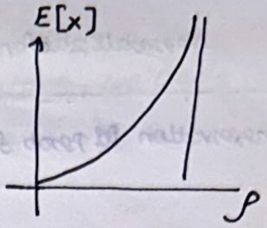
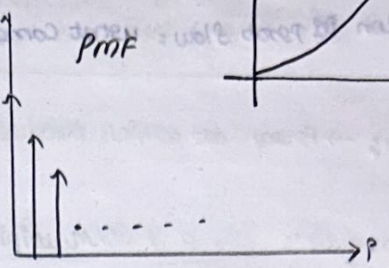
The diagram will be



1 customers : $\rho(1-\rho)$

0 customers : $(1-\rho)$

2 customers : $\rho^2(1-\rho)$



'queing theory'

'shifted - starting from 0'

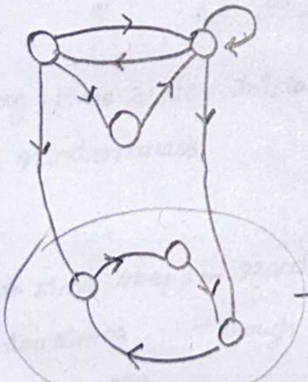
As ρ is small, E[no. of customers] = Finite

E["] = large, as $\rho \approx 1$ (or) 0.99 like

$\boxed{\rho = 1 = \frac{\lambda}{\nu} = 1}$ → '↑ = ↓' , As m is large , Expected no. of may ↑

Expectation [geometric] = $E[x_n] = \frac{1-p}{p} = \frac{1-\rho}{\rho}$ (Here.

Markov chains - III - Lecture - 18



→ Not periodic.

→ Recurrence (Lump)

Nice

* Aperiodic

$$\boxed{\lim_{n\to\infty} r_{ij}(n) = \pi_j}$$

↓

free of Initial cond

* Settles at Steady State value

$$\lim_{n\to\infty} P(x_n = j \mid x_0 = i) = \pi_j$$

$\boxed{\lim_{n\to\infty} P(x_n = j) = \pi_j}$ → As Independent of $x_0 = i$

x_0, x_n are appr independent as $n \to \infty$

↓

Steady state theorem.

∴ unique solution found by solving the balance eqn

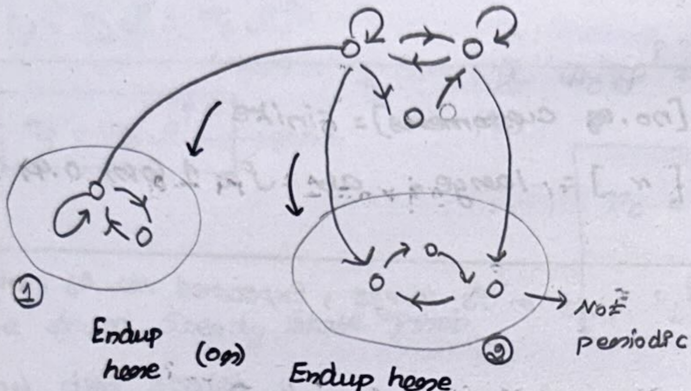
$$\pi_j = \sum \pi_k P_{kj} \quad , \quad j = 1, \dots m$$

Normalization equation $\sum_j \pi_j = 1$

Conservation of prob flow: what comes in must flow out

$\pi_K P_{Kj} \rightarrow$ Freq at which transition of these particular $K \rightarrow j$ occurs.

Multiple recurrent classes



If end up in chain 1 (recurrence): Solve this system of eqn for this chain alone

end up in 2: S.S prob inside the sub chain alone

solve linear eqn $\rightarrow$ The S.S eqn for the two chains seperately.

Starting at chain 1: S.S chain 1

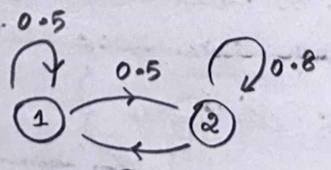
" 2: S.S chain 2

Above (Random) $\swarrow$ chain 1 $\searrow$ chain 2

(Prob it may go anywhere)

$\pi_1 = 2/7$

$\pi_2 = 5/7$



Assume process starts at 1

i) $P(X_1 = 1 \text{ and } X_{100} = 1) = P(X_1 = 1 \mid X_0 = 1) \cdot P(X_{100} = 1 \mid X_1 = 1, X_0 = 1)$

$= P_{11} \cdot P(X_{100} = 1 | X_1 = 1)$

↳ ∴ we need $X_1 = 1$ alone. we don't need how we went to X_1

$= P_{11} \cdot r_{11}(99)$

$X_0 = 1, X_1 = 1$

X_1 to $X_{100} = 1$ (99 steps)

$\boxed{r_{ij}(n) = \sum_k r_{ik}(n-1) P_{kj}}$

$= P_{11} \cdot (\pi_1)$

$P(X_{100} = 1 \text{ and } X_{101} = 2) = P(X_{100} = 1 | X_0 = 1) \cdot P(X_{101} = 2 | X_0 = 1)$

$= \pi_1 P_{12}$

$P(X_{100} = 1, X_{200} = 1) = r_{11}(100)\, r_{11}(100) \approx \pi_1 \pi_1 = \pi_1^2$

↓ Next 100 steps.

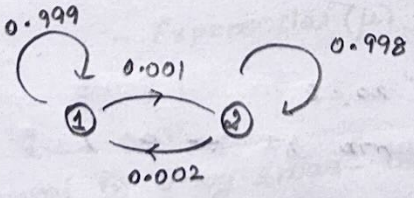
"we can appr — multistep trans prob by S.S.prob" when n is big

Is $n = 99$ or 100 is big?

* How long time scale: Initial State to be forgotten
* Enough randomness.

Every 10 time steps — randomness $(0.5 = P_{12})$

100 times — Enough randomness

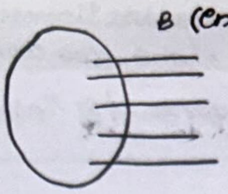


* It takes 1000 Steps (appro) to go to ② from ①
* $n = 100$ won't make sense → (we sure that we will be in 1)
* $n = 10000$ (<) → Before using this app

Erlang - Danish method — Famous problem

* How to set up: what it take to setup a phone system that how many lines should you set up for the community to communicate outside the world!

B (enough wire - not too much (expensive))



(No one wants to get a busy signal)

→ Calls originate as a poisson process rate λ,

→ Each call distributed exponentially (parameter μ)

→ B lines available

→ Discrete time intervals of (small) length δ.

B - Large enough (certainity)

modelling:

* How phone call gets initiated? → Poisson Process (No dependancies) (Different people Time independent)

* How long a call?

