarrow \sigma\{(-\infty, a] \mid a \in \mathbb{R}\} = \mathcal{B}(\mathbb{R}) \subseteq \sigma(\Psi) = \Psi$

or $\mathcal{B}(\mathbb{R}) \subseteq \Psi$

so $\forall B \in \mathcal{B}(\mathbb{R})$

$\Rightarrow B \in \Psi$

or $X^{-1}(B) \in \mathcal{F} \quad \forall B \in \mathcal{B}(\mathbb{R}$

Exe: suppose $f : \mathbb{R} \to \mathbb{R}$ is continuous. Show that $f^{-1}(B) \in \mathcal{B}(\mathbb{R})$
$\forall B \in \mathcal{B}(\mathbb{R})$

$\mathcal{G} = \{A \mid f^{-1}(A) \in \mathcal{B}(\mathbb{R})\}$

$\mathcal{G}$ is also σ-field $\to$ show

now as any open set in $\mathbb{R}$ is in $\mathcal{G}$

$\{(a,b) \mid -\infty < a < b < \infty\} \subseteq \mathcal{G}$

$\sigma\{(a,b) \mid -\infty < a < b < \infty\} = \mathcal{B}(\mathbb{R}) \subseteq \mathcal{G}$

so, $\forall \; B \in \mathcal{B}(\mathbb{R})$
$B \in \mathcal{G}$ or

$f^{-1}(B) \in \mathcal{B}(\mathbb{R}) \; \forall \; B \in \mathcal{B}(\mathbb{R})$

Note: $(\Omega = \mathbb{R}, \mathcal{F} = \mathcal{B}(\mathbb{R}), P)$
we need a new probability measure

Lebesgue measure on $\mathbb{R}$:

It's a map λ, from $\mathcal{B}(\mathbb{R}) \longrightarrow [0, \infty)$ s.t

(i) $\lambda((a,b)) = \lambda((a,b]) = \lambda([a,b]) = b-a$

(ii) $\lambda\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \lambda(A_n)$

where $A_n \in \mathcal{B}(\mathbb{R})$ and $A_i \cap A_j = \phi$, $i \neq j$

Such a map is called Lebsgue measure

($P : \mathcal{F} \to [0,1]$
$P(\Omega) = 1$
$P(\cup A_n) = \Sigma P(A_n)$

can be seen very simiar to Lebsgue measure)

$\mathcal{S} = \{(a,b] \mid -\infty < a \leq b < \infty\} \cup \{(a,\infty) \mid -\infty < a < b\}$

closed under finite intersection and compliment

then $\mathcal{S}$ called semi algebra
$\downarrow$
Algebra | field
$\downarrow$
σ-Algebra | field

unique extension called: Carathedory extension theorm

Exe: Show that Cantor sets is a borel set.

$\mathbb{C} = \bigcap_{k=1}^{\infty} C_k$ (a) $\{C_k\}_{k=1}^{\infty}$ is dec seq of closed sets

(b) C_k is disjoint union of 2^k closed sets each of length $1/3^k$

Cantor set

To show: $\mathbb{C} \in \mathcal{B}(\mathbb{R})$

now as

$\mathbb{C} = \bigcap_{k=1}^{\infty} C_k$ of disjoint 2^k closed sets

$(\mathbb{C})^c = \bigcup_{k=1}^{\infty} C_k^c$ of open 2^k sets of length $\left(2/3\right)^k$

so as $C_k^c \in \mathcal{B}(\mathbb{R})$

$\Rightarrow \bigcup_{k=1}^{\infty} C_k^c \in \mathcal{B}(\mathbb{R})$ (countable union)

$\Rightarrow \left(\bigcup_{k=1}^{\infty} C_k^c\right)^c \in \mathcal{B}(\mathbb{R})$

$\Rightarrow \bigcap_{k=1}^{\infty} C_k \in \mathcal{B}(\mathbb{R})$ $\therefore$ $\mathbb{C} \in \mathcal{B}(\mathbb{R})$ is a borel set

Ine : $\mathcal{B}(\mathbb{R}) = \sigma\{[a,b) : -\infty < a < b < \infty\}$

$= \sigma\{[a,b] : -\infty < a \leq b < \infty\}$

$= \sigma\{(-\infty, a] ; a \in \mathbb{R}\}$

proof : now $[a,b] \in \mathcal{B}(\mathbb{R})$

as $(a,b] \in \mathcal{B}(\mathbb{R})$

and $[a,b] \in \mathcal{B}(\mathbb{R})$

then $[a,b] \in \mathcal{B}(\mathbb{R})$

now $\{[a,b] \mid -\infty < a \leq b < \infty\} \subseteq \mathcal{B}(\mathbb{R})$

$\Rightarrow \sigma\{[a,b] \mid -\infty < a \leq b < \infty\} \subseteq \mathcal{B}(\mathbb{R})$ —— ①

similarly $(a,b) \in \sigma\{[a,b] \mid -\infty < a \leq b < \infty\}$
then $= \mathcal{F}_2$

as $[a,b] \in \mathcal{F}_2$

then $(a,b) \in \mathcal{F}_2$

similarly other 2.

8th Oct:

$\mathcal{B}(\mathbb{R})$ * $m : \mathcal{B}(\mathbb{R}) \longrightarrow [0,\infty)$
* $m((a,b]) = m([a,b)) = m((a,b]) \cdots = b-a$
* $m(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} m(A_n)$ $\quad A_i \cap A_j = \phi \quad i \neq j$

$\Omega = (0,1)$

$\mathcal{B}((0,1)) = \sigma\{(a,b) \mid (a,b) \subseteq (0,1)\}$

Exercise: let $B \in \mathcal{B}(\mathbb{R}) \Rightarrow B \cap (0,1) \in \mathcal{B}((0,1))$

Theorem: let $\Omega_0 \subseteq \Omega$

(i) If $\mathcal{F}$ is a σ-field of subsets of Ω then $\mathcal{F}_0 = \{A \cap \Omega_0 \mid A \in \mathcal{F}\}$ is a sigma field}

(ii) suppose $\mathcal{C}$ is a class of subsets of Ω and $\mathcal{F} = \sigma(\mathcal{C})$

$\mathcal{C}_0 = \{A \cap \Omega_0 \mid A \in \mathcal{C}\}$

then $\sigma(\mathcal{C}_0) = \{A \cap \Omega_0 \mid A \in \sigma(\mathcal{C})\}$

Result: $\mathcal{B}(\mathbb{R}) = \sigma\{\underbrace{(a,b) \mid -\infty < a < b < \infty}_{\mathcal{C}}\}$

$\mathcal{B}((0,1)) = \sigma\{(a,b) \mid 0 < a < b < 1\}$

$= (0,1) \cap \mathcal{C} = \mathcal{C}_0$

Notation: $\sigma(\mathcal{C}_0) = \{A \cap \Omega_0 \mid A \in \sigma(\mathcal{C})\}$

$\sigma(\mathcal{C}_0) = \sigma(\mathcal{C}) \cap \Omega_0$

$\sigma(\mathcal{C}_0) = \{A \cap (0,1) \mid A \in \mathcal{B}(\mathbb{R})\}$

so $\mathcal{B}((0,1)) = \sigma(\mathcal{C}_0) = \{A \cap (0,1) \mid A \in \underbrace{\mathcal{B}(\mathbb{R})}_{= \sigma(\mathcal{C})}\}$

Corr: $\mathcal{B}((0,1)) = \{A \cap (0,1) \mid A \in \mathcal{B}(\mathbb{R})\}$

$\mathcal{B}((0,1)) = \mathcal{B}(\mathbb{R}) \cap (0,1)$

Proof: (i) Exercise $\rightarrow$ done, see down

(ii) $\sigma(\mathcal{C}_0) = \sigma(\mathcal{C}) \cap \Omega_0$

Note
$\sigma(\mathcal{C}) \cap \Omega_0$ is a σ-field from (i)

now, $\mathcal{C}_0 \subseteq \sigma(\mathcal{C}) \cap \Omega = \{A \cap \Omega_0 \mid A \in \sigma(\mathcal{C})\}$

$A \in \mathcal{C} \subseteq \sigma(\mathcal{C})$

$A \cap \Omega_0 \in \sigma(\mathcal{C}) \cap \Omega_0$

so, $\mathcal{C}_0 \subseteq \sigma(\mathcal{C}) \cap \Omega_0$

as

$\mathcal{C}_0 = \{A \cap \Omega_0 \mid A \in \mathcal{C}\}$

$\Rightarrow \sigma(\mathcal{C}_0) \subseteq \sigma(\mathcal{C}) \cap \Omega_0$

now to show $\sigma(\mathcal{C}) \cap \Omega_0 \subseteq \sigma(\mathcal{C}_0)$

define $\mathcal{G} = \{A \subseteq \Omega \mid A \cap \Omega_0 \in \sigma(\mathcal{C}_0)\}$

$\mathcal{G}$ is a σ-field (trivial)

<u>Note</u>: $\mathcal{C} \subseteq \mathcal{G}$ since if $A \in \mathcal{C}$ then $A \cap \Omega_0 \in \mathcal{C}_0 \subseteq \sigma(\mathcal{C}_0)$

$\mathcal{C} \subseteq \mathcal{G}$ then

$$\sigma(\mathcal{C}) \subseteq \sigma(\mathcal{G}) = \mathcal{G}$$

$$\sigma(\mathcal{C}) \subseteq \mathcal{G}$$

$\therefore$ if $A \in \sigma(\mathcal{C})$ then

$A \cap \Omega_0 \in \sigma(\mathcal{C}_0)$ By construction of $\mathcal{G}$

$\forall\, A \in \sigma(\mathcal{C})$, $A \cap \Omega_0 \in \sigma(\mathcal{C}_0)$

or $\sigma(\mathcal{C}) \cap \Omega_0 \subseteq \sigma(\mathcal{C}_0)$

$\therefore \quad \sigma(\mathcal{C}) \cap \Omega_0 = \sigma(\mathcal{C}_0)$

<u>Cor</u>: $\mathcal{B}((0,1]) = \mathcal{B}(\mathbb{R}) \cap (0,1]$
$\mathcal{B}([0,1]) = \mathcal{B}(\mathbb{R}) \cap [0,1]$

<u>Example of Prob. spaces</u>:

(i) $\Omega = (0,1)$ $\mathcal{F} = \mathcal{B}((0,1))$ $P : \mathcal{F} \to [0,1]$

$P(A) = m(A)$, for $A \in \mathcal{F}$

(ii) $\Omega = (a,b)$ $\mathcal{F} = \mathcal{B}((a,b))$ $P : \mathcal{F} \to [0,1]$

$$P(A) = \frac{1}{b-a}\, m(A)$$

<u>Exe</u>: Find $m(\{1\})$, $m\left(\{\frac{1}{n} \mid n \in \mathbb{N}\}\right)$, $m\left(\{x \mid |x-n| < \frac{1}{2n} \text{ for some } n \in \mathbb{N}\}\right)$

$m(C)$ ← Cantor set

$$m(\{1\}) = m([1,1]) = 1-1 = 0$$

$$m\left(\{\tfrac{1}{n} \mid n \in \mathbb{N}\}\right) = m\left(\{[\tfrac{1}{n}, \tfrac{1}{n}] \; \forall n \in \mathbb{N}\}\right) = \sum \frac{1}{n} - \frac{1}{n} = 0$$

$$\text{or} = \sum m\left(\{\tfrac{1}{n}\}\right) = 0$$

$$m\left(\{x \mid |x-n| < \tfrac{1}{2n} \text{ some } n \in \mathbb{N}\}\right)$$

$$= m\left(\bigcup_{n=1}^{\infty} \left(n - \tfrac{1}{2n}, n + \tfrac{1}{2n}\right)\right)$$

$= \sum \frac{1}{2^{n-1}} = 1 + \frac{1}{2} + \frac{1}{4} + \cdots = 2$, for $m(\mathbb{C}) = m(\bigcap_{n=1}^{\infty} C_k)$
$\leq m(C_k) = (2/3)^k \to 0$

Recall : suppose p is a pmf then $\exists (\Omega, \mathcal{F}, P)$ and a random variable $X : \Omega \longrightarrow \mathbb{R}$

s.t pmf of X is p.

$\Omega = \mathbb{N}$
$\mathcal{F} = P(\mathbb{N})$
$P(\{n\}) = p(x_n)$

$X(n) = x_n$

now $F(x) = P(X \leq x)$ $\quad F : \mathbb{R} \longrightarrow \mathbb{R}$
(i) F non-decreasing
(ii) F is right cont
(iii) left limit exist
(iv) $\lim_{x \to \infty} F(x) = 1$
$\lim_{x \to -\infty} F(x) = 0$

Theorem: suppose F is a function, that satisfies cond^n (i), (ii), (iii) and (iv)
Then $\exists$ a probability space $(\Omega, \mathcal{F}, P)$ and r.v
$X : \Omega \to \mathbb{R}$ s.t dist function of X is F.

