Recall :

Theorem 1 :- (Weak law of large numbers

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a Probability space. Let $\{X_n\}_{n\geq 1}$ be a sequence of independent random variables such that

$\forall n \geq 1$, $E[X_n] = \mu < \infty$ and $Var(X_n) \leq \tilde{\sigma} < \infty$

Then, $\forall\, \varepsilon > 0$

$$\mathbb{P}\left(\left| \frac{1}{n} \sum_{i=1}^{n} X_i - \mu \right| > \varepsilon \right) \longrightarrow 0 \text{ as } n \to \infty.$$

Theorem1 (Monotone Convergence Theorem)

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

let $\{X_n\}_{n\geq 1}$, X be random variables on it.

Suppose $E[X_1] > -\infty$ and $X_n \uparrow X$

Then $E[X_n] \uparrow E[X]$.

Theorem 2 (Bounded Convergence Theorem) :

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

let $\{X_n\}_{n\geq 1}$, X be random variables on it.

Suppose - $\exists\, K > 0$ st $|X_n| \leq K$ $\forall\, n \geq 1$

and $X_n \longrightarrow X$ as $n \to \infty$.

Then $E[X_n] \longrightarrow E[X]$ as $n \to \infty$

13 First and Second Moment Method.

In general the Moment method is a method to control the probability that a random variable takes values far away from its expectation by its moments (i.e. by $E[X^k]$ $k \geq 1$).

Already seen in Section 8, Proposition 1

$X \geq 0$ r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$ then

(Markov's Inequality)

$$\mathbb{P}(X \geq \alpha) \leq \frac{E[X]}{\alpha} \quad - (1)$$

The first moment method, is essentially Markov's inequality for discrete random variables. More precisely, $X \geq 0$ integer valued then

$$\mathbb{P}(X > 0) = \mathbb{P}(X \geq 1) \underset{(1)}{\leq} E[X]$$

$$\Rightarrow \quad \boxed{\mathbb{P}(X > 0) \leq E[X] \quad - \text{1st Moment Method}} \quad - (2)$$

The second moment method controls the $\mathbb{P}(X > 0)$ by its first $(E[X])$ and Second $(E[X^2])$ moment. More precisely,

Let X be a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$ such that

$X \geq 0$, $0 < E[X] < \infty$ and $0 < E[X^2] < \infty$.

then

$$E[X^2] = E[X \cdot X \mathbb{1}_{X > 0}]$$

(C.B.S. inequality Proposition 8.2) $\leq \sqrt{E[X^2]} \sqrt{E[1_{X>0}]}$

$$\Rightarrow \quad \mathbb{P}(X>0) \geq \frac{(E[X])^2}{E[X^2]} \qquad \left(\text{Second moment method}\right)$$

These seemingly elementary inequalities are given their respective names because they are very useful in many applications

Example 1 :- (Coupon Collector Problem) Fix $n \geq 2$.
$Y_1, Y_2, \ldots$ be i.i.d. uniform $\{1, \ldots, n\}$. Let

$$T_n = \inf\{k \geq 1 \mid \{Y_1, \ldots, Y_k\} = \{1, \ldots, n\}\}$$

Coupon Collec' problem

- Think of $1, 2, \ldots, n$ as Coupons in a Tinku azad company chips packet.
- each packet has a randomly chosen Coupon placed in it
- T_n = # of chips packets you need to buy to collect all n coupons.

Has many variations

Question :- What is the order of T_n?

Answer(s) :- · Fix $m \geq 1$ - to denote # of days
· let $k \in \{1, \ldots, n\}$

and $$X_{m,k} = \begin{cases} 1 & \text{if } k^{th} \text{ coupon has not been picked until } m\text{th day} \\ 0 & \text{otherwise} \end{cases}$$

Note $\mathbb{P}(X_{m,k}=1) = \left(1-\frac{1}{n}\right)^m$ $\forall\, m \geq 1$, $1 \leq k \leq n$

$$\Rightarrow \quad E[X_{m,k}] = \left(1-\frac{1}{n}\right)^m \qquad \text{——— (1)}$$