* $\boxed{\lambda = ?}$ → Rate of phonecall

Assume: Duration of phone call → R.V → exponential (μ)

$$E[\text{Duration}] = \frac{1}{\mu}$$

(most phone calls are short (so exp will suit in this)

*∴ many phones are short

Markov process

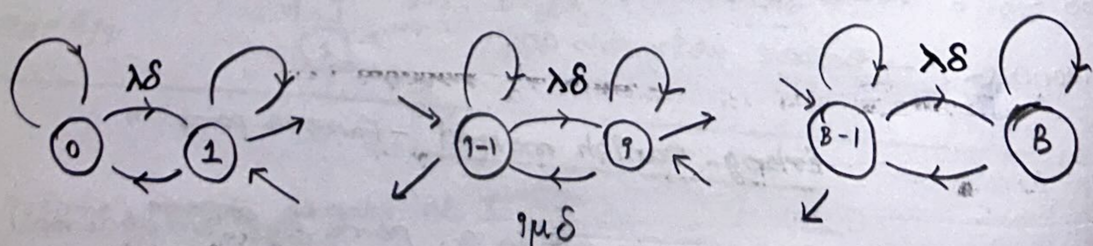
* Develop the theory of continuous time markov chain (or) → out of our scope

* Discretize time

0 → All phone lines are Free

B → All phone lines are busy

State
↓
ON
(or)
OFF



* ↑ by 1 if a phone call get placed
* ↓ by 1 if terminated
* No change - At a same time (negligible as time slot is small)

Arrival : depends upon rate

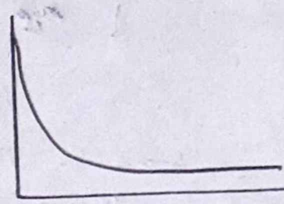
Termination: Depends on distribution of call.

$\lambda\delta$ = Rate × time

∴ Duration of call → Exponential (μ) distribution.

↓

exponential R.V can be thought of as an R.V (First arrival)



exponential (μ)

For a single phone call : $\mu \times \delta$ (mean × time)

For i calls active : $i\mu\delta$ [Termination prob]

(μ - Average call mean × time slot) × i calls

↓

scaling

* A Normal phone call has some mean of termination. We want to know - how many phone calls can be terminated in our time slot. (scale) then multiply by the total number of calls active.

<u>why $i\mu\delta$</u>

P → probability density in continuous case (usually multiplied by a small interval.

∴ Exponential (μ) = $\mu \times \delta \times i$

↓

i number of calls may going.

'As time interval is very small - collision is negligible'

<u>Note</u>

1) Incoming call is completely random ($\lambda \times \delta$)

2) Termination of call depends upon no. of calls.

Phone call: Exponential duration (I arrival in poisson Process)

* Single phone call : $\mu\delta$

* i numbers of phone calls : prob of termination: $i\delta\lambda$

$\downarrow$

Each phone call has a prob of $\mu\delta$ of being terminated.

Collectively one of them terminates: $i\mu\delta$'

$$\boxed{\text{Phone call terminate} = P(1 \text{ call terminate}) + P(2 \text{ call terminate}) + \cdots}$$

Note: 'we ignores the possibility: two calls terminates at same time'

Thinking

'poisson — Independent - memoryless'

$\mu\delta, \mu\delta, \mu\delta, \mu\delta \rightarrow \xrightarrow{i\mu\delta}$ Poisson process

'i times $\mu\delta$ running together'

After the first phone call terminates (1 arrival) - we then start fresh.

Chain

Balance equation:

'Comes left from right = from left'

'upward = downward transitions'

$\therefore$ for $i-1$ to i $\rightarrow$ we need to be in $i-1$ state with prob $\lambda\delta$ to go to i [So only one is going to be terminated

$$(\pi_{i-1})\lambda\delta = (\pi_i) i\mu\delta$$

$$\boxed{\pi_{i-1}\lambda = \pi_i i\mu}$$

$\sum \pi_i = 1$ (sub π_i)

$$\pi_i = \pi_0 \frac{\lambda^i}{\mu^i i!}, \quad \cancel{\pi_0 = \frac{1}{\sum_{i=0}^{B}}}$$

$$\boxed{\pi_0 = \frac{1}{\sum_{i=0}^{B} \frac{\lambda^i}{\mu^i i!}}}$$

If we know π_1, we are able to know P_{+1}

$\pi_B \rightarrow$ prob of being busy. (Small in a system: Real life)

$\pi_B = P(\text{busy}) = ?$

$\boxed{\begin{array}{l}\lambda = 30 \rightarrow \text{calls/minute} \\ \mu = 1/3 \rightarrow \text{Rate/minute}\end{array}}$

$E[\mu] = \frac{1}{\mu} = 3$ minutes

A call lasts for 3 minute

$\boxed{\lambda = 30 \text{ calls/minute}}$

Active on the Average: each minute 90 min of talking.

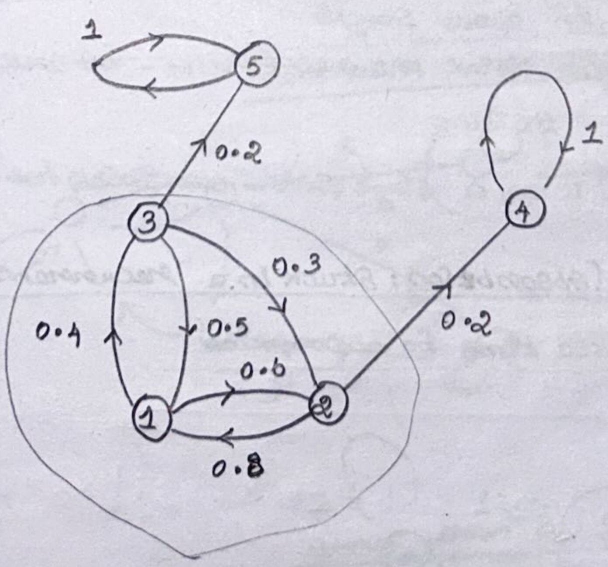
$\therefore B \approx 90$ lines will satisfy (But mean $\approx$ less or more)

$\pi_B = P(\text{busy}) = 1\% (0.01) [B \approx 106]$

what's the value of B - make my $P(\text{bust}) \approx 1\%$.

'Erlang - worked on this before markov chains were invented'

Absorption probabilites



Transition

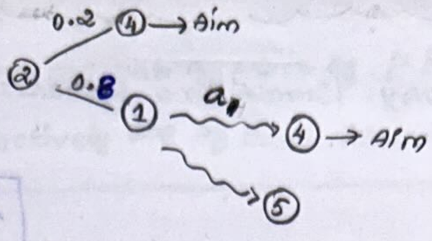
what's the prob of that: process eventually settles in state 4, given that the initial state is ?

* we have two recurrences to end up

Initial state : 4 : ending at 4 = 1

5 : " " = 0

Initial State : 2



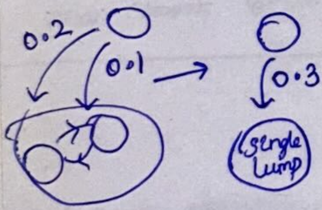
∴ Two ways → Directly (2 to 4)
→ Indirectly (2 to 1 to to 2 = then to 4)

$a_2 = 0.2 + 0.8\, a_1$

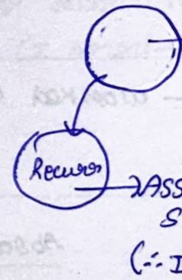
$$\boxed{a_i = \sum_j P_{ij} a_j \quad \text{for all others ;} \quad \to \text{Unique solution}}$$

∴ Seperate eqn for 1, 2, 3 → 3x3 system of eqn
↓
Once-solved - prob(absorbed) at State 4.