$\Omega = (0,1)$
$\mathcal{F} = \mathcal{B}((0,1))$
$P(A) = m(A)$, $A \in \mathcal{F}$

Then we can define X appropriately

Que : let $\Omega_0 \subseteq \Omega$, **Prove** :
If $\mathcal{F}$ is a σ-field of subsets of Ω then $\mathcal{F}_0 = \{A \cap \Omega_0 \mid A \in \mathcal{F}\}$
is a sigma field}

As $\mathcal{F}$ is a σ-field $\quad \mathcal{F}_0 = \{A \cap \Omega_0 \mid A \in \mathcal{F}\}$

as if $A_1 \cap \Omega_0 \in \mathcal{F}_0$
$A_2 \cap \Omega_0 \in \mathcal{F}_0$
then all three satisfied

11^{th} oct :

Suppose F is the distribution f^n of a r.v X

(i) $\lim_{x \to \infty} F(x) = 1 \qquad \lim_{x \to -\infty} F(x) = 0$

(ii) F is non-decreasing

(iii) $\lim_{y \downarrow x} F(y) = F(x)$, right continuous

Theorem: suppose $F : \mathbb{R} \to [0,1]$ and F satisfies (i), (ii), (iii).
Then $\exists$ a prob. space $(\Omega, \mathcal{F}, P)$ and a random variable $X : \Omega \to \mathbb{R}$
s.t $F_X(x) = P(X \leq x) = F(x) \quad \forall x \in \mathbb{R}$

proof:

$\Omega = (0,1)$
$\mathcal{F} = \mathcal{B}((0,1))$

$P(A) = m(A)$ for $A \in \mathcal{B}((0,1))$
lebesgue measure

suppose F is strictly inc and cont
$X : \Omega \to \mathbb{R}$
$X(\omega) = F^{-1}(\omega)$
special case

Define $X(\omega) = \sup\{ y \mid F(y) < \omega \}$

$\Leftrightarrow$
$X(\omega) = \inf\{ y \mid F(y) \geq \omega \}$ (generalized inverse of F)

claim: $\{\omega \in \underset{(0,1)}{\Omega} \mid X(\omega) \leq x\} = \{\omega \in \Omega \mid \omega \leq F(x)\}$

i.e
$P(A = \{\omega \in \Omega \mid X(\omega) \leq x\})$
$= P(B = \{\omega \in \Omega \mid \omega \leq F(x)\})$

or
$P(A) = P(B) =$ lebuge measure $(A) = m(B)$

where
$X(\omega) \leq x \Leftrightarrow \omega \leq F(x)$
where $X(\omega)$ is from our own defination

we $X(\omega) \leq x$ makes sense

$\omega \leq F(x)$ mean that
$P((0, F(x)]) = P(\{\omega \in \Omega \mid \omega \leq F(x)\})$
$= m((0, F(x)]) = F(x)$
or $P(\{X \leq x\}) = F(x)$

$\therefore$ if our claim is true then $F(x) = P(\{X \leq x\})$

proof of claim:

X is a r.v then $X : \Omega \to \mathbb{R}$

$\{\omega \mid X(\omega) \leq x\} \in \mathcal{F} = \mathcal{B}((0,1))$

$$X(\omega) = \sup\{y \mid F(y) < \omega\}$$

to show $B \subseteq A$

let $\omega \in B = \{\omega \in \Omega \mid \omega \leq F(x)\}$

$F(x) \geq \omega$

$\Rightarrow x \notin \{y \mid F(y) < \omega\}$

$\Rightarrow \sup\{y \mid F(y) < \omega\} \leq x$ ($\because F$ is non-decreasing)

$\Rightarrow X(\omega) \leq x$ (from defn)

$\Rightarrow \omega \in A$

to show $A \subseteq B$
or
$B^c \subseteq A^c$

let $\omega \in B^c$

$\Rightarrow F(x) < \omega$

$F(x)$ ω

as F is right cont

$\exists \varepsilon > 0$ s.t

$F(x) \leq F(x+\varepsilon) < \omega$

$\sup\{y \mid F(y) < \omega\} \geq x + \varepsilon > x$

$\Rightarrow \sup\{y \mid F(y) < \omega\} > x$

strictly greater (as A is $X(\omega) \leq x$)
so A^c is $X(\omega) > x$

$\Rightarrow X(\omega) > x$

$\Rightarrow x < X(\omega)$

$\Rightarrow x \in A^c$

$\therefore B^c \subseteq A^c$

or $A = B$

Algebra : collection of subsets of Ω which satisfies the following:

($\mathcal{A}$)

(i) $\phi \in \mathcal{A}$

(ii) $A \in \mathcal{A} \Rightarrow A^c \in \mathcal{A}$

(iii) $A \in \mathcal{A}$, $B \in \mathcal{A} \Rightarrow A \cup B \in \mathcal{A}$

σ-Algebra: (i), (ii) there
($\mathcal{F}$)
(iii) $A_i \in \mathcal{F}$, $\forall i \in C \leftarrow$ countable set
then

$$\bigcup_{i \in C} A_i \in \mathcal{F}$$

semi-Algebra: collection of subsets of Ω which satisfies the following
($\mathcal{S}$)
(i) If $A, B \in \mathcal{S} \Rightarrow A \cap B \in \mathcal{S}$
(ii) $A \in \mathcal{S}$, then A^c can be expressed as finite disjoint union of sets in $\mathcal{S}$.

Example : $\Omega = \mathbb{R}$

$$\mathcal{S} = \phi \cup \{(a,b] \mid -\infty \leq a < b < \infty\}$$

$\left(\cup \{(a,\infty) \mid -\infty \leq a < \infty\}\right.$ is also preset$\left.\right)$

Exercise : Check $\mathcal{S} = \phi \cup \{(a,b] \mid -\infty \leq a < b < \infty\}$ is a semi algebra
$\rightarrow$ done down

Def^n : A set fn $\mu : \mathcal{S} \rightarrow [0,\infty]$ is said to be a measure if it follows:
(a) $\mu(A) \geq \mu(\phi) = 0 \quad \forall A \in \mathcal{S}$
(b) if $A_1, A_2, \dots \in \mathcal{S}$, $A_i \cap A_j = \phi$
and
$\bigcup_{i=1}^{\infty} A_i \in \mathcal{S}$, then

$$\mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Def^n : A set f^n: $\mu : \mathcal{A} \rightarrow [0,\infty]$ is a measure:
(i) $\mu(A) \geq \mu(\phi) = 0$, $\forall A \in \mathcal{A}$
(ii) $A_1, A_2, \dots \in \mathcal{A}$, $A_i \cap A_j = \phi$
for $i \neq j$
and
$\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$

$$\text{then } \mu\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mu(A_n)$$

Def^n : A measure μ on $\mathcal{A}$ is called σ-finite if $\exists \{A_n\}$, $A_n \in \mathcal{A}$
s.t
$\bigcup_{n=1}^{\infty} A_n = \Omega$ and
$\mu(A_n) < \infty \ \forall n$

suppose $\mathcal{J}$ is a semi-Algebra

Define : Algebra generated by $\mathcal{J}$, $\mathcal{A}(\mathcal{J}) = \{ \bigcup_{i=1}^{m} A_i \mid A_i \in \mathcal{J}, i=1,2,\dots m, m \in \mathbb{N} \}$

Theorm : (Extension theorm) let $\mathcal{J}$ be a semi-Algebra of subsets of Ω. And μ be a measure on $\mathcal{J}$.

Then there is a unique extension $\bar{\mu}$ of μ to $\mathcal{A}(\mathcal{J})$.

measure on $\mathcal{A}(\mathcal{J})$

$\bar{\mu} : \mathcal{A}(\mathcal{J}) \longrightarrow [0,\infty]$

further, if $\bar{\mu}$ is a σ-finite on $\mathcal{A}(\mathcal{J})$ then there is a unique extension μ^* of $\bar{\mu}$ to $\sigma(\mathcal{J})$ where μ^* is a measure on $\sigma(\mathcal{J})$

(Note : second part of extension theorm is known as caratheodory's extension theorm)

Ref : (i) Measure theory, Athreya and Lahiri (TRIM series)

(ii) Real Analysis, Rudin

(iii) Probability, Rick Durret

proof : exercise

$(\Omega, \mathcal{F}, \mu) \to$ measure space

$(\Omega, \mathcal{F}, P) \to$ probability space

Completeness of measure space :

$(\Omega, \mathcal{F}, \mu) \to$ is called complete if following holds :
suppose $A \in \mathcal{F}$ and $\mu(A) = 0$, then $B \in \mathcal{F}$ $\forall B \subseteq A$

Exercise : Prove cantor sets are not countable (Hint trinerise sets)
$\to$ do

Facts :

① $\mathcal{C}$ = cantor set $\in \mathcal{B}(\mathbb{R})$
$m(\mathcal{C}) = 0$
$\hookrightarrow$ cantor set

② card of $\mathcal{B}(\mathbb{R})$ = card of $\mathbb{R}$

③ $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m)$ is not complete ($\because$ $\mathcal{C}$ is not-countable then as $\mathcal{C} \in \mathcal{B}(\mathbb{R})$ if complete then $\mathcal{P}(\mathcal{C}) \subseteq \mathcal{B}(\mathbb{R})$ but $|\mathcal{P}(\mathcal{C})| > |\mathcal{B}(\mathbb{R})|$ $*$)

④ $(\Omega, \mathcal{F}, \mu) \to$ measure space

$\mathcal{N} = \{B \mid B \subseteq N \text{ where } N \in \mathcal{F} \text{ and } \mu(N) = 0\}$

$\tilde{\mathcal{F}} = \{A \cup B \mid A \in \mathcal{F} \text{ and } B \in \mathcal{N}\}$

completion of $\mathcal{F}$

or $(\Omega, \tilde{\mathcal{F}}, \mu)$ is complete (completion of $(\Omega, \mathcal{F}, \mu)$)

Ⓢ $(\mathbb{R}, \mathcal{B}(\mathbb{R}), m) \longrightarrow (\mathbb{R}, \widetilde{\mathcal{B}(\mathbb{R})}, m)$ is complete

$\widetilde{\mathcal{B}(\mathbb{R})} = \{A \cup B \mid A \in \mathcal{B}(\mathbb{R}), B \in \mathcal{N}\}$

where $\mathcal{N} = \{B \mid B \subseteq N, N \in \mathcal{B}(\mathbb{R}), m(N) = 0\}$

$\underbrace{\widetilde{\mathcal{B}(\mathbb{R})} = \mathcal{L}(\mathbb{R})}_{\text{Lebesgue } \sigma\text{-Algebra}}$

$\underbrace{(\mathbb{R}, \mathcal{L}(\mathbb{R}), m)}_{\text{Lebesque measure space}}$

Cont r.v : Dist fn F is cont

Absolutely cont r.v : $X : \Omega \to \mathbb{R}$, if $\exists f : \mathbb{R} \to \mathbb{R}_{\geq 0}$ s.t

$P(X \leq x) = F(x) = \int_{-\infty}^{x} f(t)\,dt$

f is called the density fn of X.

Note : $f(x) \neq P(X = x) = 0$

Exercise : check $\mathcal{S} = \phi \cup \{(a, b] \mid -\infty \leq a < b < \infty\}$ is a semi algebra

let $A \in \mathcal{S}$ then $A = (a_1, a_2]$
if $B \in \mathcal{S}$ then $B = (b_1, b_2]$

then $A \cap B = (a_1, b_2]$ if they overlap,
else ϕ

as $(a_1, b_2] \in \mathcal{S}$
and
$\phi \in \mathcal{S}$

$\Rightarrow A \cap B \in \mathcal{S}$

now if $A \in \mathcal{S}$ then $\underset{\in \mathcal{S}}{(-\infty, a_1]} \cup \underset{\in \mathcal{S}}{(a_2, \infty)}$

so $A^c \in \mathcal{S}$, $\therefore \mathcal{S}$ is a semi algebra

16th oct :

Recap:

cont. random variable

- Ab. cont r.v
- density funtion of a r.V X

$$F(x) = \int_{-\infty}^{x} f(t)\,dt$$

Note: density funtion is not unique

eg: finite points – changed but integration remains same

cont random vector :

(X, Y)

$$F(x, y) = P(X \le x, Y \le y)$$

$$f(x, y) : \mathbb{R}^2 \longrightarrow \mathbb{R}_{\ge 0}$$

$$F(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} \underbrace{f(s, t)}_{\text{joint density function}}\,ds\,dt$$

- marginal density fn
- conditional density fn
- expectation

change of variable formula :

X , density f

$Y = g(X)$ → density of Y we have to find

$(X_1, X_2) \longrightarrow (g(X_1, X_2), h(X_1, X_2))$

$g, h : \mathbb{R}^2 \longrightarrow \mathbb{R}$

& $f(X_1, X_2) \rightarrow$ joint density of (X_1, X_2)

→ will find joint density of (Y_1, Y_2)

(Note: Tutorial will cover all question)

Convergence of random variables :

suppose $\{f_n\}_{n \ge 1}$ is a seq of functions

$f_n : [0,1] \longrightarrow \mathbb{R}$

Pointwise convergence: f_n converges to some funtion f pointwise

$(f : [0,1] \longrightarrow \mathbb{R})$

if $f_n(x) \longrightarrow f(x)$ as $n \rightarrow \infty$ $\forall x \in [0,1]$

$\{ f : f : [0,1] \to \mathbb{R} \} = \mathcal{F}$

$\| \cdot \| \to$ appropriate norm on $\mathcal{F}$

$\{f_n\} \subseteq \mathcal{F}$, $f \in \mathcal{F}$

$\|f_n - f\| \to 0$

$(\Omega, \mathcal{F}, P)$ $X : \Omega \to \mathbb{R}$

$\{ \omega \in \Omega \mid X(\omega) \leq x \} \in \mathcal{F}$

- Pointwise convergence here in probability ($\{X_n\}$ s.t $\|X_n - X\| \to 0$) sequence

$\Downarrow$

- Almost sure convergence

$\Downarrow$

- conv in probability

$\Downarrow$

- conv in distribution

} important conv in probability

Important theorms to cover :

strong law of large numbers

weak law of large numbers

central limit theorm

Almost sure convergence :

let $\{X_n\}_{n \geq 1}$ be a sequence of r.v's defined on a prob space $(\Omega, \mathcal{F}, P)$ and let X be a random variable defined on $(\Omega, \mathcal{F}, P)$. we say X_n converges to X almost surely

if $P(\{ \omega \mid \lim_{x \to \infty} X_n(\omega) = X(\omega) \}) = 1$

Example : $f_n(x) \to f(x)$ except at $x = 1/2$ $\to$ $P(X = \frac{1}{2}) = 0$

$\therefore$ not point-wise

$P(\frac{1}{2}) = 0$ $\to$ singleton set in like c-r-v

$B([0,1]) = \mathcal{F}$

$\therefore P(\{ \omega \in \Omega \mid \lim_{n \to \infty} f_n(x) = f(x) \}) = 1$

Note : sometimes if random variable is not given as implicit X form, but given as f or like F, then not possible to check Almost sure convergence

Convergence in probability :

let $\{X_n\}_{n \geq 1}$ and X be a r.v.s defined on $(\Omega, \mathcal{F}, P)$. we say X_n converges to X in probability if $\forall \varepsilon > 0$,

$P(|X_n - X| > \varepsilon) \to 0$ as $n \to \infty$

$$P(|X_n - X| > \varepsilon) = P\left(\{\omega \mid |X_n(\omega) - X(\omega)| > \varepsilon\}\right)$$

Remark : A similar notion of convergence is "convergence in measure" will see in measure theory course. Probability measure will be replaced by a measure μ.