Also $\mathbb{P}(X_{m,k}=1, X_{m,\ell}=1) = \left(1-\frac{2}{n}\right)^m$ $\forall\, m \geq 1$, $1 \leq k \neq \ell \leq n$

$$\Rightarrow \quad E[X_{m,k}\, X_{m,\ell}] = \left(1-\frac{2}{n}\right)^m \qquad \text{——— (2)}$$

$$\mathbb{P}(T_n > m) = \mathbb{P}(X_{m,1} + \cdots + X_{m,n} \geq 1)$$

1st moment method

$$\leq E[X_{m,1} + \cdots + X_{m,n}]$$

$$\overset{(1)}{=} n\left(1-\frac{1}{n}\right)^m$$

$1-x \leq e^{-x}$ $\forall\, x > 0$

$$\leq n\, e^{-\frac{m}{n}} \qquad \text{——— (3)}$$

If $m \cong n\log n$ then $n e^{-\frac{m}{n}} \cong 1$.

From this intuition take $m = n\log n + n\theta_n$ for some sequence of real numbers θ_n, $0 < \theta_n < \log n$.

Therefore,

$$\mathbb{P}(T_n > n\log n + n\theta_n)$$

$$\overset{(3)}{\leq} n\, e^{-\frac{\lfloor n\log n + n\theta_n \rfloor}{n}}$$

$x-1 < \lfloor x \rfloor \leq x$ $\forall\, x \in \mathbb{R}$

$$\leq e^{-\theta_n} \qquad \text{——— (4)}$$

Now on the other hand,

$$\mathbb{P}(T_n > m) = \mathbb{P}(X_{m,1} + \cdots + X_{m,n} \geq 1)$$

2nd moment method

$$\geq \frac{\left(E[X_{m,1} + \cdots + X_{m,n}]\right)^2}{E[(X_{m,1} + \cdots + X_{m,n})^2]}$$

$$\overset{(1)}{=} \frac{n^2 (1-\frac{1}{n})^{2m}}{\left(\sum_{k=1}^{n} E[X_{m,k}^2] + \sum_{\substack{k,l=1 \\ k \neq l}}^{n} E[X_{m,k} X_{m,l}]\right)}$$

$E[X_{m,k}^2] = E[X_{m,k}]$

(1) & (2)

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

$1-x \leq e^{-x}$

& $1-x \geq e^{-x-x^2}$

for $-\frac{1}{2} \leq x \leq \frac{1}{2}$

$$\geq \frac{n\, e^{-\frac{2m}{n} - \frac{2m}{n^2}}}{n\, e^{-\frac{m}{n}} + n(n-1)\, e^{-\frac{2m}{n}}}$$

$\therefore \mathbb{P}(T_n > n\log n - n\theta_n)$

Ex.

$$\geq \frac{e^{2\theta_n - \frac{2}{n}\log n + \frac{2\theta_n}{n}}}{e^{\theta_n} + (n(n-1)e^{-2\log n})\, e^{2\theta_n}}$$

$$\geq \frac{e^{\theta_n - \frac{2}{n}(\log n - \theta_n)}}{1 + e^{\theta_n}}$$

$$= \frac{e^{\theta_n(1+\frac{2}{n})}}{n^{2/n}(1+e^{\theta_n})} \quad \text{--- (5)}$$

$$\mathbb{P}(|T_n - n\log n| > n\theta_n)$$

$$= \mathbb{P}(T_n > n\log n + n\theta_n) + 1 - \mathbb{P}(T_n \geq n\log n - n\theta_n)$$

$$\overset{(4),(5)}{\leq} e^{-\theta_n} + 1 - \frac{e^{\theta_n(1+\frac{2}{n})}}{n^{2/n}(1+e^{\theta_n})} \qquad — (6)$$

$\therefore$ From (6)

if $0 \leq \theta_n \leq \log n$ and $\theta_n \to \infty$ as $n \to \infty$

then

$$\mathbb{P}(|T_n - n\log n| > n\theta_n) \longrightarrow 0 \quad \text{as} \quad n \to \infty. \qquad — (7)$$

$\Rightarrow$ T_n concentrates around $n\log n$ in a window of size n.