Recurrent → multiple stage

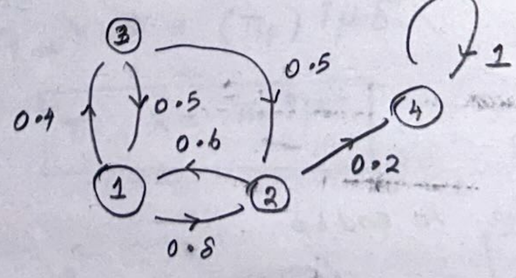


* No matter about stages
* once get in to the recurrence - we stuck inside the lump
* (view it as a single state - even it has stages)



How long its going to take? (Absorbtion : Stuck in a recurrence):

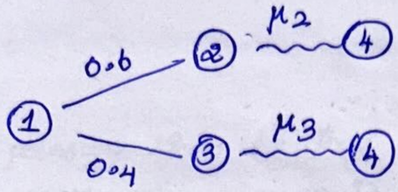
Expected time to absorption



Here: end up in 4.

* Find expected no. of transitions μ_i, until reaching the absorbing state, given that the initial state is i.

$\mu_1 \to$ Initial state: 1, end up at 4 (time)



$\boxed{\mu_1 = 0.6\mu_2 + 0.4\mu_3 + 1}$ → for first transition

↓

Account for first transition.

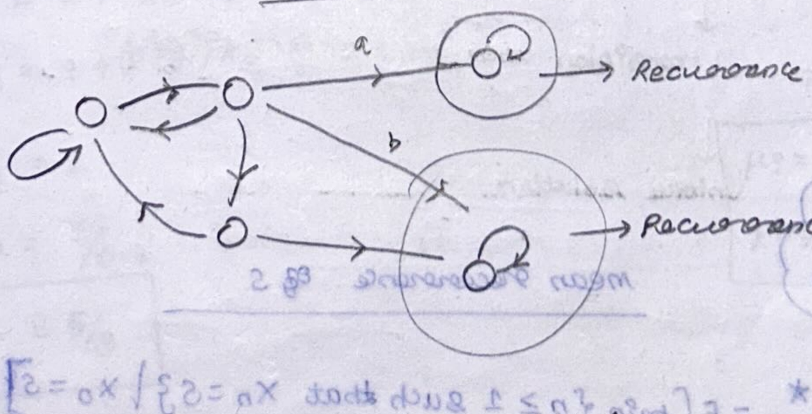
4 ∴ Already there 0.

* $\mu_i = 0$, for $i=4$ → Already there

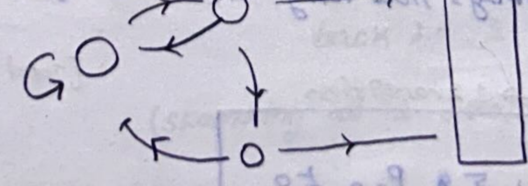
* All other i: $\boxed{\mu_i = 1 + \sum_j p_{ij}\,\mu_j}$ (Recursive → after reaching 4)

$\boxed{\text{Total transition time} = \text{time of first state} + \text{Expectation time from next time.}}$

more than 1 recurrence.

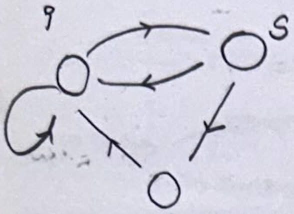


group them as 1



'Lump them as 1'

Another one

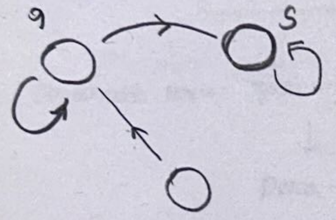


$S \rightarrow$ special state

i - initial state

'How long it takes for me to visit S for the first time?'

'we don't care after going to S'



'Remove transition out of S - Make self-transition'

we changed: what happens after S: — No change (as per our problem)

we want: Before S

Starting at S : $t_S = 0$ [No transition needed]

$$t_i = 1 + \sum_j P_{ij} t_j \, , \quad \text{for all } i \neq S$$

$\downarrow$

1 transition then next stages E[value]

$\downarrow$

unique solution

Transition: From one state to other

mean recurrence of S

$$t_S^* = E[\min\{n \geq 1 \text{ such that } X_n = S\} \mid X_0 = S]$$

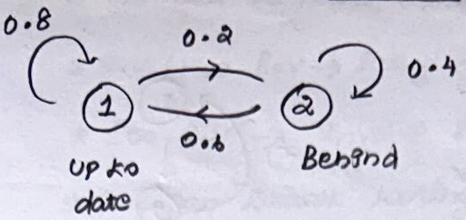
Start at S : chain randomly: How long it takes for me to back to S.

[Total prob theory]

$\rightarrow$ 1 transition.

$$\boxed{t_S^* = 1 + \sum_j P_{Sj} t_j}$$

Recitation



i) $t = 0, 1, 2, \ldots$

$x(t) = 2\ 2\ 2\ 1$

Sample path.

First passage time - First time he enters in to the new state.

solu:

$T_i \rightarrow$ First passage time to state 1. (Starting from state i) at $t=0$.

key: $t_2 = E[T_2] \rightarrow$ Expected value of Transition from 2.

$$\boxed{t_2 = 1 + \sum_{j=1}^{2} P_{2j}\, t_j}$$

Assume state 2

After one step we will be in other state (next step)

+

Expected time from that point on to enters 1

$t_2 = 1 + P_{21} t_1 + P_{22} t_2$

= After one transition we will be in 1

Now I'm in 1

$\boxed{\therefore t_1 = 0} \rightarrow$ No need for transient.

$t_1 = 0$

$= 1 + P_{22} t_2$

$t_2 = 1 + (0.4) t_2$

$0.6\, t_2 = 1$

$t_2 = 1/0.6$

$$\boxed{t_2 = 5/3}$$

Formula

Total transient time

$$\mu_i = 1 + \sum_j P_{ij}\, \mu_j$$

$$t_i = 1 + \sum_j P_{ij}\, t_j$$

2 → Revolve it; move to 1

b) T_1^* = first time to visit state 1, again back to 1.

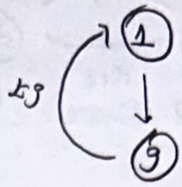
(starting at 1 from $t=0$)

$t = 0\ \ 1\ \ 2\ \ 3\ \ 4$

$x(t) = 1\ \ 2\ \ 2\ \ 1\ \ 1$ → Time 3 is the first time we back to where we started

$t_1^* = E[T_1^*]$

$$t_1^* = 1 + \sum_{g=1}^{2} P_{1g}\, t_g$$



$$t_1^* = 1 + P_{11} t_1 + P_{12} t_2$$

why works: Markov chain depends on current chain for future chains. Able to breakdown as recursion

$$t_1^* = 1 + P_{11} t_1 + P_{12} t_2$$

$$= 1 + 0 + 0.2\, t_2$$

As we are in state 1

$$= 1 + 0 + 0.2\left(\frac{5}{3}\right)$$

$$= 4/3$$

Unit - 4

'Laws of large numbers & inference'

Situation: Population (Penguin: at random chosen 1)

$$E[\text{height}] = \text{Average} = \frac{x_1 + \dots + x_n}{n}$$

(Equally likely to chose) $n \to \infty$

① measure each & everything → Tedious

② Pick some (Few) → measure - calculate average. [Estimate]

$$M_n = \frac{x_1 + \dots + x_n}{n} \quad \text{[Sample mean]}$$

$E[X]$ → over the entire population

↳ variable

M_n → Average over the smaller sample

↳ Random variable [As we have measured in random]