Example : suppose $X \sim Ber(p)$
define
$$X_n(\omega) = \left(1 + \frac{1}{n}\right) X(\omega), \quad n \geq 1$$

$$X_n = \left(1 + \frac{1}{n}\right) X \quad \text{for } n \geq 1$$

$\{X_n\}$ is the sequence of r.v s.t
$$X_n = \left(1 + \frac{1}{n}\right) X$$

$$\{X_n\}_{n \geq 1}$$

now let $\varepsilon > 0$,
$$P(|X_n - X| > \varepsilon)$$
$$= P(|X| > n\varepsilon)$$

$\exists\, n_0$ s.t
$n_0 \varepsilon > 1$

so, $P(|X| > n\varepsilon > 1)$
$$= P(|X| > 1)$$
$$= 0$$

$\therefore \forall\, n \geq n_0 \quad P(|X| > n\varepsilon) = 0$
$$\Rightarrow P\left(|X_n - X| > \varepsilon\right) = 0$$

$\therefore$ conv in probability

Note : Notion of $X_n \to X$ in probability is
$$X_n \xrightarrow{P} X$$

Example : $\{X_n\}$ is a sequence of random variable s.t
$$P(X_n = n) = \frac{1}{n}$$
$$\& \; P(X_n = 0) = 1 - \frac{1}{n}$$

converges to $X \equiv 0$ (const random Variable)
$X(\omega) = 0, \forall\, \omega \in \Omega$

let $\varepsilon > 0$

$$P(|X_n - X| > \varepsilon) = P(|X_n| > \varepsilon) \leq P(X_n = n) = \frac{1}{n} \to 0 \text{ as } n \to \infty$$

if $\varepsilon < \frac{1}{n}$ then

$$X_n \xrightarrow{P} 0$$

<u>Example</u>:

$\{X_n\}_{n \geq 1}$ $\quad P(X_n = 1) = \frac{1}{2}$

$P(X_n = n) = \frac{1}{2}$

$X_n \xrightarrow{P} 1$ NO

as say $\varepsilon = 1/2$ for $n \geq N$

$$P(|X_n - 1| > \frac{1}{2}) = P(X_n - 1 > \frac{1}{2} \text{ \& } X_n - 1 < -\frac{1}{2})$$

$$= P(X_n > \frac{3}{2} \text{ \& } X_n < \frac{1}{2})$$

$$= \frac{1}{2} \quad X_n \not\xrightarrow{P} 1 \text{ (see)}$$

as $N - 1 > \varepsilon$

$$P(|X_n - 1| > \varepsilon) = P(X_n - 1 = \varepsilon) + P(X_n - 1 = \varepsilon + 1) + \cdots = \frac{1}{2}$$

$$= 1/2$$

moreover, we will show that X_n does not converge to any r.v X

<u>Result</u>: Suppose $\{X_n\}_{n \geq 1}$ be a sequence of rv with mean μ $(E(X_n) = \mu, \forall n)$

$\text{Var}(X_n) \to 0$ as $n \to \infty$

<u>proof</u>: let $\varepsilon > 0$

$$P(|X_n - \mu| > \varepsilon) \leq \frac{\text{Var}(X_n)}{\varepsilon^2} \to 0 \text{ as } n \to \infty$$

← Chebyshev's inequality

Result: suppose $\{X_n\}$ be seq of r.v with $E(X_n) = \mu_n$, and $\mu_n \to \mu$ and random variable $Var(X_n) \to 0$ as $n \to \infty$

Then $X_n \xrightarrow{P} \mu$

proof: let $\varepsilon > 0$, $P(|X_n - \mu| > \varepsilon) = P((X_n - \mu)^2 > \varepsilon^2)$

$$\leq \frac{1}{\varepsilon^2} E[(X_n - \mu)^2]$$

where $E[(X_n - \mu)^2]$

$$= E[\{(X_n - \mu_n) + (\mu_n - \mu)\}^2]$$

$$\leq 2E[(X_n - \mu_n)^2 + (\mu_n - \mu)^2]$$

$$= 2[var X_n + (\mu_n - \mu)^2]$$

$$\to 0 \text{ as } n \to \infty$$

Result: ① $X_n \xrightarrow{P} X$, $Y_n \xrightarrow{P} Y \Rightarrow X_n + Y_n \xrightarrow{P} X + Y$

proof: as $\forall \varepsilon > 0$, $n \to \infty$

$$P(|X_n - X| > \varepsilon/2) \to 0$$

& $$P(|Y_n - Y| > \varepsilon/2) \to 0$$

$$P(|X_n + Y_n - (X + Y)| > \varepsilon)$$

$$\leq P(|X_n - X| + |Y_n - Y| > \varepsilon)$$

as $\{|X+Y| > \varepsilon\} \subseteq \{|X| + |Y| > \varepsilon\}$

$P(A \cup B) \leq P(A) + P(B)$

$$\leq P\left(|X_n - X| > \frac{\varepsilon}{2}\right) + P\left(|Y_n - Y| > \frac{\varepsilon}{2}\right)$$

as $\{|X_n - X| + |Y_n - Y| > \varepsilon\} \subseteq \underbrace{\{|X_n - X| > \frac{\varepsilon}{2}\}}_{=A} \cup \underbrace{\{|Y_n - Y| > \frac{\varepsilon}{2}\}}_{=B}$

as $n \to \infty$, $P(A) \to 0$
& $P(B) \to 0$

$$P(|(X_n + Y_n) - (X + Y)| > \varepsilon) \to 0$$

② $X_n - Y_n \xrightarrow{P} X - Y$

proof : now

$\forall \varepsilon > 0$

$$P(|X_n - Y_n - (X - Y)| > \varepsilon)$$

$$= P(|(X_n - X) - (Y_n - Y)| > \varepsilon)$$

$$\leq P(|X_n - X| + |Y_n - Y| > \varepsilon)$$

$$\leq \underbrace{P\left(|X_n - X| > \frac{\varepsilon}{2}\right)}_{\to 0} + \underbrace{P\left(|Y_n - Y| > \frac{\varepsilon}{2}\right)}_{\to 0}$$

as $n \to \infty$

18^{th} oct :

convergence in probability:

$$X_n \xrightarrow{P} X$$

we say X_n converges to X in probability if for $\forall \varepsilon > 0$

$$P(|X_n - X| > \varepsilon) \to 0 \text{ as } n \to \infty$$

Observe :

$$X_n \xrightarrow{P} X \Leftrightarrow X_n - X \xrightarrow{P} 0$$

Result: ① $X_n \xrightarrow{P} X$ & $Y_n \xrightarrow{P} Y$

$$\Rightarrow X_n \pm Y_n \xrightarrow{P} X \pm Y$$

② $X_n \xrightarrow{P} X \Rightarrow cX_n \xrightarrow{P} cX$ where $c \in \mathbb{R}$

proof : $P(|cX_n - cX| > \varepsilon)$

$$= P\left(|X_n - X| > \frac{\varepsilon}{|c|}\right) \longrightarrow 0$$

$$n \to \infty$$

③ $X_n \xrightarrow{P} c \Rightarrow X_n^2 \xrightarrow{P} c^2$

proof : as $X_n \xrightarrow{P} c$ $\quad X_n - c \xrightarrow{P} 0$

$$\Rightarrow (X_n - c)^2 \xrightarrow{P} 0$$

$$\Rightarrow X_n^2 - 2cX_n + c^2 \xrightarrow{P} 0$$

$$2cX_n \xrightarrow{P} 2c^2$$

$$\Rightarrow X_n^2 - c^2 \xrightarrow{P} 0$$

$$\Rightarrow X_n^2 \xrightarrow{P} c^2$$

Exe: X is a r.v , there for $\varepsilon > 0$, $\exists M > 0$ s.t $P(|X| > M) < \varepsilon$

$$\lim_{x \to +\infty} F(x) = 1 = \lim_{x \to \infty} P(X \leq x)$$

$$\lim_{x \to -\infty} F(x) = 0 = \lim_{x \to -\infty} P(X \leq x)$$

$\exists M_1 < 0$ s.t

$$F(x) < \varepsilon/2 \quad \forall x \leq M$$

$$P(X < M_1) \leq P(X \leq M_1) = F(M_1) < \varepsilon/2$$

$\exists M_2 > 0$ s.t

$$1 - F(M_2) < \varepsilon/2$$

$$\Rightarrow P(X > M_2) < \varepsilon/2$$

$$M = \max\{-M_1, M_2\}$$

$$P(M_1 \geq X > M_2) < \varepsilon/2 + \varepsilon/2$$

$$\Rightarrow P(|X| > M) < \varepsilon/2$$

④ $X_n \xrightarrow{P} X$ and Y is a r.v

then $X_nY \xrightarrow{P} XY$ in probability

proof: let $P(|X_nY - XY| > \varepsilon)$

$$= P(|Y||X_n - X| > \varepsilon)$$

now as $\exists M > 0$ s.t $\forall \delta > 0$ $P(|Y| > M) < \delta$

$$= P(|Y||X_n - X| > \varepsilon, |Y| > M)$$

$\left(P(A \cap (B \cup B^c)) = P(A \cap B) + P(A \cap B^c) \right)$

$$+ P(|Y||X_n - X| > \varepsilon, |Y| \leq M)$$

$$\leq P(|Y| > M) + P\left(|X_n - X| > \frac{\varepsilon}{M}\right)$$

as $\underbrace{\{\omega \mid |Y||X_n - X| > \varepsilon, |Y| \leq M\}}_{A} \subseteq \underbrace{\{\omega \mid |X_n - X| > \frac{\varepsilon}{M}\}}_{B}$

so $P(A) \leq P(B)$

now $P(|Y||X_n - X| > \varepsilon) \leq P(|Y| > M) + \underbrace{P\left(|X_n - X| > \frac{\varepsilon}{M}\right)}_{\to 0}$

as $n \to \infty$

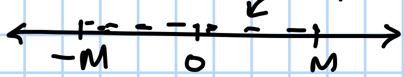
⑤ Suppose $X_n \xrightarrow{P} X$. Then for $\varepsilon > 0$, $\exists M > 0$ s.t

$P(|X_n| > M) < \varepsilon$ $\underline{\forall n \in \mathbb{N}}$ (Note: $\forall n$ is given)

proof: let $\varepsilon > 0$

since X is a r.v $\exists M > 0$ s.t

$P(|X| > M) < \varepsilon$



Most of mass of x is sitting here

as $P(|X_n - X| > \delta) \to 0$ as $n \to \infty$
most of mass of X_n for $n \to \infty$ also sits $-M$ to M.

now $P(|X_n|>M+1) = P(|X_n|>M+1, |X|\leq M) + P(|X_n|>M+1, |X|>M)$

$\leq P(|X_n|>M+1, |X|\leq M) + P(|X|>M)$

$\leq \underset{\to 0}{P(|X_n - X|>1)} + P(|X|>M) < \varepsilon + \varepsilon$

as $n\to\infty$ $\longrightarrow 0$

as $|X_n|>M+1 > |X|+1$

$\Rightarrow |X_n| - |X| > 1$

$\subseteq |X_n - X|>1$

$\Rightarrow P(|X_n|>M+1, |X|\leq M) \leq P(|X_n - X|>1)$

Hence, $P(|X_n|>M+1) < 2\varepsilon \quad \forall n\geq N$

$X_1, \dots X_{N-1}$ are random variables
for each $i = 1,2 \dots N-1$
$\exists M_i$ s.t

$P(|X_i|>M_i) < \varepsilon < 2\varepsilon$

$\tilde{M} = \max\{M_1, M_2, \dots M_{N-1}, M+1\}$

$\Rightarrow P(|X_n|>\tilde{M}) < 2\varepsilon \quad \forall n\in\mathbb{N}$

⑥ $X_n \xrightarrow{P} X \Rightarrow X_n^2 \xrightarrow{P} X^2$ (similar to c.r.v)

$\Rightarrow X_n - X \xrightarrow{P} 0$

$\Rightarrow (X_n - X)^2 \xrightarrow{P} 0$

$\Rightarrow X_n^2 + X^2 - 2X_nX \xrightarrow{P} 0$

now if $X_n \xrightarrow{P} X$ & Y is a R.V
then $X_nY \xrightarrow{P} XY$

we $X_nX \xrightarrow{P} X^2$

$\Rightarrow X_n^2 + X^2 - 2X^2 \xrightarrow{P} 0$

$\Rightarrow X_n^2 - X^2 \xrightarrow{P} 0$

$\Rightarrow X_n^2 \xrightarrow{P} X^2$

⑦ $X_n \xrightarrow{P} X$
$Y_n \xrightarrow{P} Y \Rightarrow X_nY_n \xrightarrow{P} XY$

proof : $P(|X_nY_n - XY|>\varepsilon)$ $\qquad X_n$ is a r.v

$= P(|X_nY_n - X_nY + X_nY - XY|>\varepsilon) \leq P(|X_nY_n - X_nY|>\varepsilon/2) \;\longrightarrow 0$