Corollary 1 :-

$$\frac{T_n}{n\log n} \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty$$

Proof :- Given $0 < \varepsilon < 1$, take $\theta_n = \varepsilon \log n$

$$\Rightarrow \mathbb{P}\left(\left|\frac{T_n}{n\log n} - 1\right| > \varepsilon\right) = \mathbb{P}(|T - n\log n| > n\theta_n).$$

Use (7) to complete the proof. $\square$

Example 2 : Bond - Percolation on $\mathbb{T}_2$

let $\mathbb{T}_2^0 = \{0\}$ & $\mathbb{T}_2^n = \{0,1\}^n \quad \forall n \geq 1$

(Vertices) $\mathbb{T}_2$ - binary tree $= \bigcup_{n=0}^{\infty} \mathbb{T}_2^n$

$\forall x \in \mathbb{T}_2^n$ let $a(x) = (x_1, x_2, .., x_{n-1}) \in \mathbb{T}_2^{n-1}$ be the ancestor of x in $\mathbb{T}_2$.

(Edge)

$$E := \{ \{x, a(x)\} \mid x \in \mathbb{T}_2 \setminus \mathbb{T}_2^0 \}$$

Fix $0 \leq p \leq 1$; declare each edge $e \in E$ to be

- open with probability p
- closed with probability $1-p$

in an i.i.d. manner

[each $e \in E$ is independently open or closed with probability p]

Setup:

- $\Omega = \{0, 1\}^E$ ("closed" ↙ 0, 1 ↘ "open")

$$\mu_e(\omega(e) = 1) = p \qquad \forall e \in E$$

$$\mu_e(\omega(e) = 0) = 1-p$$

- $\mathcal{A}$:= σ-field generated cylinder sets
- $\mathbb{P}_p = \bigotimes_{e \in E} \mu_e$

...

- **Percolation :-**

$E = \{$ there exists an infinite connected component of open edges $\}$

$$p_c = \inf \{ p \in [0,1] \mid \mathbb{P}_p(E) = 1 \}$$

E - Tail event & $\mathbb{P}_p(E) \in \{0, 1\}$

(Path)

$x = x_0 \to x_1 \quad \cdots \quad \to x_n = y$ is path in $\mathbb{T}^2$ between x and y if $\{x_i, x_{i+1}\} \in E$. $\forall\, i = 0, 1, \dots, n-1$.

Ex: Between $x, y \in \mathbb{T}^2$ $\exists$ a unique path in $\mathbb{T}^2$.

let $C_0 = \left\{ x \in \mathbb{T}^2 \;\middle|\; \text{Path between } x \rightsquigarrow 0 \text{ is open} \right\}$

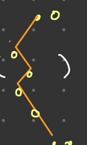
let $X_n = \sum_{x \in \mathbb{T}_2^n} \mathbb{1}(x \in C_0)$

(i.e X_n = # of vertices in $\mathbb{T}_2^n \cap C_0$)

$$\mathbb{P}(X_n > 0) \geq \frac{(\mathbb{E}[X_n])^2}{\mathbb{E}[X_n^2]}$$

(Second moment method)

$$= \frac{\left(\sum_{x \in \mathbb{T}_2^n} \mathbb{P}(x \in C_0) \right)^2}{\sum_{y \in \mathbb{T}_2^n} \sum_{x \in \mathbb{T}_2^n} \mathbb{P}(x \in C_0, y \in C_0)}$$

- $\sum_{x \in \mathbb{T}_2^n} \mathbb{P}(x \in C_0) = \sum_{x \in \mathbb{T}_2^n} \mathbb{P}($  $) = \sum_{x \in \mathbb{T}_2^n} p^n = |\mathbb{T}_2^n|\, p^n = 2^n p^n$

— (8)

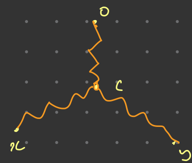
- $\forall\, x, y \in \mathbb{T}_2^n \quad n \geq 1$

$x \neq y$

$\exists$ a unique $c = c(x, y; 0) \in \mathbb{T}_2$ s.t $c \in \mathbb{T}_2^s$ $s \leq n-1$ & c lies on the path $x \rightsquigarrow 0$ and $y \overset{\text{path}}{\rightsquigarrow} x$.