As $n \to \infty$ → close to $E[X]$

* we need to have closer $M_n \approx E[x]$
* One/Two R.V → Easy to handle
* 100 R.V → Involve summations (Difficult)
* (we can start taking limits → Simplify formulas)

Limit theorems - Relate prob & $E[x]$

* markov inequality
* chebyshev's - related inequality
* convergence
* converges (M_n) to true mean convergence? ⟶ 'weak law'

'Because there is a 'strong law' - Abstract'

---

markov inequality

* $X \geq 0$ (R.V non negative) → Assume discrete.

$$\boxed{E[x] = \sum_x x \, P_x(x)} \quad x \geq 0 \text{ (non negative)}$$

sum add all things $\geq$ sum add a group bigger than a constant.

$$\sum_x x \, P_x(x) \geq \sum_{x \geq a} x \, P_x(x)$$

↓

x starts from 0

Assume

$$Y_a = \begin{cases} 0 , & x < a \\ a , & x \geq a \end{cases}$$

$$\therefore \sum_x x \, P_x(x) \geq \sum_{x \geq a} a \cdot P_x(x)$$

↓

Above $x \geq a \rightarrow Y_a = a$

$x < a \rightarrow Y_a = 0$

$\therefore Y_a \leq X$

$$\boxed{E[Y_a] \leq E[X]}$$

$$\therefore E[X] \geq a \, P(X \geq a)$$

$$\boxed{P(X \geq a) = P(X = a) + \cdots}$$

'smallness of expected value to a statement about smallness of probability'

$E[(x-\mu)^2] \geq P((x-\mu)^2 \geq a^2) \cdot a^2$

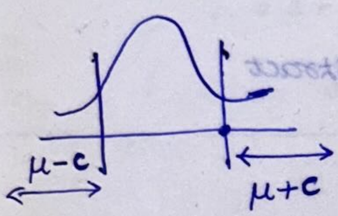
$Var(x) \geq P(|x-\mu| \geq a) \cdot a^2$

$\therefore$ Variance small - Prob far away from mean also small.

R.V $\rightarrow X$

(with finite mean μ & variance σ^2)

$$\sigma^2 = \int (x-\mu)^2 f_X(x)\,dx \geq \int_{-\infty}^{\mu-c} (x-\mu)^2 f_X(x)\,dx + \int_{\mu+c}^{\infty} (x-\mu)^2 f_X(x)\,dx$$



$$\geq c^2 \int_{-\infty}^{\mu-c} f_X(x)\,dx + c^2 \int_{\mu+c}^{\infty} f_X(x)\,dx$$

$\therefore$ when $Y_a = \begin{cases} c, & |x| \geq c \\ 0, & |x| < c \end{cases}$

$$\geq c^2 \cdot P(|x-\mu| \geq c)$$

$\downarrow$

Absolute value.

$\therefore \sigma^2 \geq c^2 \cdot P(|x-\mu| \geq c)$

$$\boxed{P(|x-\mu| \geq c) \leq \frac{\sigma^2}{c^2}}$$

when c is small, $\frac{\sigma^2}{c^2}$ is big.

'variance small - prob of being far away is small'

Let:

$c = K\sigma$

$$P(|x-\mu| \geq c) \leq \frac{\sigma^2}{(K\sigma)^2}$$

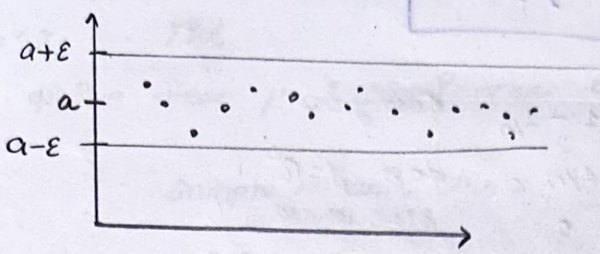
$$P(|x-\mu| \geq K\sigma) \leq \frac{1}{K^2}$$

$\downarrow$

what fraction of the class is 3 standard deviation away from the mean.

{chebyshev's inequality - handy - relate prob & E[X] s.}

limits of sequences



'look at interval'

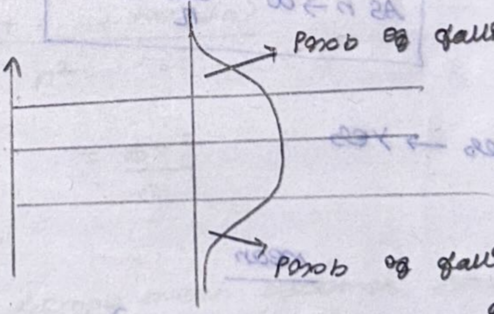
'since they converge to a — the values must lie inside a+ε & a−ε'

'Band of +ve band', for every $\varepsilon > 0$, there exists n_0, such that for every $n \ge n_0$, we have

$$|a_n - a| \le \varepsilon$$

As $\varepsilon \to$ smaller, wait for more time. [very small band]

Dealing with probabilities — R.V

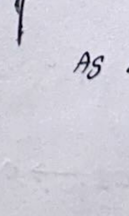


Prob of falling outside the band falls eventually to zero

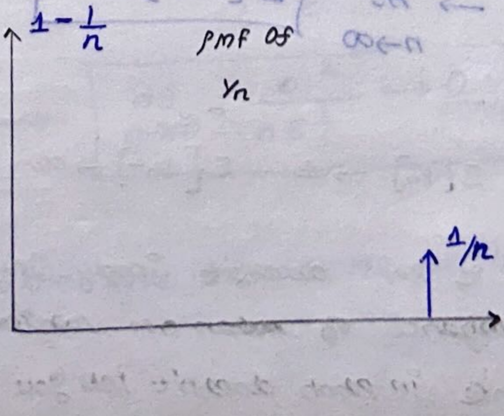
Prob of falling outside the band falls eventually to zero.

For every $\varepsilon > 0$,

$$\boxed{\lim_{n\to\infty} P(|Y_n - a| \ge \varepsilon) = 0}$$



As time goes - will lie entirely in the band

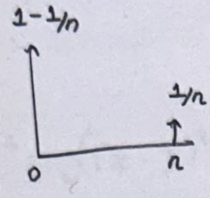


For every $\varepsilon > 0$:

(limit is zero).