$+P(|X_nY - XY|>\varepsilon/2)$

$\longrightarrow 0 \quad n\to\infty$

Weak law of large number : (WLLN)

let $\{X_n\}$ be a seq of independent and identically distributed (same dist function) (i.i.d) r.v with mean μ then

$$\frac{S_n}{n} \xrightarrow{P} \mu \quad \text{where } S_n = \sum_{i=1}^{n} X_i$$

Note : identically dist $\not\Rightarrow$ point-wise same

$E(X_i) = \mu \ \forall i$

$S_n = \sum_{i=1}^{n} X_i \rightarrow$ Also a random variable

Example : $X_n \sim Ber(P)$

$\{X_n\}$ are iid bernoulli

$P(X_n = 1) = p$
$P(X_n = 0) = 1 - p$

$E(X_n) = p$

$\frac{\# H}{n} \rightarrow p$ for n large enough

weaker version of WLLN :

let $\{X_n\}$ be a sequence of ind random variable with

$E(X_n) = \mu$
$Var(X_n) = \sigma^2 \quad \forall n \in \mathbb{N}$ (we are arguing second moment exist)

then $\frac{S_n}{n} \xrightarrow{P} \mu$

proof : Here $S_n = \sum_{i=1}^{n} X_i$

then $\frac{S_n}{n} = \sum_{i=1}^{n} \frac{X_i}{n} = \frac{X_1}{n} + \cdots \frac{X_n}{n}$

$E(S_n) = n\mu$

$$P\left(\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right) = P\left(|S_n - n\mu| > n\varepsilon\right) \leq \frac{1}{n^2\varepsilon^2} Var(S_n)$$

$Var(S_n) = Var(\sum X_i)$

$= \sum Var(X_i) + 2\sum_{i<j} Cov(X_i, X_j)$ (as $Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2\,Cov(X_1, X_2)$)

$Var(S_n) = n\sigma^2 + 0$ as X_i, X_j are ind

or $P\left(|S_n - \mu n| > n\varepsilon\right) \leq \frac{n\sigma^2}{n^2\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2} \rightarrow 0$

Note : ind $\Rightarrow$ $cov(x_i, x_j) = 0$
$cov(x_i, x_j) = 0 \not\Rightarrow$ ind

$\therefore$ full independence is not required

$\{X_n\}$ s.t for $i \neq j$ ① $cov(x_i, x_j) = 0$

② $E(x_i) = \mu$

③ $var(x_i) = \sigma^2$

(This is weaker condition then even ind)

then

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

weaker version of weaker version of weaker law of large numbers

23rd oct:

WLLN:

$\{X_n\}_{n\geq 1}$ iid with mean μ. Then $\frac{S_n}{n} \xrightarrow{P} \mu$ when $S_n = \sum_{i=1}^{n} X_i$

weaker version of WLLN:

$\{X_n\}_{n\geq 1}$ are ind with mean μ, var σ^2

then

$$\frac{S_n}{n} \xrightarrow{P} \mu$$

Ex: suppose $X_n \xrightarrow{P} X$ and $f: \mathbb{R} \to \mathbb{R}$ continuous. Then show that $f(X_n) \xrightarrow{P} f(X)$

$X_n \xrightarrow{P} X$ means that

$$P(|X_n - X| > \varepsilon) \to 0 \quad \text{as } n \to \infty$$

To show:

$$P(|f(X_n) - f(X)| > \varepsilon') \to 0 \quad \text{as } n \to \infty$$

now as f is cont at x, for $\varepsilon > 0$, $\exists \delta > 0$ s.t

$$|f(x) - f(y)| < \varepsilon \text{ whenever } |x - y| < \delta$$

f is uniformly cont. For $\varepsilon > 0$, $\exists \delta > 0$ s.t

$$|f(x) - f(y)| < \varepsilon \text{ whenever } |x - y| < \delta$$

Suppose f is uniformly cont, let $\varepsilon > 0$. Then $\exists \delta > 0$ s.t

$$|f(x) - f(y)| < \varepsilon \text{ when } |x - y| < \delta$$

as $X_n \xrightarrow{P} X$

$$1 - P(|X_n - X| \leq \delta) = P(|X_n - X| > \delta)$$

as $n \to \infty$ $P(|X_n - X| \leq \delta) \to 1$

or $P(|X_n - X| < \delta) \to 1$

as $n \to \infty$

$$\{\omega \mid |X_n(\omega) - X(\omega)| < \delta\} \subseteq \{\omega \mid |f(X_n) - f(X)| < \varepsilon\}$$

$$1 \geq P(|f(X_n) - f(X)| < \varepsilon) \geq P(|X_n - X| < \delta) \to 1$$

or $P(|f(X_n) - f(X)| > \varepsilon) \to 0$

Exe: If f is cont but not uniformly cont then show above → done
(f is cont on $\mathbb{R}$ then uniformly cont on a closed bounded interval)
(Also use conv in probability)

Almost sure convergence: $X_n \xrightarrow{a.s} X$

$$P(\{\omega \mid X_n(\omega) \to X(\omega) \text{ as } n \to \infty\}) = 1$$

Equivalently, $P(\{\omega \mid X_n(\omega) \not\to X(\omega)\}) = 0$

Example: $\Omega = [0,1]$
$\mathcal{F} = \mathcal{B}([0,1])$
$P(A) = m(A)$
$\quad\uparrow$
lebesgue measure

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

$\{X_n\}_{n\geq 1}$ as $X_n(\omega) = \omega^n$, $\omega \in \Omega$

$X_n(\omega) \to 0$ if $\omega \in [0,1)$
$\quad \to 1$ if $\omega = 1$

or $X_n \to Y$ point wise
Y: discrete random variable

pmf of Y: $P(Y=0) = P([0,1)) = 1$
$P(Y=1) = P(\{1\}) = 0$

$X_n \to 0$ almost surely
as if $0 = X$ then
$X_n \to X$

$$P(\{\omega \mid X_n(\omega) \to X(\omega)\}) = P([0,1)) = 1$$

Note $X_n(1) \not\to X(1)$

Note: X_n is cont. random variable and X is discrete random variable

Theorm: $X_n \xrightarrow{a.s} X$ almostsurely iff for any $\varepsilon > 0$, $P(\overline{\lim}\{|X_n - X| > \varepsilon\}) = 0$

Denote $A_n = \{\omega : |X_n(\omega) - X(\omega)| > \varepsilon\}$

$$\overline{\lim} A_n = \bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k = \{\omega \mid \omega \in A_n \text{ for infinitely often}\}$$

$$P(\overline{\lim} A_n) = P\left(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_n\right) = \lim_{n\to\infty} P\left(\bigcup_{k=n}^{\infty} A_k\right)$$

by cont of probability

cor: $X_n \xrightarrow{a.s} X \Rightarrow X_n \xrightarrow{P} X$

proof: let $\varepsilon > 0$

$$P(|X_n - X| > \varepsilon) \leq P\left(\bigcup_{k=n}^{\infty} \{|X_n - X| > \varepsilon\}\right) = P\left(\bigcup_{k=n}^{\infty} A_k\right)$$

$\rightarrow 0$ as $n \rightarrow \infty$

$\left(\because \text{A.C mas } P(\overline{\lim} A_n) \rightarrow 0\right)$

proof: $(\Rightarrow)$ $X_n \xrightarrow{A.S} X$

$$P(\{\omega | X_n(\omega) \rightarrow X(\omega)\}) = 1$$

$$\approx P(\{\omega | X_n(\omega) \not\rightarrow X(\omega)\}) = 0$$

Denote $A^{\varepsilon} = \overline{\lim} A_n$

if $\omega \in A^{\varepsilon}$, then does $X_n(\omega)$ conv to $X(\omega)$?

$x_n \rightarrow x$. For any $\varepsilon > 0$, $\exists N$ s.t $x_n \in (x - \varepsilon, x + \varepsilon)$ $\forall n \geq N$

$x_n \not\rightarrow x$. $\exists \varepsilon > 0$ s.t $x_n \notin (x - \varepsilon, x + \varepsilon)$ for infinitely n.

if $A^{\varepsilon} = \overline{\lim} A_n = \{\omega | \omega \in A_n \text{ for i.o}\}$

if $\omega \in A^{\varepsilon} \Rightarrow \omega \in A_n$ for infinitely many n

$\Rightarrow |X_n(\omega) - X(\omega)| > \varepsilon$ for inf many n

$\Rightarrow X_n(\omega) \not\rightarrow X(\omega)$

or $\omega \in \{\omega | X_n(\omega) \not\rightarrow X(\omega)\}$

$\Rightarrow P(\{\omega | X_n(\omega) \not\rightarrow X(\omega)\}) = 0$

$\forall \omega \in A^{\varepsilon} \Rightarrow \omega \in \{\omega | X_n(\omega) \not\rightarrow X(\omega)\}$

$\Rightarrow A^{\varepsilon} \subseteq \{\omega | X_n(\omega) \rightarrow X(\omega)\}$

$P(A^{\varepsilon}) \leq P(\{\omega | X_n(\omega) \not\rightarrow X(\omega)\}) = 0$

$\Rightarrow P(A^{\varepsilon}) = 0$

$(\Leftarrow)$ for any $\varepsilon > 0$, $P(A^{\varepsilon}) = 0$

In particular $P(A^{1/m}) = 0$, $m \in \mathbb{N}$

$$A^{1/m} = \overline{\lim} \left\{\omega \,\middle|\, |X_n(\omega) - X(\omega)| > \frac{1}{m}\right\}$$

$$P\left(\bigcup_{m=1}^{\infty} A^{1/m}\right) \leq \sum_{m=1}^{\infty} P(A^{1/m}) = 0$$

denote $A = \bigcup_{m=1}^{\infty} A^{1/m}$, To show $P(\{\omega | X_n(\omega) \not\rightarrow X(\omega)\}) = 0$

Let $\omega \in \{\omega | X_n(\omega) \not\rightarrow X(\omega)\}$

as $X_n(\omega) \not\to X(\omega)$

$\exists\ \varepsilon > 0$ s.t

$|X_n(\omega) - X(\omega)| > \varepsilon$ for infinitely many n.

For $\varepsilon > 0$, $\exists M$ s.t $\frac{1}{M} < \varepsilon$

Therefore $|X_n(\omega) - X(\omega)| > \frac{1}{M}$ for infinitely many n.

Hence $\omega \in A^{1/M} \Rightarrow \omega \in A = \bigcup_{m=1}^{\infty} A^{1/m}$

or $\{\omega \mid X_n(\omega) \not\to X(\omega)\} \subseteq A$

$\Rightarrow P(\{\omega \mid X_n(\omega) \not\to X(\omega)\}) \leq P(A) = 0$

$\Rightarrow P(\{\omega \mid X_n(\omega) \not\to X(\omega)\}) = 0$

or $X_n \xrightarrow{a.s} X$

Ex: Suppose $X_n \xrightarrow{a.s} X$ and $Y_n \xrightarrow{a.s} Y$

then (i) $\alpha X_n + \beta Y_n \xrightarrow{a.s} \alpha X + \beta Y$

(ii) $X_n Y_n \xrightarrow{a.s} XY$

(iii) $f(X_n) \xrightarrow{a.s} f(X)$ where f is a cont function

(ii) $X_n Y_n \xrightarrow{a.s} XY$

i.e $P(\{\omega \mid X_n(\omega) Y_n(\omega) \to X(\omega) Y(\omega)\}) = 1$

as $P(\{\omega \mid X_n(\omega) \to X(\omega)\}) = 1$

& $P(\{\omega \mid Y_n(\omega) \to Y(\omega)\}) = 1$

now $P(\{\omega \mid X_n(\omega) Y_n(\omega) \to X(\omega) Y(\omega)\})$

$\geq P(\{\omega \mid X_n(\omega) \to X(\omega)\} \cap \{\omega \mid Y_n(\omega) \to Y(\omega)\})$

$= 1$

$\left(\begin{array}{l} X_n \to X \ \& \ Y_n \to Y \\ \Rightarrow X_n Y_n \to XY \\ \text{proof is already known} \end{array} \right)$

$\left(\because P(X_n Y_n \to X_n Y_n) \geq P(X_n \to X \cap Y_n \to Y) \right)$

Example: $X_n \xrightarrow{P} X$ let $X_n \not\xrightarrow{a.s} X$

$\left(\begin{array}{c} P \not\Rightarrow A.S \\ \text{proof} \end{array} \right)$

$\Omega = [0,1]$

$\mathcal{F} = \mathcal{B}([0,1])$

$P = m(A)$

define $\{X_n\}_{n \geq 1}$ in the following:

$X_1(\omega) = 1 \quad \forall\ \omega \in [0,1]$

$X_2(\omega) = \begin{cases} 1 & ; \omega \in [0, \frac{1}{2}] \\ 0 & ; \omega \in (1/2, 1] \end{cases}$

$X_3(\omega) = \begin{cases} 0 & ; \omega \in [0, 1/2) \\ 1 & ; \omega \in [1/2, 1] \end{cases}$

$$X_4(\omega) = \begin{cases} 1 & ; \omega \in [0, \frac{1}{4}] \\ 0 & ; \omega \in (\frac{1}{4}, 1] \end{cases}$$

$$X_5(\omega) = \begin{cases} 1 & ; \omega \in [1/4, 1/2] \\ 0 & ; \omega \in \text{otherwise} \end{cases}$$