- if $x = y$ $\quad c(x, y, 0) = x = y \in \mathbb{T}_2^n$

So

$$\mathbb{P}(x \in C_0, y \in C_0) = \mathbb{P}\Big(\ $$  $$\ \Big) = p^{2n-s}$$

$$\Rightarrow \sum_{y \in \Pi_2^n} \sum_{\substack{x \in \Pi_2^n \\ x \neq y}} \mathbb{P}(x \in C_0, y \in C_0)$$

$$= \sum_{s=0}^{n-1} \sum_{y \in \Pi_2^n} \sum_{x \in \Pi_2^n} \mathbb{P}(x \in C_0, y \in C_0, c \in \Pi_2^s)$$

$$= \sum_{s=0}^{n-1} \sum_{y \in \Pi_2^n} \sum_{x \in \Pi_2^n} p^{2n-s} \, \mathbb{1}(x \in \Pi_2^n, y \in \Pi_2^n, c \in \Pi_2^s)$$

$$= \sum_{s=0}^{n-1} p^{2n-s} \, 2^s \cdot 2^{n-s} \, 2^{n-s-1}$$

$$= \frac{(2p)^{2n}}{2} \sum_{s=0}^{n-1} (2p)^{-s}$$

if $p > \frac{1}{2}$

$$= \frac{1}{2} \cdot (2p)^{2n} \cdot \frac{1 - (1/2p)^n}{1 - (1/2p)} \leq (2p)^{2n} \frac{p}{2p-1} \quad -(9)$$

$$\Rightarrow \sum_{y \in \Pi_2^n} \sum_{\substack{x \in \Pi_2^n \\ x = y}} \mathbb{P}(x \in C_0, y \in C_0) = p^n \cdot 2^n \quad -(10)$$

if $p > \frac{1}{2}$

$$\therefore \mathbb{P}(X_n > 0) \underset{\substack{(9) \\ (10)}}{\overset{(8)}{\geq}} \frac{(2^n p^n)^2}{(2p)^{2n} \cdot \frac{p}{2p-1} + (2p)^n} = \frac{1}{\frac{p}{2p-1} + (2p)^{-n}}$$

$$\Rightarrow \quad \mathbb{P}(X_n > 0) \geq \frac{1}{\frac{p+(2p)^{-n}}{2p-1}} \quad \text{if } p \in (\tfrac{1}{2}, 1] \quad -(11)$$

$$\therefore \quad \mathbb{P}(|C_0| = \infty) \geq \mathbb{P}(\limsup_{n\to\infty} \{X_n > 0\})$$

$$\underset{\text{(Proposition 4.1)}}{\geq} \limsup_{n\to\infty} \mathbb{P}(\{X_n > 0\})$$

$$\underset{(11)}{\overset{\text{Ex.}}{\geq}} \frac{2p-1}{p} > 0$$

So the second moment method implies that the probability that the cluster size at o is infinite is positive $\forall\, p \in (\tfrac{1}{2}, 1]$ $\square$

14. Chernoff bounds

These bounds are a variation of Markov's inequality that have meaningful applications.

Let X be a r.v. on $(\Omega, \mathcal{F}, \mathbb{P})$. Let $I \equiv I(X)$ be an interval such that

Moment generating function $\left\{ \begin{array}{l} M_X : I \to \mathbb{R}_+ \quad \text{given by} \\ M_X(r) = E[e^{rX}] \quad \text{is well-defined.} \end{array} \right.$

Let, $r \in I$; $Y = e^{rX}$. The Markov inequality implies

$$\mathbb{P}(Y \geq a) \leq \frac{E[Y]}{a} \quad \forall\, a > 0$$

$$\Rightarrow \quad \mathbb{P}(e^{rX} \geq a) \leq \frac{M_X(r)}{a} \quad \forall a > 0, \ \forall r \in I$$

Now let $a = e^{rb}$ (say)

$$\left.\begin{array}{l} r>0,\ r \in I \Rightarrow \mathbb{P}(X \geq b) \leq e^{-rb} M_X(r) \\ r<0,\ r \in I \Rightarrow \mathbb{P}(X \leq b) \leq e^{-rb} M_X(r) \end{array}\right\} (C)$$

- Chernoff bounds -

These provide a family of upper bounds on the tails of the random variable X.