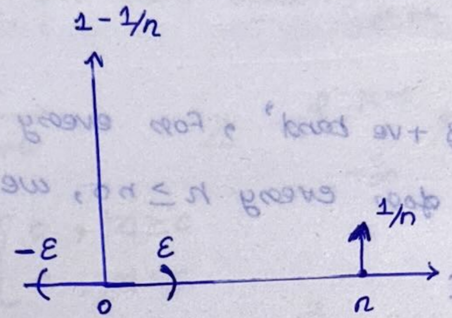
for any $\varepsilon' > 0 \rightarrow$ exists a n_0 such that every $n > n_0$

$$\boxed{P(\quad) \leq \varepsilon'}$$



$$\rightarrow \begin{cases} 1 - 1/n & \text{for } y = 0 \\ 1/n & \text{for } y = n \\ 0 & \text{elsewhere.} \end{cases}$$

As time goes $\rightarrow$ Prob of falling outside the band ≈ 0



* 'Define a range'

* Falling outside the range is

$$\boxed{\text{As } n \rightarrow \infty \quad \frac{1}{n} \rightarrow 0}$$

Does Y_n converges $\rightarrow$ yes

<u>mean</u>

$$E[Y_n] = n\left(\frac{1}{n}\right) + 0\left(1 - \frac{1}{n}\right)$$

$$= 1$$

$$E[Y_n^2] = n^2\left(\frac{1}{n}\right) + 0^2\left(1 - \frac{1}{n}\right)$$

$$= n \rightarrow \underset{n \rightarrow \infty}{\text{As}} \quad \boxed{E[Y_n^2] = \infty}$$

<u>Note</u>: Convergence $\rightarrow 0$, $E[Y_n] \rightarrow 1$, $E[Y_n^2] \rightarrow \infty$

'Convergence to 0 of R.V doesn't imply anything about convergence of mean or variance'

'Convergence in prob doesn't tell you whole story' "convergence → tells: Tail prob will be very small" → Doesn't tell how fast it goes $\left[\lim_{n\to\infty} P(\text{tail prob}) = 0\right]$

Sequence of R.V

$X_1, X_2, \ldots$ iid

finite mean μ & variances σ^2

Sample mean $= M_n = \dfrac{X_1 + \cdots + X_n}{n}$ → w.k. how to find distribution of sums of Ind. variable (Convolution over & over)

$$E[M_n] = \frac{E[X_1] + \ldots + E[X_n]}{n}$$

$$= \frac{n\mu}{n} = \mu$$

* Expectations: Some kind of average
* Sample mean: Average of heights of penguin (we collected).

$$\text{var}(M_n) = \text{var}\left(\frac{X_1}{n}\right) + \cdots + \text{var}\left(\frac{X_n}{n}\right)$$

$$= \frac{\text{var}(X_1) + \cdots + \text{var}(X_n)}{n^2}$$

$$= \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n}$$

The variance of the sample mean becomes smaller & smaller (Large sample - randomness will be small).

Chebyshev's Inequality

$$P(|M_n - \mu| \geq \varepsilon) \leq \frac{\text{var}(M_n)}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon^2}$$

more prob away from band will be zero.

$$\boxed{\text{As } n\to\infty \quad \frac{\sigma^2}{n\varepsilon^2} \to 0}$$

Approaches the true mean: As we have large sample mean.

'Sample mean - converges in prob to true mean' as large sample size.

<u>what fraction prefers coke to pepsi</u>

i^{th} person polled:

$$x_i = \begin{cases} 1, & \text{yes} \\ 0, & \text{No} \end{cases}$$

f : fraction of population that chooses coke over pepsi.

$$M_n = \frac{X_1 + \cdots X_n}{n}$$ [yes in our sample]

Goal: 95% confidence of $\le 1\%$ error:

$$P(|M_n - f| \ge 0.01) \le 0.05$$

40% → may be 41% or 39%,

until we take poll to the entire population — we can't able to tell exactly.

Goal: 95% confidence of 1% error: (less than)

'can't give hard guarantee - I may took wrong group'

'I can say confidence'

$P(|M_n - f|) \ge 0.01 \le 0.05$
↳ (1 - confidence)

$Var(M_n) = \frac{\sigma^2}{n}$

<u>Chebsheb's inequality</u>

P (getting an answer that's more than 0.01 away from the true answer) $= P(|M_n - f| \ge 0.1) \le \frac{\sigma^2_{M_n}}{(0.01)^2}$

$$\le \frac{\sigma_S^2}{n(0.01)^2}$$ → standard

$$\le \frac{1}{4n(0.01)^2}$$

∴ probability I am wrong must be less than 5%.

$$P(|M_n - f| \geq 0.01) \leq 0.05.$$

M_n → my average

f → Actual

≥ 0.01 → more than 1% difference

∴ $M_n \leq 1$

$f \leq 1$

'chebsheb's inequality'

$$\boxed{P(|x-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2} = \frac{var(x)}{\varepsilon^2}}$$

$$P(|M_n - f| \geq 0.1) \leq \frac{var(M_n)}{(0.01)^2}$$

$$\leq \frac{\sigma_x^2}{n(0.01)^2}$$

$$\leq \frac{\left(\frac{1}{4}\right)}{n(0.01)^2} \leq 0.05$$

only when $n \approx 50000$

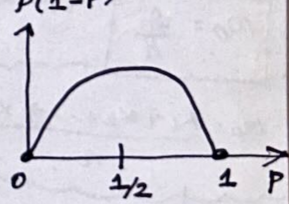
variance = ?

'Bernoulli distribution

$var P(x) = P(1-P)$'

'worst case'

$P(1-P)$



$0 \leq P \leq 1$

$E[P] = 1/2$

worst case : 1/2

variance: $P(1-P)$

: $1/2(1-1/2)$

: $1/4$

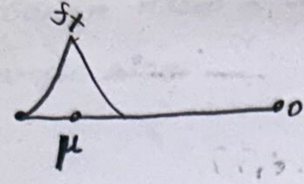
Realworld:

In election polls : Sample size will be around 1000. (50,000 - Too much)

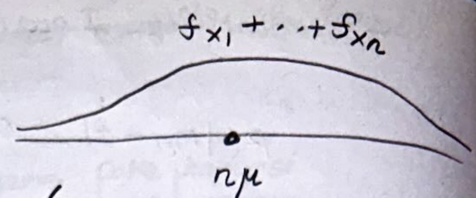
As 50,000 - Too much

we can cut corners (3 places)
→ Accuracy (3% instead of 1%) [Saves you a factor of 10]
→ Confidence (10% say)
→ other ways of calculating inequality other than chebsheb's inequ.

∴ chebsheb's inequality - Not so 'tight' → Not accurate.



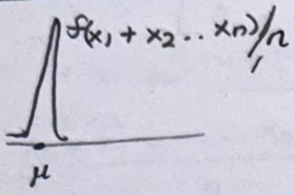
$\longrightarrow$ Add distributions (x's)



No scaling.

Variance also ↑ (a wide distributi)

weak law:



$\longrightarrow$ scaled

(distribution is concentrated)

weak law told that very conc. around the original mean.

Different scalings of M_n

① sample mean ≈ original mean $\left(\frac{\mu+\mu+\mu+\dots\mu}{n}\right) \approx \mu$ → Scaled

② Adding all distributions: $(f_{x_1} + \dots + f_{x_2})$ = New mean = $n\mu$ → Not scaled.