(note: $AS \Rightarrow P$
$P \not\Rightarrow AS$)

$$X_6(\omega) = \begin{cases} 1 & ; \omega \in [1/2, 3/4] \\ 0 & ; \text{o.w} \end{cases}$$

$$X_7(\omega) = \begin{cases} 1 & ; \omega \in [3/4, 1] \\ 0 & ; \text{o.w} \end{cases}$$

we cont in this way

$$P(|X_n - 0| > \varepsilon) \leq \frac{1}{2k} \longrightarrow 0 \text{ as } \underset{k \to \infty}{n \to \infty}$$

if we fix ω then
$X_n(\omega)$ keeps on fluctuating b/w 0 & 1

$$P(\{\omega \mid X_n(\omega) \not\longrightarrow \underset{\substack{\parallel \\ 0}}{X(\omega)}\}) = 1$$

$$\text{or } X_n \xrightarrow{P} X \not\Rightarrow X_n \xrightarrow{A.S} X$$

Que: If f is cont but not uniformly cont then show above
(f is cont on $\mathbb{R}$ then uniformly cont on a closed bounded interval)
(Also we conv in probability)

f is cont, then $\forall \varepsilon > 0$, $\forall P \in X$, $\exists \delta > 0$ s.t
$|x - P| < \delta \Rightarrow |f(x) - f(P)| < \varepsilon$

$$\text{or } |x - P| \leq \delta - \frac{1}{n} \Rightarrow |f(x) - f(P)| < \varepsilon \quad \forall n \in \mathbb{N}$$

as $X_n \xrightarrow{P} X$ as $n \to \infty$

Note: for $|x - P| \leq \delta - 1/n$ f becomes ab. cont

$$P(|X_n - X| > \delta) \longrightarrow 0$$

$$\text{or } P(|X_n - X| < \delta) \longrightarrow 1$$

now $\{\omega \mid |X_n - X| \leq \delta - \frac{1}{n}\} \subseteq \{\omega \mid |f(X_n) - f(X)| < \varepsilon\}$

$$\Rightarrow \bigcup_{n=1}^{\infty} \{\omega \mid |X_n - X| \leq \delta - \frac{1}{n}\} \subseteq \{\omega \mid |f(X_n) - f(X)| < \varepsilon\}$$

$$\Rightarrow \{\omega \mid |X_n - X| < \delta\} \subseteq \{\omega \mid |f(X_n) - f(X)| < \varepsilon\}$$

$$\Rightarrow P(\{\omega \mid |X_n - X| < \delta\}) \leq P(\{\omega \mid |f(X_n) - f(X)| < \varepsilon\})$$

$$\Rightarrow P(\{\omega \mid |f(X_n) - f(X)| > \varepsilon\}) \longrightarrow 0$$

as $n \to \infty$

25th oct :

Recall : Theorem $X_n \xrightarrow{a.s} X$ iff for any $\varepsilon > 0$ $P(\overline{\lim}\{|X_n - X| > \varepsilon\}) = 0$

Cor : $X_n \xrightarrow{a.s} X \Rightarrow X_n \xrightarrow{P} X$

But converse is not true

Example : $\Omega = [0,1]$ $\mathcal{F} = \mathcal{B}([0,1])$ $P = m(A)$

$\{X_n\}_{n\geq 1}$ suppose $2^k \leq n < 2^{k+1}$ for some k

$$n = 2^k + l$$
$$0 \leq l < 2^k$$

define $X_n(\omega) = \begin{cases} 1 & ; \omega \in [l/2^k, l+1/2^k] \\ 0 & ; \text{otherwise} \end{cases}$

now $P(|X_n - 0| > \varepsilon) \leq \frac{1}{2^k} \longrightarrow 0$ (for large enough n it depends on ε) as $n \to \infty$

$X_n \xrightarrow{P} 0$

But $X_n(\omega) \not\to 0$ $\forall \omega \in \Omega$

Theorem : $X_n \xrightarrow{P} X$ iff given any subsequence $\{X_{n_k}\}$ of $\{X_n\}$ there exist a further subseq $\{X_{n_{k_l}}\}$ s.t $X_{n_{k_l}} \xrightarrow{a.s} X$

(optional exercise)

Exe : $X_n \xrightarrow{P} X$ and f is a cont. f^n then $f(X_n) \xrightarrow{P} f(X)$

$Y_n = f(X_n)$

$Y = f(X_n)$

To show : $Y_n \xrightarrow{P} Y$, let $\{Y_{n_k}\}$ be a subsequence of $\{Y_n\}$

$Y_{n_k} = f(X_{n_k})$

$\{X_{n_k}\}$ is a subsequence of $\{X_k\}$

as $X_n \xrightarrow{P} X$, $\{X_{n_k}\}$ has a subseq which $\{X_{n_{k_l}}\}$ s.t $X_{n_{k_l}} \xrightarrow{a.s} X$

As f is a cont function $f(X_{n_{k_l}}) \xrightarrow{a.s} f(X)$

Hence, $Y_{n_{k_l}} \xrightarrow{a.s} Y$. By the last theorem

$Y_n \xrightarrow{P} Y$

First-Borel-Cantelli Lemma :

Let $\{A_n\}$ be a seq^n of events. Suppose $\sum_{n=1}^{\infty} P(A_n) < \infty$, then

$P(\overline{\lim} A_n) = 0$

proof : $P(\overline{\lim} A_n) = P(\bigcap_{1}^{\infty} \bigcup_{n}^{\infty} A_n)$

$$= \lim_{n\to\infty} P(\bigcup_{k=n}^{\infty} A_n)$$

$$\leq \lim_{n\to\infty} \left(\sum_{k=n}^{\infty} P(A_k)\right) = 0$$

as $\sum_{n=1}^{\infty} P(A_n) < \infty$

Exe: $\{X_n\}_{n\geq 1}$, $P(X_n=n) = \frac{1}{n^p}$

$P(X_n=0) = 1-\frac{1}{n^p}$ where $p>0$ (fixed)

Let $\varepsilon>0$ $\quad P(|X_n-0|>\varepsilon) \leq P(X_n=n) = \frac{1}{n^p}$

$\Rightarrow P(|X_n|>\varepsilon) \leq \frac{1}{n^p} \to 0$ as $n\to\infty$

$X_n \xrightarrow{P} 0$ $\quad$ To show:

for A.S: $\quad P(\{\omega \mid X_n \to X\}) = 1$

using first Borel-Cantelli lemma

$$\sum_{n=1}^{\infty} P(|X_n|>\varepsilon) \leq \sum_{n=1}^{\infty} \frac{1}{(n)^p} < \infty$$

finite for $p>1$

$\Rightarrow P(\overline{\lim}\, A_n) = 0$

$\Rightarrow X_n \xrightarrow{a.s} 0$ if $p>1$

($\because P(\overline{\lim}\{|X_n-X|>\varepsilon\}) = 0 \Rightarrow X_n \xrightarrow{a.s} X$)

weaker version of SLLN:

Let $\{X_n\}_{n\geq 1}$ be a sequence of independent random variables with mean μ and $E(X_n^4) \leq M$ $\forall n\geq 1$. Then

$$\frac{S_n}{n} \xrightarrow{a.s} \mu$$

proof: wlog $\mu=0$ as if $\mu\neq 0$ then

$Y_n = X_n - \mu$

$E(Y_n) = 0$

$\frac{1}{n}\sum Y_n \xrightarrow{a.s} 0 \Rightarrow \frac{1}{n}\sum X_n \xrightarrow{a.s} \mu$

$$P\left(\left|\frac{S_n}{n}-0\right|>\varepsilon\right) \leq \frac{E(S_n^2)}{n^2\varepsilon^2} = \frac{\sum_{i=1}^{n} E(X_i^2)}{n^2\varepsilon^2}$$

$\quad ||$

$|S_n| > n\varepsilon$ $\quad$ (cov term 0)

Cauchy Schwarz:

$|E(XY)| \leq \sqrt{E(X^2)E(Y^2)}$

$E(X_i^2 \cdot 1) \leq \sqrt{E(X_i^4)E(1^2)} = \sqrt{E(X_i^4)}$

or $\quad P\left(\left|\frac{S_n}{n}-0\right|>\varepsilon\right) \leq \sum \frac{\sqrt{E(X_i^4)}}{n^2\varepsilon^2} \leq \frac{n\times\sqrt{M}}{n^2\varepsilon^2} = \frac{\sqrt{M}}{n\varepsilon^2}$

$$\Rightarrow \sum_{1}^{\infty} P\left(\left|\frac{S_n}{n}\right|>\varepsilon\right) \leq \sum_{1}^{\infty} \frac{\sqrt{M}}{\varepsilon^2 n} \to \infty$$

we cannot use first Borel lemma

new approach

$$P(|X|>\varepsilon) = P(X^2 \geq \varepsilon^2) = P(X^4 \geq \varepsilon^4) \leq \frac{E(X^4)}{\varepsilon^4}$$

$$\Rightarrow P(|X|>\varepsilon) \leq \frac{E(X^4)}{\varepsilon^4}$$

$$P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \leq \frac{1}{\varepsilon^4 n^4} E(S_n^4)$$

$$S_n = \sum_{i=1}^{n} x_i$$

$$E(S_n^4) = E\Big(\underbrace{\sum x_i^4}_{n \text{ such terms}} + \sum x_i^3 x_j^1 + \sum x_i^2 x_j^2\Big)$$

$$= \sum_{i=1}^{n} E(x_i^4) + \sum_{i \neq j} E(x_i^2 x_j^2) + \sum_{i \neq j} \underbrace{E(x_i x_j^3)}_{0}$$

$$\leq nM + \sum_{i \neq j} E(x_i^2 x_j^2) + 0$$

$\Big($as ind: $E(x_i x_j^3) = E(x_i)E(x_j^3) = 0\Big)$

now, $E(x_i^2 x_j^2) \leq \sqrt{E(x_i^4)E(x_j^4)}$
$\leq M$

so $\sum_{i \neq j} E(x_i^2 x_j^2) \leq \sum_{i \neq j} cM \leq cn^2 M$ as $i \neq j$

$$\Sigma(x_1 + x_2 + \cdots x_n)^4 = \Sigma\left(\frac{4!}{r_1!\, r_2! \cdots r_n!}\right) x_1^{r_1} x_2^{r_2} \cdots x_n^{r_n}$$

for $x_1^2 x_2^2$ $\leq cn^2$
no. of terms

so $E(S_n^4) \leq nM + cn^2 M$

$$\Rightarrow P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) \leq \frac{nM + cn^2M}{\varepsilon^4 n^4} = \frac{M}{\varepsilon^4 n^3} + \frac{cM}{n^2 \varepsilon^4}$$

$\Rightarrow \sum P\left(\left|\frac{S_n}{n}\right| > \varepsilon\right) < \infty$ as $p > 1$ $\quad p > 1$

$\Rightarrow$ using first borel lemma:

$$P\left(\overline{\lim}\left\{w \,\middle|\, \left|\frac{S_n}{n}\right| > \varepsilon\right\}\right) = 0$$

$$\Rightarrow \frac{S_n}{n} \xrightarrow{a.c} 0$$

Note: If $\mu \neq 0$ then define $Y_n = X_n - \mu$
$E(Y_n) = 0$, $\{Y_n\}$ are independent as $\{X_n\}$ are independent.

26^{th} oct:

Exe: $X_n \xrightarrow{P} X$, show that $|X_n| \xrightarrow{P} |X|$

as $X_n \xrightarrow{P} X$, $P(|X_n - X| > \varepsilon) \to 0$ as $n \to \infty$

to show: $P(||X_n| - |X|| > \varepsilon) \to 0$ as $n \to \infty$

for $\varepsilon < ||X_n| - |X|| \leq |X_n - X|$
(triangle ineq)

$\Rightarrow P(||X_n| - |X|| > \varepsilon) \leq P(|X_n - X| > \varepsilon) \to 0$ as $n \to \infty$

$(\because \{\omega \mid ||X_n| - |X|| > \varepsilon\} \subseteq \{\omega \mid |X_n - X| > \varepsilon\})$

Exe: $\{X_n\}$ are iid $U(0,\theta)$ ← uniform dist in $(0,\theta)$, show that

$\max\{X_1, X_2, \dots X_n\} \xrightarrow{P} \theta$

$X \sim U(0,\theta)$

$$f(x) = \begin{cases} \frac{1}{\theta} & ; x \in (0,\theta) \\ 0 & ; \text{otherwise} \end{cases}$$

let $Y = \max\{X_1, \dots X_n\}$

to show: $Y_n \xrightarrow{P} \theta$

let $\varepsilon > 0$

$P(|Y_n - \theta| > \varepsilon) = P(Y_n > \theta + \varepsilon) + P(Y_n < \theta - \varepsilon)$

Note: all X_i's are not same

X, Y say i.d does not mean

$X(\omega) = Y(\omega)\ \forall \omega$

does not mean $F_X(\omega) = F_Y(\omega)\ \forall \omega$

let X be uniform dist on

$S = \{-n, -n+1, \dots 0, 1, 2, \dots n\}$

$P(X = i) = \frac{1}{2n+1}$ for $i \in S$

$Y(\omega) = -X(\omega)$

then

$Y \neq X$ but $P(Y = i) = \frac{1}{2n+1}$

now $P(Y_n > \theta + \varepsilon) = 0$

$\Rightarrow P(Y_n < \theta - \varepsilon)$

$= P(X_i < \theta - \varepsilon, \forall i = 1, 2, \dots n)$

$= P\left(\bigcap_{i=1}^{n} \{X_i < \theta - \varepsilon\}\right)$

$= \prod_{i=1}^{n} P(X_i < \theta - \varepsilon)$ $(\because X_i$'s are independent$)$