- In (C) for $r > 0$, $r \in I(X)$

$\mathbb{P}(X \geq b)$ - decays exponentially to 0 at rate e^{-rb}

(Note: not true for all X : require $E[e^{rX}] < \infty$)

Example 1 : let $X_1, X_2, \ldots, X_n, \ldots$ be i.i.d X

for $n \geq 1$ $\quad S_n = \sum_{i=1}^{n} X_i$

$r \in I(X)$ $\quad E[e^{rS_n}] = \prod_{i=1}^{n} E[e^{rX_i}] = (M_X(r))^n$

(C) with $b = na$ gives

$$\mathbb{P}(S_n \geq na) \leq (M_X(r))^n e^{-rna}$$

$$\mathbb{P}(S_n \leq na) \leq (M_X(r))^n e^{-rna}$$

let $\gamma_X(r) = \log(M_X(r))$, for $r \in I(X)$.

$$\boxed{\begin{aligned} &\mathbb{P}(S_n \geq na) \leq \exp(n(\gamma_X(r) - ra)) && r > 0,\ r \in I(X) \\ &\mathbb{P}(S_n \leq na) \leq \exp(n(\gamma_X(r) - ra)) && r < 0,\ r \in I(X) \end{aligned}}$$

— (0)

Lemma 1: Let $X_1, X_2, \ldots, X_n$ be i.i.d X. Assume $0 \in I(X)^\circ$. Then,

$$\mu_X(a) = \inf_{r \in I(X)} [\gamma_X(r) - ra] < 0, \ \forall\ a \neq E[X]$$

Also,

$$\mathbb{P}(S_n \geq na) \leq e^{n\mu_X(a)} \quad \text{for } a > E[X]$$

$$\mathbb{P}(S_n \leq na) \leq e^{n\mu_X(a)} \quad \text{for } a < E[X]$$

Proof:- let $I(X) = [r_-, r_+]$. Fix $a \in \mathbb{R}$.

$h: I \to \mathbb{R}$ be given by

$$h(r) = \gamma_X(r) - ra$$

$$h'(r) = \gamma_X'(r) - a$$

$\therefore$ h attains its minimum at

r_-, $r_0 : \gamma_X'(r_0) = a$ or r_+.

let $\mu_X(a) := \min(h(r_-), h(r_0), h(r_+))$

$$= \inf_r [\gamma_X(r) - ra]$$

Now: $$\gamma_X'(r) = \frac{d}{dr} \ln(E(e^{rX}))$$

(to be covered soon) $\doteq$ Ex $\frac{1}{E[e^{rX}]} E[X e^{rX}]$

$$\Rightarrow \boxed{h'(0) = E[X] - a} \quad -(1)$$

$$h''(r) = r_X''(r) = \frac{E[X^2 e^{rX}] - \left(E[X e^{rX}]\right)^2}{\left(E[e^{rX}]\right)^2}$$

Ex :- $r_X''(r) \geq 0$ ...

$$\Rightarrow \boxed{h''(r) \geq 0} \quad -(2)$$

Keys :-

- $a < E[X] \Rightarrow h'(0) > 0$; $h(0) = 0$

$h'(\cdot)$ (is continuous) $\Rightarrow \exists\ \delta > 0 \quad -\delta < r < 0 \Rightarrow h(r) < 0$

$h''(r) \geq 0 \Rightarrow \mu_X(a) < 0$ and

$$\mu_X(a) = \inf_{r<0} \{ r_X(r) - ra \}$$

- $a > E[X] \Rightarrow h'(0) < 0$; $h(0) = 0$

$h'(\cdot)$ (is continuous) $\Rightarrow \exists\ \eta > 0 \quad 0 < r < \eta \Rightarrow h(r) < 0$