③ Scale by $\sqrt{n}$ $\boxed{S_n = x_1 + \dots + x_n \quad \text{var}(S_n) = n\sigma^2)}$

$$\frac{S_n}{\sqrt{n}} \Rightarrow \text{var}\left(\frac{S_n}{\sqrt{n}}\right) = \frac{\text{var}(S_n)}{n} = ?$$

$$\boxed{M_n = \frac{S_n}{n}}$$

$$M_n = \frac{x_1 + x_2 + \dots + x_n}{n}$$

$$\downarrow$$

$$\text{var}(M_n) = \frac{\sigma^2}{n}$$

$$\boxed{\begin{aligned} \text{var}(S_n) &= \text{var}(X_1) + \dots + \text{var}(X_n) \\ &= n\,\text{var}(X_1) \\ &= n\sigma^2 \end{aligned}}$$

$$\boxed{\text{var}\left(\frac{S_n}{\sqrt{n}}\right) = \frac{\text{var}(S_n)}{n} = \frac{n\sigma^2}{n} = \sigma^2}$$

In this way: S_n changes → variance doesn't change (doesn't become narrower or wider)

$$\text{scale} \begin{cases} \dfrac{S_n}{n} & \text{(narrow)} \\ S_n & \text{(wide)} \\ S_n/\sqrt{n} & \text{(No change)} \end{cases}$$

Central limit theorem

$\boxed{var(S_n) = \sigma^2 = n\sigma^2}$

'Standardized' $S_n = X_1 + \cdots + X_n$.

$$Z_n = \frac{S_n - E[S_n]}{\sigma_{S_n}} = \frac{S_n - nE[X]}{\sqrt{(n\sigma^2)}} = \frac{S_n - nE[X]}{\sqrt{n}\cdot\sigma}$$

* zero mean [$\because S_n$ has a mean of $E[S_n]$]
* unit variance. [S_n has a variance of σ_{S_N}]

If S_n has +ve mean

$$\frac{S_n - nE[X]}{\sqrt{n}\,\sigma}$$

→ shifting right

→ standardized.

→ How many S.D away from the mean

How much S_n away from $E[S_n]$

* Z → standard normal r.v (zero mean, unit variance)

Theorem

$$P(Z_n \le c) \rightarrow P(Z \le c)$$ [For every c]

↓ standard normal.

$P(Z \le c)$ is the std. normal CDF, $\Phi(c)$ available from the normal table.

$\therefore$ standarize Z_n to Z

'Than calculate for Z' → shortcut

S_n → we are interested.

$$\boxed{S_n = \sigma\sqrt{n}\,Z_n + nE[X]}$$ → If we calculate Z_n approximately, we can for S_n.

Central limit theorem:

As $n \to \infty$ (large)

$$P(Z_n \leq c) \to P(Z \leq c)$$

we can pretend Z_n is Z

Then

$S_n \to$ standard normal variable.

Functions of standard normal is also a standard normal

S_n is a linear function of Z_n

Key:

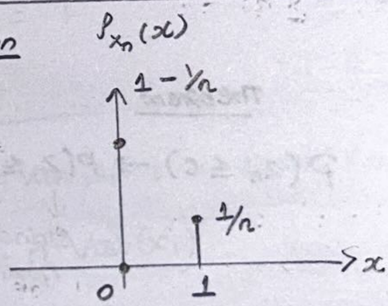
mathematically central limit theorem

* doesn't talk about distribution of S_n
* cumulative distribution of Z_n s.

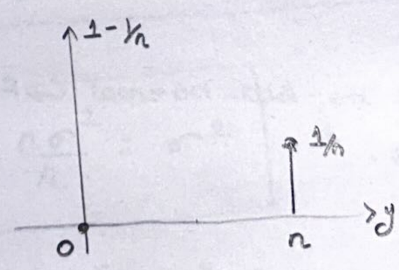
pretending S_n is normal

Recitation

X_n : $P(X_n = 0) = 1 - 1/n$
$P(X_n = 1) = 1/n$



Y_n : $P(Y_n = 0) = 1 - 1/n$
$P(Y_n = n) = 1/n$



Convergence & mean

$E[X_n] = 1(1/n) + 0(1 - 1/n)$
$= 1/n$

$var(X_n) = E[(X_n - 1/n)^2]$
$= (0 - \frac{1}{n})^2 \cdot (1 - 1/n) +$
$(1 - 1/n)^2 (1/n)$

$E[Y_n] = 0(1 - 1/n) + n(1/n)$
$E[Y_n] = 1.$

$var(Y_n) = E[(Y_n - 1)^2]$
$= (0-1)^2 (1 - 1/n) +$
$(n-1)^2 (1/n)$

$$\text{var}(X_n) = \sum_n (x_n - \mu)^2 \cdot P_X(n)$$

$$= \frac{n-1}{n^2}$$

$$= \frac{1}{n} - \frac{1}{n^2}$$

AS $n \to \infty$

$$\text{var}(X_n) = 0$$

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

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

$$= \left(1 - \cancel{\frac{1}{n}}\right) + \left(n - 2 + \cancel{\frac{1}{n}}\right)$$

$$= n - 1$$

AS $n \to \infty$

$\text{var}(Y_n) =$ goes to ∞

<u>chebsheb's inequality</u>

$$P(|X_n - \mu| \geq \varepsilon) \leq \frac{\text{var}(X_n)}{\varepsilon^2}$$

$$P(|X_n - 1/n| \geq \varepsilon) \leq \frac{n-1}{n^2 \varepsilon^2}$$

For $\varepsilon > 0$:

$$\lim_{n \to \infty} P(|X_n - 0| \geq \varepsilon) = 0$$

<u>Take</u>

$$\lim_{n \to \infty} P(|X_n| \geq \varepsilon) \leq P\left(\left|X_n - \frac{1}{n}\right| \geq \varepsilon + \frac{1}{n}\right)$$

$\downarrow$ chebsheb's inequality

$\therefore \frac{1}{n} \to$ some constant

$$\text{var}\left(\left|X_n - \frac{1}{n}\right|\right) = \text{var}(|X_n|)$$

$$\leq \frac{n-1}{n^2} \cdot \left(\frac{1}{(\varepsilon + 1/n)^2}\right)$$

$$\leq \text{AS } n \to \infty \quad 0 \cdot \frac{1}{\varepsilon^2}$$

$$\leq 0$$

<u>why?</u>

$X_n \geq \varepsilon \cap$
$-X_n \geq \varepsilon$
$X_n \leq \varepsilon \cap$

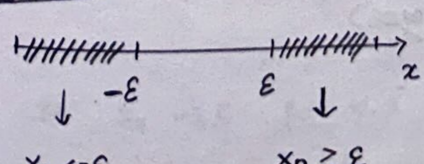
$$P(|X_n| \geq \varepsilon) \leq P\left(\left|X_n - \frac{1}{n}\right| \geq \varepsilon + 1/n\right)$$

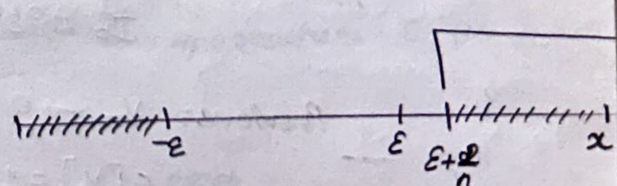
$X_n - \frac{1}{n} \geq \varepsilon + \frac{1}{n}$

$\boxed{X_n \geq \varepsilon + \frac{2}{n}}$

$-X_n + \frac{1}{n} \geq \varepsilon + \frac{1}{n}$

$X_n \leq -\varepsilon$





$\therefore$ The $P(|X_n - \frac{1}{n}| \geq \varepsilon + \frac{1}{n})$ captures less values than $P(|X_n| \geq \varepsilon)$.

$$\boxed{P(|X_n| \geq \varepsilon) \leq P(|X_n - \tfrac{1}{n}| \geq \varepsilon + \tfrac{1}{n})}$$

ii) $P(|Y_n - \mu| \geq \varepsilon) \leq \dfrac{var(Y_n)}{\varepsilon^2}$

$$P(|Y_n - 1| \geq \varepsilon) \leq \frac{n-1}{\varepsilon^2} \xrightarrow{n \to \infty} \infty$$