$= \left(\frac{\theta - \varepsilon}{\theta}\right)^n \to 0$ as $n \to \infty$ $\quad \frac{\theta - \varepsilon}{\theta} < 1$

que: $\{X_n\}$ are iid with common density function $f(x) = \begin{cases} e^{-x+\theta} & ; x \geq \theta \\ 0 & ; \text{otherwise} \end{cases}$

show that

(i) $\frac{1}{n}\sum_{i=1}^{n} X_i \xrightarrow{P} (1+\theta)$

(ii) $\min\{X_1, \ldots X_n\} \xrightarrow{P} \theta$

(i) As $\{X_n\}$ are i.i.d with $f(x) = \begin{cases} e^{-x+\theta} & ; x \geq \theta \\ 0 & ; \text{otherwise} \end{cases}$

$$F(x) = \int_{-\infty}^{x} f(x)\,dx$$

$$P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt$$

$$E[X] = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{\theta}^{\infty} e^{-x+\theta} \cdot x\,dx$$

$$= e^{\theta}\left[x \frac{e^{-x}}{-1} + \int [1]\left[\frac{e^{-x}}{+1}\right] dx\right]$$

$$= e^{\theta}\left[-xe^{-x} + \int e^{-x}\,dx\right]$$

$$= e^{\theta}\left[-xe^{-x} - e^{-x}\right]_{\theta}^{\infty}$$

$$= \cancel{e^{\theta}}\left[0 - (-\theta \cancel{e^{-\theta}} - \cancel{e^{-\theta}})\right]$$

$$E[X] = \theta + 1$$

or by WLLN

$$\frac{S_n}{n} \xrightarrow{P} \theta + 1 \quad \text{as } n \to \infty$$

(ii) $\min\{X_1, X_2, \ldots X_n\} \xrightarrow{P} \theta$

let $Y_n = \min\{X_1, \ldots X_n\}$

then

$$P(|Y_n - \theta| > \varepsilon) = P(Y_n > \theta + \varepsilon) + \cancel{P(Y_n < \theta - \varepsilon)}^{\,0}$$

$$= P(Y_n > \theta + \varepsilon)$$

$$= P(X_i > \theta + \varepsilon, \forall i = 1, 2 \ldots n)$$

$$= P\left(\bigcap_{i=1}^{n} \{X_i > \theta + \varepsilon\}\right)$$

$$= \prod_{i=1}^{n} P(\{X_i > \theta + \varepsilon\})$$

$$= \prod_{i=1}^{n} \left[\int_{\theta+\varepsilon}^{\infty} e^{-x+\theta} dx \right]$$

$$= \prod_{i=1}^{n} \left[\cancel{e^{\theta}} \right] \left[\cancel{e^{-\theta}} {}^{-\varepsilon} \right]$$

$$= \prod_{i=1}^{n} \left[e^{-\varepsilon} \right] = e^{-n\varepsilon} \to 0 \text{ as } n \to \infty$$

or $P(|Y_n - \theta| > \varepsilon) \to 0$

Sme: $\{X_n\}$ suppose $\max_{1 \le k \le n} |X_k| \xrightarrow{P} 0$

to show: $\frac{S_n}{n} \xrightarrow{P} 0$

now $\left| \frac{S_n}{n} \right| = \frac{1}{n} |\Sigma X_i| \le \frac{1}{n} \Sigma |X_i| \le \frac{1}{n} \Sigma \max |X_i|$

$= \max |X_i| = Y_n$

let $\varepsilon > 0$

$P\left(\left| \frac{S_n}{n} \right| > \varepsilon \right) \le P(Y_n \ge \varepsilon) \to 0$ as $n \to \infty$

Sme: let $\{X_n\}_{n \ge 1}$ be iid r.v's s.t $E(X_1^2) < \infty$

find probability limit of

$$\frac{1}{n} \Sigma X_j^2 - \left(\frac{1}{n} \Sigma X_j \right)^2$$

$\frac{1}{n} \Sigma X_j^2$

$Y_n = X_n^2$

$\{Y_n\}_{n \ge 1}$

$$F_{Y_1}(y) = P(X_1^2 \le y)$$

$$= P(-\sqrt{y} \le X_1 \le \sqrt{y})$$

$$= \int_{-\sqrt{y}}^{\sqrt{y}} f(x) dx$$

$\{Y_n\}_{n \ge 1}$ are iid with finite mean

as $E(Y_1) < \infty$ $\left(E(Y_1) = E(X_1^2) < \infty \right)$

$\Rightarrow E(Y_i) = E(Y_1) < \infty$

By WLLN

$\frac{1}{n} \Sigma Y_n \xrightarrow{P} E(Y_1) = E(X_1^2)$

by WLLN $\frac{1}{n} \Sigma X_i \xrightarrow{P} E(X_1)$

now $\frac{1}{n} \Sigma X_j^2 - \left(\frac{1}{n} \Sigma X_j \right)^2 \xrightarrow{P} E(X_1^2) - (E(X_1))^2 = \text{var}(X_1)$

Exe: $\{X_n\}_{n\geq 1}$ where $P(X_n=1)=P_n$
$P(X_n=0)=1-P_n$

(i) show that $X_n \xrightarrow{P} 0$ iff $P_n \to 0$ as $n\to\infty$

(ii) suppose $\sum P_n < \infty$, Does X_n conv a.s ?

(i) Let $\varepsilon>0 \Rightarrow P(|X_n-0|>\varepsilon) \leq P_n \to 0$ as $n\to\infty$
$\Leftrightarrow P_n \to 0$ as $n\to\infty$

(ii) $X_n \xrightarrow{a.s} 0$ iff $\forall \varepsilon>0$
$P(\overline{\lim}\{\omega \mid |X_n-0|>\varepsilon\})=0$

now as $\sum P(|X_n|>\varepsilon) \leq \sum P_n < \infty$
$\Rightarrow P(\overline{\lim}\{|X_n|>\varepsilon\})=0$
$\Rightarrow X_n \xrightarrow{a.s} 0$

Exe: Suppose $X_n \xrightarrow{P} X$, show that $\{X_n\}$ is "cauchy in probability"

here cauchy in probability means that $\forall \varepsilon>0$
we have $P(|X_n - X_m|>\varepsilon) \to 0$ as $n,m\to\infty$

Here $P(|X_m - X_n|>\varepsilon) = P(|X_n - X + X - X_m|>\varepsilon)$
$\leq P(|X_n - X| + |X_m - X|>\varepsilon)$
$\leq \underbrace{P(|X_n - X|>\varepsilon/2)}_{\to 0} + \underbrace{P(|X_m - X|>\varepsilon/2)}_{\to 0}$
as $n,m\to$

Exe: Suppose $X_n \xrightarrow{P} X$ and $X_n(\omega) \leq X_{n+1}(\omega)$ $\forall n,\omega \in \mathbb{N},\Omega$
show that $X_n \xrightarrow{a.s} X$

as $X_n \xrightarrow{P} X$ iff $\forall \{X_{n_k}\}$ s.t $X_{n_{k_l}} \xrightarrow{a.s} X$

here $\{X_n\}$ is also a subseq
so,
$\exists \{X_{n_k}\}$ s.t
$P(\{\omega \mid \lim_{k\to\infty} X_{n_k}(\omega) = X(\omega)\})=1$

now $\omega \in \{\omega \mid \lim_{k\to\infty} X_{n_k}(\omega) = X(\omega)\}$

as X_n is a mono inc seq with a subseq s.t
$X_{n_k}(\omega) \to X(\omega)$
$\Rightarrow X_n(\omega) \to X(\omega)$
$\Rightarrow P(\{\omega \mid \lim_{n\to\infty} X_n(\omega) = X(\omega)\})=1$

$\Rightarrow X_n \xrightarrow{a.s} X$

Second Borel cantelli lemma:

suppose $\{A_n\}_{n\geq 1}$ a seq of ind events and $\sum P(A_n) = \infty$ then

$$P(\overline{\lim} A_n) = 1$$

Exe: suppose $\{x_n\}$ are ind & $P(x_n = n^2) = \frac{1}{n}$

$$P(x_n = 0) = 1 - \frac{1}{n}$$

does $x_n \xrightarrow{a.s} 0$

let $\varepsilon = 1/2$ then

$$\sum P(|x_n - 0| > \varepsilon) = \sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

$$A_n = \{|x_n| > \varepsilon\}$$

$$\Rightarrow P(\overline{\lim} A_n) = 1$$

By second B.C lemma

$$x_n \stackrel{a.s}{\not\to} x$$

Note: $x_n \xrightarrow{P} 0$ as $P(|x_n - 0| > \varepsilon) \leq \frac{1}{n} \to 0$ as $n \to \infty$

as $x_n \xrightarrow{a.s} 0 \Rightarrow x_n \xrightarrow{P} 0$

here x_n to conv a.s it has to conv to 0.

29th oct :

Almost sure conv $\Rightarrow$ conv in probability
(SLLN) (WLLN)
$\not\Leftarrow$

Convergence in distribution :

X_n conv to X in distribution as $n\to\infty$, if $F_n(x) \longrightarrow F(x)$ at all continuity point x of F, where F_n and F are distribution f^n of X_n and X respectively.

$X_n \xrightarrow{\mathcal{D}} X$ or $X_n \xrightarrow{d} X$

Exe : ① $\{X_n\}$ is a seq of random variable with dist fn $\{F_n\}$

$$F_n(x) = \begin{cases} 0 ; & \text{if } x<0 \\ 1-\frac{1}{n} ; & \text{if } 0\le x<n \\ 1 ; & x \geq n \end{cases}$$

$$F_n(x) \longrightarrow F(x) = \begin{cases} 0 ; & x<0 \\ 1 ; & x\geq 0 \end{cases}$$

(F is a valid dist function)

F corresponds to $X\equiv 0$ where $P(X=0)=1$

so $X_n \xrightarrow{\mathcal{D}} 0$

$X\equiv 0$ means $\forall x \in \mathbb{R}$

$\{\omega | \underset{0}{\underset{\shortparallel}{X(\omega)}} \le x\} \in \mathcal{F}$

where $\{\omega | \underset{0}{\underset{\shortparallel}{X(\omega)}} \le x\} = \phi$ for $x<0$

& $\{\omega | \underset{0}{\underset{\shortparallel}{X(\omega)}} \le x\} = \Omega$ for $x\geq 0$

or $F(x) = \begin{cases} 0 ; & x<0 \\ 1 ; & x\geq 0 \end{cases}$

② $\{X_n\}$ be set of r.v

$$F_n(x) = \begin{cases} 0 ; & x<n \\ 1 ; & x\geq n \end{cases}$$

$F_n(x) \longrightarrow 0 \quad \forall x \in \mathbb{R} \qquad 0 = F(x)$
as $n\to\infty$

here limit of F_n is F ; $F(x) = 0 \ \forall x\in\mathbb{R}$

as $\lim_{n\to\infty} 0 = 0$

F is not a dist function

X_n does not conv in dist

③ $$F_n(x) = \begin{cases} 0 & ; x<0 \\ (x/\theta)^n & ; 0\le x<\theta \\ 1 & ; x\geq\theta \end{cases}$$

$$F_n(x) \longrightarrow F(x) = \begin{cases} 0 ; & x<0 \\ 0 ; & 0\le x<\theta \\ 1 ; & x\geq\theta \end{cases}$$

$$= \begin{cases} 0 ; & x<\theta \\ 1 ; & x\geq\theta \end{cases}$$

here F is a dist function of cost r.v $X\equiv\theta$

$X_n \xrightarrow{d} \theta$

④ $\{X_n\}$ where X_n takes values

$$P\left(X_n = (-1)^n \frac{1}{n}\right) = 1$$

$$F_{2n}(x) = \begin{cases} 0; & x < 1/2n \\ 1; & x \geq 1/2n \end{cases} \qquad F_{2n+1}(x) = \begin{cases} 0; & x < -1/2n+1 \\ 1; & x \geq -1/2n+1 \end{cases}$$

$$F_{2n}(x) \longrightarrow \begin{cases} 0; & x \leq 0 \\ 1; & x > 0 \end{cases} = F(x) \qquad F_{2n+1}(x) \longrightarrow \begin{cases} 0; & x < 0 \\ 1; & x \geq 0 \end{cases} = G(x)$$

(not a dist funtion) (dist. function of const $X \equiv 0$)

Here $F_n(x) \to G(x)$ at all cont points of G

$F_n(0) \not\to G(0)$ is okay as 0 is not a continuity point of G.

$$X_n \xrightarrow{d} 0$$

Note: ④ uses definition of distribution convergence

Result: conv in probability $\Rightarrow$ conv in distribution

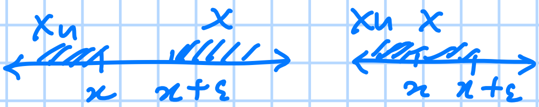
proof: Suppose $X_n \xrightarrow{P} X$. Let F_n, F be the distribution f^n of X_n and X resp.

We have to show $F_n(x) \longrightarrow F(x)$ at all continuity point of x of F.