$h''(r) \geq 0 \Rightarrow \mu_X(a) < 0$ and

$$\mu_X(a) = \inf_{r>0} \{ r_X(r) - ra \}$$

From (0) result follows $\square$

Example 1: $X_1, X_2, \dots, X_n, \dots$ iid $X \sim \text{Bernoulli}(p)$

$$M_X(r) = pe^r + 1-p \quad \forall\, r \in (-\infty, \infty)$$

$$h(r) = \gamma_X(r) - ra \qquad a \in \mathbb{R}$$

$$= \ln(pe^r + 1-p) - ra$$

$$h'(r) = 0 \iff \frac{pe^r}{pe^r + 1-p} = a$$

$$\iff r = \ln\left(\frac{a(1-p)}{p(1-a)}\right)$$

$$\Rightarrow \mu_X(a) = \ln\left(p\,\frac{a(1-p)}{p(1-a)} + (1-p)\right) - a\ln\left(\frac{a(1-p)}{p(1-a)}\right)$$

$$= (1-a)\ln\left(\frac{1-p}{1-a}\right) + a\ln\left(\frac{p}{a}\right)$$

$\therefore$

$$a > p \Rightarrow \mathbb{P}(S_n \geq na) \leq \exp\left(n\left[(1-a)\ln\left(\frac{1-p}{1-a}\right) + a\ln\left(\frac{p}{a}\right)\right]\right)$$

$$a < p \Rightarrow \mathbb{P}(S_n \leq na) \leq \exp\left(n\left[(1-a)\ln\left(\frac{1-p}{1-a}\right) + a\ln\left(\frac{p}{a}\right)\right]\right) \quad -(3)$$

(Markov) $$\mathbb{P}(S_n \geq na) \leq \frac{p}{a} \quad -(4)$$

(Tchebyschev) $$\mathbb{P}(S_n \geq na) = \mathbb{P}((S_n - np) \geq n(a-p))$$

$$\leq \mathbb{P}(|S_n - np| \geq n(a-p))$$

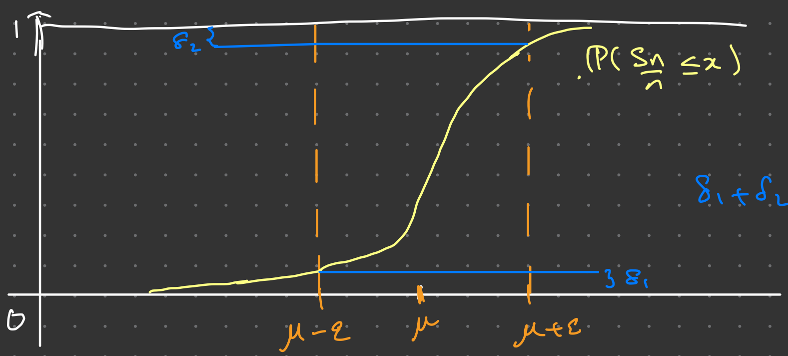
$$\leq \frac{np(1-p)}{n^2(a-p)^2} = \frac{p(1-p)}{n(a-p)^2}$$

Example 2 : $X_1, X_2, \ldots, X_n, \ldots$ be i.i.d X with

$E[X] = \mu < \infty$ and $M_X(t) < \infty$ in $(-\alpha, \alpha)$

$Var[X] = \sigma^2 < \infty$

$$\mathbb{P}\left(\left| \frac{S_n}{n} - \mu \right| > \varepsilon \right) \le \frac{\sigma^2}{\varepsilon^2 n}$$



$$\delta_1 + \delta_2 \le \frac{\sigma^2}{n\varepsilon^2}$$

$$\mathbb{P}\left(\frac{S_n}{n} \ge \mu + \varepsilon \right) \le \exp(-n\, \mu_X(\mu+\varepsilon))$$

$$\mathbb{P}\left(\frac{S_n}{n} \le \mu - \varepsilon \right) \le \exp(-n\, \mu_X(\mu-\varepsilon))$$

$$\Rightarrow \quad \mathbb{P}\left(\left| \frac{S_n}{n} - \mu \right| \ge \varepsilon \right) \le \underbrace{\exp(-n\, \mu_X(\mu+\varepsilon)) + \exp(-n\, \mu_X(\mu-\varepsilon))}_{\delta_1 + \delta_2}$$