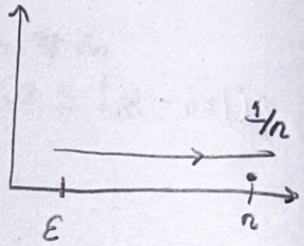
Chebyshev's inequality doesn't tell me anything new?

c) $P(|Y_n| \geq \varepsilon) = \dfrac{1}{n}$, $\varepsilon < n$

$$\lim_{n \to \infty} P(|Y_n - 0| \geq \varepsilon)$$

$$= \lim_{n \to \infty} \frac{1}{n} = 0$$

$Y_n \to 0$



As 0 is skipped, Y_n is converged to 0.

skipping 0,
$P(|Y_n| \geq \varepsilon) = 1/n$

d) If a sequence of R.V converges in probability to a, does the corresponding sequence of expected values converge to a? prove or give a counter example.

solu:

$Z_n \to c$, $n \to \infty$ (In prob)

$\lim_{n \to \infty} E[Z_n] \to c$ (Is this true)

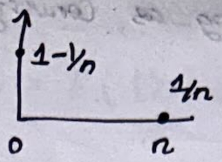
True: Because intuitively it looks that Z_n almost concentrates on c. So Expected value may be also c.

Is this true?

Previous: $Y_n \to 0$ (converges)

$E[Y_n] = 1$

overview:



$\therefore$ $E[X] = 1$ (As At 0, $n\uparrow$, going towards 1)

At 1, going towards 1)

$\downarrow$

more likely $0\cdot(1) + \frac{1}{n}\cdot n = 1$.

From image:

mostly $Y_n \to$ converges to 0 (position)

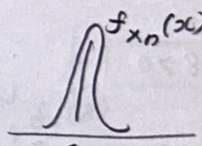
$\therefore$ our inequality is not enough strong to tell us whether the $E[x]$ converges to same as our x do.

---

c) $X_n \to c$ in mean square.

$\lim_{n\to\infty} E\left[(X_n - c)^2\right] = 0$, $E[X_n] = c$

$\lim_{n\to\infty} \mathrm{var}(X_n) = 0 \longrightarrow$ Intuitively: $E[X_n] = c \to \therefore$ concentrates on c. So variance will be zero.



a) use markov inequality to show that converges in the mean square implies convergence in probability.

Sol:

$X_n \to c$ in mean square

$X_n \to c$ in probability

$P(|X_n - c| \geq \varepsilon) \to 0$ as $n \to \infty$

Squaring

$$\boxed{P\left((X_n - c)^2 \geq \varepsilon^2\right) \leq \frac{E\left[(X_n - c)^2\right]}{\varepsilon^2}} \to \text{markov inequality.}$$

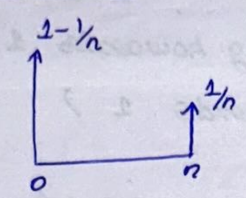
$\to$ Approaches 0 as $n\to\infty$

If $X_n \to c$ in M.S $\Rightarrow X_n \to c$ in prob

'Is this true'

Prob convergence in prob is not strong as Convergence in mean square

From previous example:



As $n \to \infty$, Y_n converges to 0 $[1 - 0 = 1]$

$$E[(Y_n - 0)^2] = 0\left(1 - \frac{1}{n}\right) + n^2 \cdot \frac{1}{n}$$

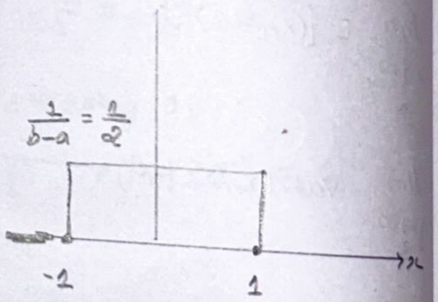
$$= n \to \text{as } n \to \infty$$

$$= \infty$$

∴ Convergence of Y_n ≠ Convergence of $\text{var}(Y_n)$.

Recitation

X ~ uniform [−1, 1] → Find convergence.



X_1, X_2

a) $X_i \to c$ in prob, as $i \to \infty$?

$\boxed{\therefore\ \varepsilon > 0}$

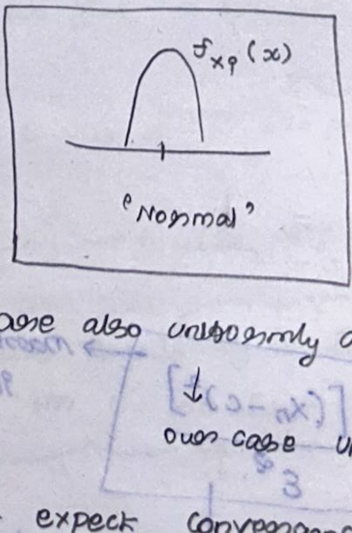
$$\lim_{i \to \infty} P(|X_i - c| \ge \varepsilon) = 0$$

mean = 0

$X_i = ?$

$Y_i = \frac{X_i}{i}$

$Z_i = (X_i)^i$



* $X_1, X_2, \ldots$ are also uniformly distributed & independent.

↓

our case uniform.

"we can't expect convergence in this case"

b) $Y_i = \frac{X_i}{i}$ $\qquad |X_i| \le 1$

$|Y_i| \le 1/i$ $[Y_i \to 0$ as $i \to \infty]$

verify

$P(|Y_q - 0| \geq \varepsilon) = P(|Y_q| \geq \varepsilon) = P\left(\frac{1}{q} \geq \varepsilon\right)$

$\lim_{q \to \infty} P(|Y_q - 0| \geq \varepsilon) = 0$ $\quad \therefore \begin{cases} 1 & \text{if } q \leq 1/\varepsilon \\ 0 & \text{if } q > 1/\varepsilon \end{cases}$

$\therefore$ In the $\lim_{q\to\infty}$ $\rightarrow$ goes to zero $(q > 1/\varepsilon)$

$Y_q \rightarrow 0$ as $q \rightarrow \infty$

c) $Z_q = (X_q)^q$ $\qquad$ mean $= 0$ $\qquad$ $|X_q| \leq 1$

$Z_q = (X_q)^q$

AS $X_q \rightarrow 0$

$Z_q \rightarrow 0$

$P(|Z_q - 0| \geq \varepsilon) = P(|X_q^q| \geq \varepsilon)$

$= P(X_q \leq -\varepsilon^{1/q}$ (or) $X_q \geq \varepsilon^{1/q})$

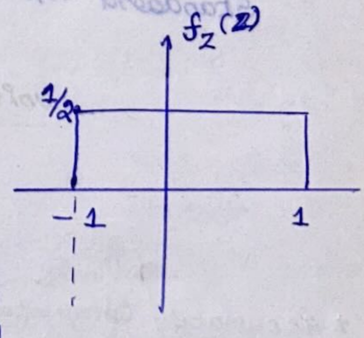
$\begin{cases} 0 \\ 2 \cdot \frac{1}{2} \cdot (1 - \varepsilon^{1/q}) \end{cases}$

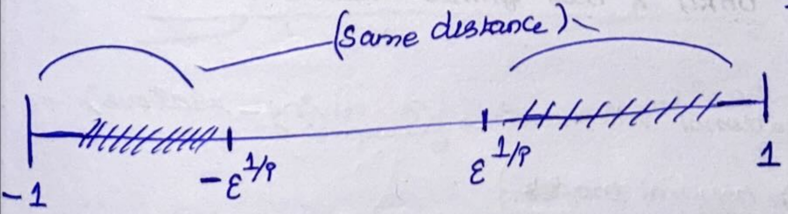
$\varepsilon > 1$ $[\therefore X_q \rightsquigarrow [-1, 1]]$

$\hookrightarrow \varepsilon^{1/q} > 1$ [Beyond $|1|$. X_q is zero]

$0 < \varepsilon < 1 \Rightarrow \varepsilon^{1/q} < 1$

AS $q \rightarrow \infty$, $(1 - \varepsilon^{1/q}) \rightarrow 0$.