Let $\varepsilon > 0$

To show: $F(x-\varepsilon) \leq \lim\limits_{n\to\infty} F_n(x) \leq F(x+\varepsilon)$

$$F_n(x) = P(X_n \leq x) = P(X_n \leq x, X > x+\varepsilon) + P(X_n \leq x, X \leq x+\varepsilon)$$



$$\leq P(|X - X_n| > \varepsilon) + P(X \leq x+\varepsilon)$$

$$F(x-\varepsilon) = P(X \leq x-\varepsilon) = P(X \leq x-\varepsilon, X_n > x) + P(X \leq x-\varepsilon, X_n \leq x)$$

$$\leq P(|X_n - X| > \varepsilon) + P(X_n \leq x)$$

$$F(x-\varepsilon) \leq P(|X_n - X| > \varepsilon) + F_n(x)$$

$$\Rightarrow F(x-\varepsilon) - P(|X_n - X| > \varepsilon) \leq F_n(x) \leq F(x+\varepsilon) + P(|X_n - X| > \varepsilon)$$

$$\Rightarrow F(x-\varepsilon) \leq \lim_{n\to\infty} F_n(x) \leq F(x+\varepsilon)$$

If x is a point of cont of F, as $\varepsilon \to 0$ we get $\lim\limits_{n\to\infty} F_n(x) = F(x)$

Note: convergence in probability $\underset{\not\Leftarrow}{\Longrightarrow}$ conv in dist
(WLLN) (CLT)

Result: If $X_n \xrightarrow{D} c$, then $X_n \xrightarrow{P} c$

(c is the point of discontinuity)

proof: let F be a dist function of c then

$$F(x) = \begin{cases} 0; & x < c \leftarrow \text{point of discontinuity} \\ 1; & x \geq c \end{cases}$$

let $\varepsilon > 0$, $P(|X_n - c| > \varepsilon) = P(X_n > c+\varepsilon) + P(X_n < c-\varepsilon)$

$\leq 1 - P(X_n \leq c+\varepsilon) + P(X_n \leq c-\varepsilon)$

$= 1 - F_n(c+\varepsilon) + F_n(c-\varepsilon)$

$\to 1 - F(c+\varepsilon) + F(c-\varepsilon)$

since $X_n \xrightarrow{d} c$ and $c+\varepsilon$, $c-\varepsilon$ are point of cont of F

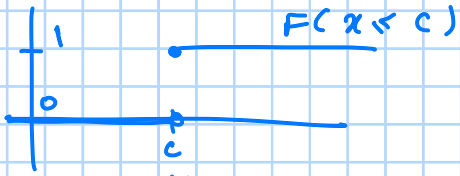
$\therefore$ $P(|X_n - c| > \varepsilon) \to 1 - (1) + (0)$
$= 0$

or $P(|X_n - c| > \varepsilon) \to 0$
$\forall \varepsilon$

$\therefore$ $X_n \xrightarrow{P} c$

Note: $X_n \xrightarrow{d} X$ and X is const (say c)
then and then only
$X_n \xrightarrow{P} X$ for X (const)

so $X_n \xrightarrow{d} c \Rightarrow X_n \xrightarrow{P} c$
for c to be a const random varible



Point of discontinuity

1^{st} Nov :

Convergence in distribution:

$$X_n \xrightarrow{d} X$$

X_n converges in distribution to x if $F_n(x)$ converges to $F(x)$ at all cont. points x of F, where F_n and F are the distribution f^n of X_n and X, respectively.

Result : $X_n \xrightarrow{P} X \Rightarrow X_n \xrightarrow{d} X$
$X_n \xrightarrow{d} c \Rightarrow X_n \xrightarrow{P} c$

Exe : X is a r.v

$$P(X=1) = P(X=-1) = \frac{1}{2}$$

define $X_n = X \quad \forall n \geq 1$

does X_n conv in dist?

$$F(x) = P(X \leq x) = \begin{cases} 0 ; & x < -1 \\ 1/2 ; & -1 \leq x < 1 \\ 1 ; & 1 \leq x \end{cases}$$

$$F_n(x) = P(X_n \leq x) = F(x) = \begin{cases} 0 ; & x < -1 \\ 1/2 ; & -1 \leq x < 1 \\ 1 ; & 1 \leq x \end{cases}$$

$\therefore F_n(x) \longrightarrow F(x)$ to all $x \in \mathbb{R}$

Therefore $X_n \xrightarrow{d} X$

Exe : If $Y = -X$, where X is same as above, then does $X_n \xrightarrow{d} Y$?

$$P(Y=1) = \frac{1}{2} = P(Y = -\frac{1}{2})$$

$$F_y(y) = \begin{cases} 0 ; & x < -1 \\ 1/2 ; & -1 \leq x < 1 \\ 1 ; & x \geq 1 \end{cases}$$

$$= F(x)$$
$$= F_n(x)$$

so $F_n(x) \longrightarrow F_y(x)$, $\forall x \in \mathbb{R}$ (at all continuity points)

$\therefore$

$$X_n \xrightarrow{d} Y$$

Exe : does $X_n \xrightarrow{P} Y$ (same R.V defined above)?

for $X_n \xrightarrow{P} Y$

we would have

$$P(|X_n - Y| > \varepsilon)$$
$$= P(2|X_n| > \varepsilon)$$
$$= P(|X_n| > \varepsilon/2)$$

for $\varepsilon < 2$ say 1

then $P(|X_n| > 1/2) = \frac{1}{2} + \frac{1}{2} = 1 \neq 0$

$\therefore \quad X_n \overset{P}{\not\to} Y$

Exe: $X_n \xrightarrow{d} X, \; Y_n \xrightarrow{d} Y$

$\Rightarrow X_n + Y_n \xrightarrow{d} X + Y$?

for $X_n = X$

s.t $P(X=1) = P(X=-1) = \frac{1}{2}$

$Y_n = X$

$Y = -X$ then ① $X_n \xrightarrow{d} X$

② $Y_n \xrightarrow{d} Y$

(already seen)

$X_n + Y_n = 2X$

$X + Y = 0$

$X_n + Y_n \overset{d}{\not\to} X + Y$

$$H_n(x) = \begin{cases} 0 & ; \; x < -2 \\ 1/2 & ; \; -2 \leq x < 2 \\ 1 & ; \; x \geq 2 \end{cases} \qquad H(x) = \begin{cases} 0 & ; \; x < 0 \\ 1 & ; \; x \geq 1 \end{cases}$$

Note: $X_n \xrightarrow{d} X$

$Y_n \xrightarrow{d} Y \;\not\Rightarrow\; X_n + Y_n \xrightarrow{d} X + Y$

Slutsky's theorm:

Suppose $X_n \xrightarrow{d} X$ and $Y_n \to c$ then

① $X_n + Y_n \xrightarrow{d} X + c$

② $X_n Y_n \xrightarrow{d} cX$

Exe: $\{X_n\}$ seqn of r.v with dist f^n as:

$$F_n(x) = \begin{cases} 0 & ; \; x < -n \\ \frac{x+n}{2n} & ; \; -n \leq x < n \\ 1 & ; \; x \geq n \end{cases}$$

does X_n conv in dist?

$F_n(x)$ for $n \to \infty$ it converges to $\frac{1}{2}$

or for $n \to \infty$

$F_n(x) \to \frac{1}{2} \quad \forall x \in \mathbb{R}$

$F(x) = \frac{1}{2} \; \forall x \in \mathbb{R}$

← Not a dist function

$\therefore$ X_n does not conv in distribution

Exe: $\{X_n\}$ iid where $X_i \sim N(0,1)$ $\forall i$

$$\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$$

from WLLN $\bar{X}_n \xrightarrow{P} \mu = 0$

or $\bar{X}_n \xrightarrow{P} 0$

as $\bar{X}_n \xrightarrow{P} 0 \Rightarrow \bar{X}_n \xrightarrow{d} 0$

Exe: $\{X_n\}$ iid $U(0,\theta)$ & $Y_n = (\min\{X_1 \ldots X_n\}) \times n$, check wether Y_n conv in dist.

$\{X_n\}$ iid $U(0,\theta)$

$$\begin{aligned} P(Y_n \le y) &= P(n \min\{X_1 \ldots X_n\} \le y) \\ &= P\left(\min\{X_1 \ldots X_n\} \le \frac{y}{n}\right) \\ &= 1 - P\left(X_1 > \frac{y}{n}, X_2 > \frac{y}{n}, \ldots X_n > \frac{y}{n}\right) \\ &= 1 - \left(\frac{\theta - \frac{y}{n}}{\theta}\right)^n \end{aligned}$$

$0 \le \frac{y}{n} < \theta$

$\Downarrow$

$0 \le y \le n\theta$

$$F_{Y_n}(y) = \begin{cases} 0 & ; \ y < 0 \\ 1 - \left(\frac{\theta - y/n}{\theta}\right)^n & ; \ 0 \le y < n\theta \\ 1 & ; \ y \ge n\theta \end{cases}$$

as $n \to \infty$ $1 - \left(1 - \frac{y}{n\theta}\right)^n \to 1 - (e^{-y/\theta})$

for $n \to \infty$ $F_n(y) \longrightarrow F(y)$

$$F(y) = \begin{cases} 0 & ; \ y < 0 \\ 1 - e^{-y/\theta} & ; \ y \ge 0 \end{cases}$$

$$F'(y) = \underbrace{\frac{1}{\theta} e^{-y/\theta}}_{\text{density}}$$

$\therefore Y_n \xrightarrow{d} Y$ s.t Y is the exponential r.v with

$$f_Y(y) = \begin{cases} \frac{1}{\theta} e^{-y/\theta} & ; \ y \ge 0 \\ 0 & ; \ \text{otherwise} \end{cases}$$

Exe: Suppose $X_n \xrightarrow{d} X$ and $\{a_n\}$ be a real sequence s.t $a_n \to \infty$ as $n \to \infty$

does $\frac{X_n}{a_n}$ conv in probability?

As $Y_n = \frac{1}{a_n}$, $Y_n \xrightarrow{a.s} 0 \Rightarrow Y_n \xrightarrow{P} 0 \Rightarrow Y_n \xrightarrow{d} 0$

by slutsky's theorm

$$X_n Y_n \xrightarrow{d} 0$$

also as

$$X_n Y_n \xrightarrow{d} 0 \Rightarrow X_n Y_n \xrightarrow{P} 0$$

$$\therefore \frac{X_n}{a_n} \xrightarrow{P} 0$$

Exe: $\{X_n\}$ sequence of random variable with following density:

$$f_n(x) = \begin{cases} 1+\sin(2\pi n x); & \text{if } 0<x<1 \\ 0 & ; \text{otherwise} \end{cases}$$

check: does X_n conv in dist or not. If yes then find the limit

$$F_n(x) = \begin{cases} 0 ; & x<0 \\ \int_0^x (1+\sin(2\pi n x))\,dx & ; 0\le x<1 \\ 1 ; & x \geq 1 \end{cases}$$

$$\int_0^x (1+\sin(2\pi n x))\,dx = x + \left(-\frac{\cos(2\pi n x)}{2\pi n}\right)_0^x$$

$$= x + \left[\frac{1}{2\pi n} - \frac{\cos(2\pi n x)}{2\pi n}\right]$$

$$F(x) = \begin{cases} 0 ; & x<0 \\ x ; & 0\le x<1 \\ 1 ; & x \geq 1 \end{cases}$$

$\therefore$ X_n conv in dist

$$F'(x) = \begin{cases} 0 ; & \text{otherwise} \\ 1 ; & 0\le x<1 \end{cases}$$

central limit theorm:
(CTL)

Suppose $\{X_n\}$ is a sequence of iid random variables with mean μ, & variance σ^2.

Then

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma} \xrightarrow{d} N(0,1)$$

mean

standard deviation

$$f(t) = \frac{1}{\sqrt{2\pi}} e^{x^2/2}$$

$$P\left(\frac{S_n - n\mu}{\sqrt{n}\,\sigma} \le x\right) \to \int_{-\infty}^{x} f(t)\,dt \qquad \forall x \in \mathbb{R}$$

Observe : suppose $\mu = 0, \sigma^2 = 1$ then

$$\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1)$$

By law of large numbers

$$\frac{S_n}{n} \xrightarrow[a.s]{P} 0$$

(SLLN / WLLN)

$$\Rightarrow \frac{S_n}{n} \xrightarrow{d} 0$$

so $\frac{S_n}{n} \xrightarrow{d} 0$ but $\frac{S_n}{\sqrt{n}} \xrightarrow{d} N(0,1)$

we are magnifying this Quantity by $\sqrt{n}$ (zooming in) then conv to $N(0,1)$

$$\sqrt{n}\left(\frac{S_n}{n}\right) \longrightarrow N(0,1)$$

(finiteness of second moment by SLLN)

Exe: $\{X_n\}$ iid with mean 0, variance 1. Find the limiting dist of U_n as $n \to \infty$ where

$$U_n = \sqrt{n}\,\frac{\sum_{i=1}^{n} X_i}{\sum_{i=1}^{n} X_i^2}$$

finity $\frac{\sum X_i}{\sqrt{n}} \xrightarrow{d} N(0,1)$

Now by LLN: $\frac{\sum (X_i)^2}{n} \xrightarrow{P} \frac{E((X_i)^2)}{1}$

$$= \frac{E((X_i)^2) - (E(X_i))^2}{1}$$

$$= var(X_i)$$

$$= 1$$

$$\frac{\sum (X_i)^2}{n} \xrightarrow{P} 1$$

$$\Rightarrow \frac{\sum (X_i)^2}{n} \xrightarrow{d} 1$$

also $\frac{\sum (X_i)}{\sqrt{n}} \xrightarrow{d} N(0,1)$ (By CLT)

$$U_n = \frac{\frac{1}{\sqrt{n}} \sum X_i}{\frac{1}{n} \sum (X_i)^2} \xrightarrow{d} \frac{N(0,1)}{1} = N(0,1)$$

(By Slutsky's theorm)