$\rightarrow X_n \rightsquigarrow [-1, 1]$

uniform.

$$\boxed{P(\text{something}) = f_Z(z) \cdot \delta}$$

$\therefore \frac{1}{2} \times \underbrace{(-\varepsilon^{1/q} + 1)}_{\text{Both distance same.}} + (1 - \varepsilon^{1/q}) \cdot 1/2$

$\therefore 2\left(\frac{1}{2}\right)(1 - \varepsilon^{1/q})$ $\longrightarrow$ AS $q \rightarrow \infty$ $\longrightarrow$ $Z_q \rightarrow 0$.

Key: AS $Y_n \to c$ doesn't mean

$var(Y_n) \to c$

$E[Y_n] \to c$

Lecture - 20

central limit theorem

$X_1, \dots, X_n \to$ iid $\to$ finite variance σ^2

* $X_1 + \dots + X_n = S_n \to$ can't be said normal distribution

$\therefore n \to \infty \to S_n$ (widely spread)

$\therefore$ Standardizing:

$$Z_n = \frac{S_n - E[S_n]}{\sigma S_n} = \frac{S_n - nE[X]}{\sqrt{n}\,\sigma}$$

$E[Z_n] = 0$, $var(Z_n) = 1$ [standard normal distribution]

Theorem:

$P(Z_n \le c) \to P(Z \le c)$

$\downarrow$

Standard normal CDF (available as table)

universal - Irrespective of X's distribution

until X has finite mean & variance

* Accurate computational shortcut $\to$ (convolving - tedious)
* Justification of normal models.

quick way
(we satisfy with appr answers)

why we are interested?

Noisy phenomenon: By lot of independent : described by a normal R.V

A particle inside a liquid $\to$ Due to random collision how much the particle is displaced $\to$ very well modeled by normal R.V why?

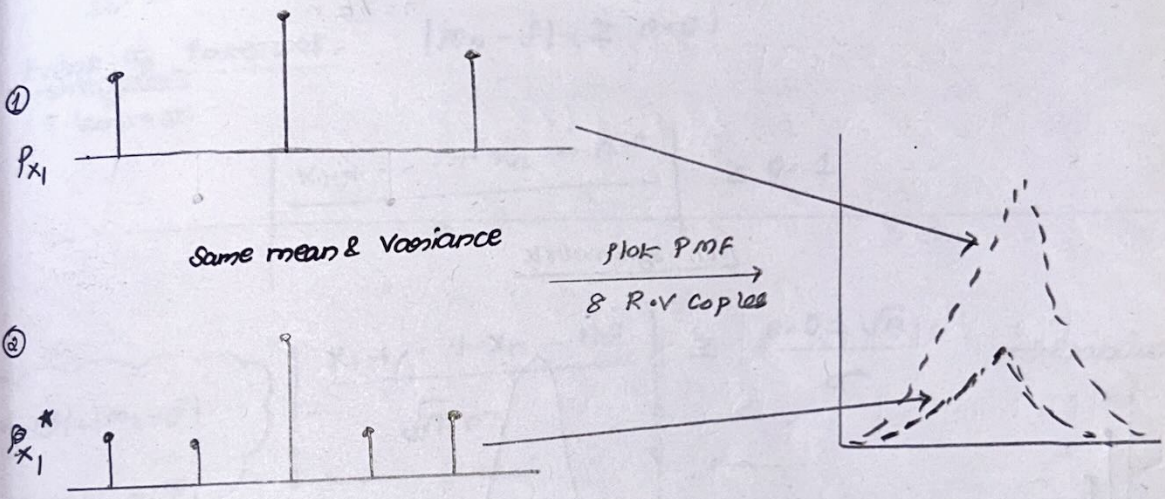
* Position of that molecule is decided by the cumulative effect of lots of random hits. (by molecules) [Brownian motion]

* movement of prices in financial market.

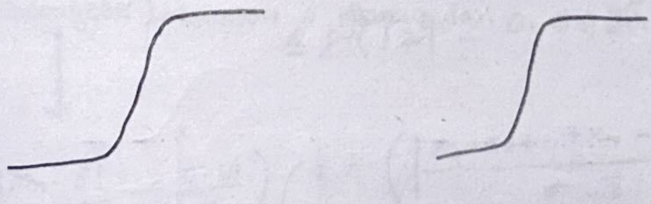
[More heavy tailed - Extreme events are more little likely to occur].

Central limit theorem

* Convergence of CDF of Z_n to normal CDF.



CDFS



Comparing with normal CDF - similar.

Practice: calculate probability.

∴ $S_n \rightarrow$ linear function of Z_n (normal distribution)

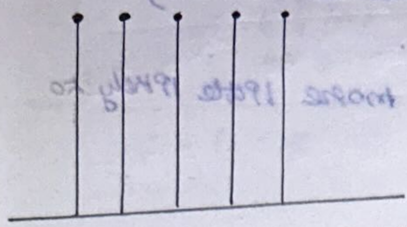
Limit theorem ($n \rightarrow \infty$)

* can we use it for moderate numbers of n?

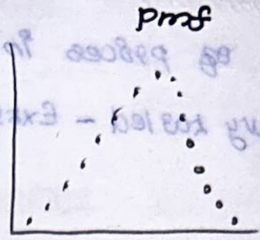
↓

Yes ($n \approx 15$ or 20 → Good answers)

plots



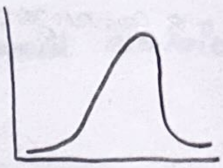
Convolve this with itself →



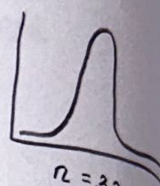
4 - discrete uniform

↓

Symmetric.



n = 16

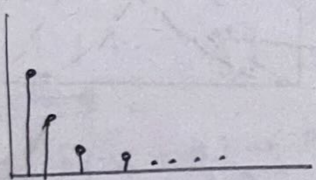


n = 32

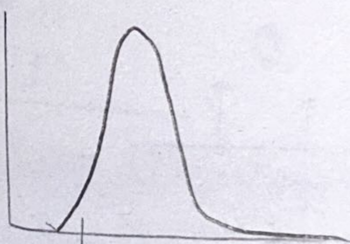
↓

Exactly like normal

Non-symmetric



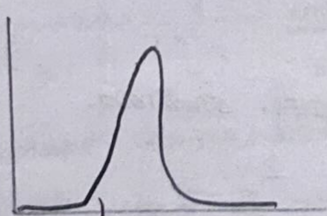
→



n=4

↓

Not exactly a normal (Assymmetry)



↓

Still (n=16) asymmetry

↓

Now better (n=32)

moral: It might vary a little bit (based on n)

larger n → great approximation

Polling

↓