Exe: If $x_n \xrightarrow{P} c$ $(c \neq 0)$
then

$$\frac{1}{x_n} \xrightarrow{P} \frac{1}{c}$$

as $x_n \xrightarrow{P} c$

$$P\left(\left|\frac{1}{x_n} - \frac{1}{c}\right| > \varepsilon\right) = P\left(\left|\frac{x_n - c}{x_n c}\right| > \varepsilon\right)$$

as $x_n \to c$ as $n \to \infty$

$P(|x_n - c| > \varepsilon') \to 0$

as $P(|x_n - c| > \varepsilon') \to 0$
as $n \to \infty$

$$P\left(\left|\frac{x_n - c}{x_n c}\right| > \varepsilon\right) = P\left(|x_n - c| > \varepsilon |x_n c|\right)$$

$$= P\left(|x_n - c| > \varepsilon |x_n||c|,\ |x_n| \geq |c|\right) \to 0$$

$$+ P\left(|x_n - c| > \varepsilon |x_n||c|,\ |x_n| < |c|\right)$$

$$\leq \underset{\to 0}{(\quad)} + P(|x_n| < |c|)$$

$$P(|x_n| < |c|) = P(|x_n| - |c| < 0)$$

$$= P(|c| - |x_n| > 0)$$

$$\leq P(|c - x_n| > 0)$$

$$\to 0$$

$$\therefore \frac{1}{x_n} \xrightarrow{P} \frac{1}{c}$$

Exe: $\{x_n\}$ iid $x_i \sim U(0,1)$

$$Z_n = \left(\prod_{j=1}^{n} x_j\right)^{1/n}$$

(i) show that Z_n conv in probab. and find limit

(ii) show that $\sqrt{n}(Z_n - e) \xrightarrow{d} N(0, e^2)$

now, $Z_n = \left(\prod_{j=1}^{n} x_j\right)^{1/n}$

then as $x_j \sim U(0,1)$

Z_n is a r.v
$\Rightarrow f(Z_n)$ is a r.v for f cont

then

$$\log(Z_n) = \log\left(\prod_{j=1}^{n} x_j\right)^{\frac{1}{n}}$$

$$= \frac{1}{n} \log\left(\prod_{j=1}^{n} x_j\right)$$

$$= \frac{1}{n} \sum_{j=1}^{n} \log(x_j)$$

now let $y_j = \log(x_j)$
then
$$\log(Z_n) = \frac{\sum y_j}{n} \xrightarrow{P} E(y_j)$$

$$\log(Z_n) = \frac{\sum y_i}{n} \xrightarrow{P} E(y_i)$$
$$y_i = \log(x_i)$$

here $E(Y_i) = \int_0^1 \log(x) f(x)\,dx$

$$= \int_0^1 \log(x)(1)\,dx$$

$$= (x\log x - x)_0^1$$

$$= \log(1) - 1 - (0)$$

$$= -1$$

then $Z_n \xrightarrow{P} e^{-1} = \frac{1}{e}$

Note: CLN comes when $(\quad) - (\quad)$
subtraction

as $Z_n \xrightarrow{P} \frac{1}{e} = c$

$$Z_n \xrightarrow{P} \frac{1}{e} \quad \text{then}$$
doubt

6^{th} Nov :

Conv in distribution: (Recap)

- $X_n \xrightarrow{d} X \not\Rightarrow X_n \xrightarrow{P} X$
  but if $X = c \Rightarrow X_n \xrightarrow{P} c$
- $X_n \xrightarrow{d} X, \; Y_n \xrightarrow{d} Y \not\Rightarrow X_n + Y_n \xrightarrow{d} X+Y$

Slutskys theorem:

$$X_n \xrightarrow{d} X, \quad Y_n \xrightarrow{d} c$$
$$\Rightarrow X_n + Y_n \xrightarrow{d} X + c$$
$$\Rightarrow Y_n X_n \xrightarrow{d} cX$$

Theorem: $X_n \xrightarrow{d} X$ iff $E[f(X_n)] \longrightarrow E[f(X)]$ for all bounded cont function

proof is too tecnical and we will be skipping it

Application: $X_n \xrightarrow{d} X$ and $g: \mathbb{R} \to \mathbb{R}$ cont function. Then

$$g(X_n) \xrightarrow{d} g(X)$$

Here, $E[f(Y_n)] \longrightarrow E[f(Y)]$. To show

$$f(Y_n) = f(g(X_n)) = f \circ g(X_n)$$

$$E[f(Y_n)] = E[h(X_n)], \quad h = f \circ g$$

as h is bdd. cont. and $X_n \xrightarrow{d} X$ we get

$$E[h(X_n)] \longrightarrow E[h(X)]$$
$$\Rightarrow E[f(Y_n)] \longrightarrow E[f(Y)]$$
$$\Rightarrow Y_n \xrightarrow{d} Y$$

Moment generating function:

let X be r.v, moment generating function (MGF) is defined as $M_X(t) = E[e^{tX}]$ $t \in \mathbb{R}$

Note: $e^{tX} \geq 0$, and therefore $M_X(t) \geq 0$

and $\therefore M_X(t)$ could be infinity.

we say MGF of X exist if $E[e^{tX}] < \infty$ in a nbd of 0.

Example:

① $X \sim Ber(p)$
$M_X(t) = pe^t + 1 - p, \; t \in \mathbb{R}$

② $X \sim Bin(n, p)$

$$M_X(t) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} e^{tk}$$

$$= (pe^t + 1 - p)^n$$

Note: $E[g(X)] = \sum_i g(x_i)\, p(x_i)$
$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$

③ $X \sim Poi(\lambda)$

$$M_X(t) = \sum_{n=0}^{\infty} \frac{\lambda^n e^{-\lambda}}{n!} e^{tn}$$

$$= e^{-\lambda} e^{\lambda e^t} = e^{\lambda(e^t - 1)}$$

④ Exe: calculate MGF of Exponential, Normal

exponent $M_X(t) = \frac{\lambda}{\lambda - t}$ $\quad$ $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$

$f_X(x) = \lambda e^{-\lambda x} \mathbb{1}_{x>0}$

Theorm: let X be a r.v. suppose the MGF of X is fined in $|t|<\varepsilon$, for some $\varepsilon>0$ then

(i) $E|X|^k < \infty \quad \forall k \in \mathbb{N}$

(ii) $M_X(t) = \sum \frac{t^n}{n!} E(X^n)$ for $|t|<\varepsilon$

(iii) $M_X(t)$ is infinitely diff in $|t|<\varepsilon$ and in particular

$$M_X^{(r)}(t) = \frac{d}{dt^r} M_X(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} E(X^{n+r})$$

and $M_X^{(r)}(0) = E(X^r)$

proof: (i) $e^{|tX|} \geq \frac{|tX|^n}{n!}$

$$\Rightarrow E|X|^n \leq \frac{n!}{|t|^n} E(e^{|tX|})$$

Note: $e^{|tX|} \leq e^{tX} + e^{-tX}$

therefore

$$E[e^{|tX|}] \leq E(e^{tX}) + E(e^{-tX})$$

now if we fix a t when $0<t<\varepsilon$

then

$$\underbrace{E(e^{tX}) < \infty}_{\text{given}} \text{ and } \underbrace{E(e^{-tX}) < \infty}_{\text{given}}$$

condition for $|t|<\varepsilon$

$E(e^{tX}) < \infty$

$$\therefore E[e^{|tX|}] < \infty$$

$$\Rightarrow E|X|^n \leq \frac{n!}{|t|^n} E(e^{|tX|}) < \infty$$

(ii) $Y_n = \sum_{k=0}^{n} \frac{t^k}{k!} X^k \longrightarrow e^{tX}$

we want to conclude,

$$E(Y_n) \longrightarrow E(e^{tX})$$

for this we need dominated convergence theorm (DCT)

$$|Y_n| \leq e^{|tX|}$$

$$\lim_{n\to\infty} E(Y_n) = E(e^{tX})$$

$$\Rightarrow \lim_{n\to\infty} \sum_{k=0}^{n} \frac{t^k}{k!} E(X^k) = E(e^{tX})$$

$$\Rightarrow E(e^{tX}) = \sum_{n=0}^{\infty} \frac{t^n}{n!} E(X^n) \text{ for } |t|<\varepsilon$$

DCT : $X_n \to X$ & $|X_n| \leq X$ and $E(X) < \infty$ then

$$E(X_n) \to E(X)$$

(iii) from (ii) we have power series expansion of $M_X(t)$ in $|t|<\varepsilon$. so, $M_X(t)$ is infinitely differentiable and r-the derivative of $M_X'(t)$ can be obtained from the ($|t|<\varepsilon$) term-wise derivative of power series. → do

(Rudin ch 8, Th 8.1)

Result: suppose X, Y are independent and MGF of X, Y exist in a nbd of 0. then

$$M_{X+Y}(t) = M_X(t) M_Y(t)$$

Theorm : (Uniqueness). Let X, Y be two random variable. If

$$M_X(t) = M_Y(t) \text{ for } |t|<\varepsilon, \varepsilon>0$$

then $X \overset{d}{=} Y$

Applications:

① Suppose $M_X(t) = e^{2(e^t-1)}$, $t \in \mathbb{R}$
find $P(X=1)$
from uniqueness theorm ($X \sim Poi(2)$ as $M_X(t) = e^{2(e^t-1)}$)
$X \sim Poi(2)$
$P(X=1) = e^{-2}(2)$

② $M_X(t) = \frac{1}{2}e^t + \frac{1}{4}e^{4t} + \frac{1}{4}e^{8t}$ $t \in \mathbb{R}$

If $P(Y=1) = \frac{1}{2}$ $P(Y=4) = \frac{1}{4}$ $P(Y=8) = \frac{1}{4}$

$$E[e^{tY}] = e^t\left(\frac{1}{2}\right) + \frac{1}{4}(e)^{4t} + \frac{1}{4}(e)^{8t}$$

$$Y \overset{d}{=} X \Rightarrow P(X=0) = P(Y=0) = 0$$

③ $X \sim Bino(n,p)$ $Y \sim Bino(m,p)$ & X, Y are independent
Find the dist of $X+Y$.

$$M_{X+Y}(t) = M_X(t) M_Y(t) = (pe^t + 1-p)^{n+m}$$

or $X+Y \sim Bin(n+m, p)$
or
$$X+Y \overset{d}{=} Bin(n+m, p)$$

④ $X \sim Poi(\lambda_1)$, $Y \sim Poi(\lambda_2)$, X & Y are independent. Then

$$X+Y \sim Poi(\lambda_1+\lambda_2)$$

as
$$M_X(t) = e^{\lambda_1(e^t-1)}$$
$$M_Y(t) = e^{\lambda_2(e^t-1)}$$
$$\Rightarrow \quad M_X(t)M_Y(t) = e^{(\lambda_1+\lambda_2)(e^t-1)}$$

⑤ $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$, X & Y are independent
then
$$M_X(t) = E[e^{tX}]$$
$$= \int_{-\infty}^{\infty} e^{tx} f(x)\,dx$$

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx}\left[\frac{1}{\sqrt{2\pi}} \times \frac{1}{\sigma_1}\right] e^{-\frac{(x-\mu_1)^2}{2\sigma_1^2}}\,dx$$

now, $X \sim N(0,1)$
$$Y = \mu + \sigma X \sim N(\mu, \sigma^2)$$
$$\Rightarrow M_Y(t) = E[e^{t(\mu+\sigma X)}]$$
$$M_Y(t) = e^{t\mu} M_X(t\sigma) \leftarrow \text{do}$$
to get $X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$

⑦ $M_X(t) = \left(\frac{1}{4}e^t + \frac{3}{4}\right)e^{\frac{1}{4}(e^t-1)}$ $t \in \mathbb{R}$

here
$$\underset{Y}{Y \sim Ber\left(\tfrac{1}{4}\right)} \qquad \underset{Z}{Z \sim Poi\left(\tfrac{1}{4}\right)}$$
then
$$X \sim Ber\left(\tfrac{1}{4}\right) + Poi\left(\tfrac{1}{4}\right)$$
or
$$P(X=1) = P(Y+Z=1)$$
$$\begin{array}{cc} Y & Z \\ 0 & 1 \\ 1 & 0 \end{array}$$
$$\therefore \quad = P(Y=1, Z=0) + P(Z=1, Y=0)$$
$$= P(Y=1)P(Z=0) + P(Z=1)P(Y=0)$$

characteristic function:

$$\phi_X(t) = E[e^{itX}] \quad t \in \mathbb{R},\ i = \sqrt{-1}$$
$$= E[\cos(tX)] + iE[\sin(tX)]$$

① ch. fn always exist
② if X, Y are ind,
$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$$
$$\phi_X(t) = \phi_Y(t) \Leftrightarrow X \overset{d}{=} Y$$
(uniqueness theorm)

③ continuity theorm: suppose $\{X_n\}$ is a seq^n of r.v and

suppose $\{X_n\}$ is seq of r.v & X is r.v

(a) $X_n \xrightarrow{d} X$ then
$\phi_{X_n}(t) \to \phi_X(t) \;\forall t \in \mathbb{R}$

(b) conversely $\phi_{X_n}(t) \to \phi_X(t)$ for $t \in \mathbb{R}$ (iff)
then
$X_n \xrightarrow{d} X$

Remark : proof of WLLN and CLT is based on continuity theorm.

proof of Slutsky's : $X_n \xrightarrow{d} X \quad Y_n \xrightarrow{d} c$
$X_n \sim F_n, \quad X_n + Y_n \sim G_n$
$X \sim F \quad X + c \sim G$
then let x be point of continuity of G. show that

$$F_n(x - c - \delta_k) + P(|Y_n - c| > \delta_k) \leq G_n(x) \leq F_n(x - c + \delta_k) + P(|Y_n - c| > \delta_k)$$

δ_k positive numbers

for $n \to \infty$ $\quad P(|Y_n - c| > \delta_k) \to 0$

$$\lim_{n\to\infty} F_n(x - c - \delta_k) \leq \lim_{n\to\infty} G_n(x) < \lim_{n\to\infty} F_n(x - c + \delta_k)$$

$$\Rightarrow \underbrace{F(x - c - \delta_k)}_{\text{countable amount}} \leq \lim_{n\to\infty} G_n(x) \leq F(x - c + \delta_k)$$

choose $\delta_k > 0$ s.t $(x - c - \delta_k)$, $(x - c + \delta_k)$ are points of cont of F and $\delta_k \to 0$ as $k \to \infty$.
Finally take $k \to \infty$ we get

$$\lim_{n\to\infty} G_n(x) = F(x - c) = G(x)$$

For WLLN, $\frac{S_n}{n}$, we can calculate $\phi_{\frac{S_n}{n}}(t)$ and show it cong to $\phi_\mu(t)$. one more proof of WLLN. ← do

central limit theorm: $\phi_{(\;)}(t)$ & show it cong to $\phi(\;)$
what we want ← do normal

∴ CLT also done

Slutsky's theorm proof is similar to :