Session - 4

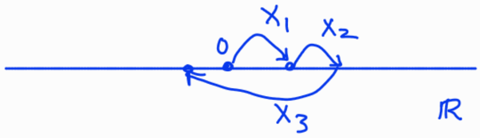
29/05/25

- X is a rv s.t. $\mathbb{E} e^{sX} < \infty \quad \forall s \in \mathbb{R}$.

  $\Rightarrow \mathbb{E} e^{sX} \leq e^{s^2 \nu/2} \quad \forall \ |s| \leq s_0$

  $\Rightarrow X$ is sub exponential.

- X_n iid X and $S_n = \sum_{i=1}^{n} X_i$.



$$\mathbb{P}(S_n - n\mu \geq c_n) \leq \exp\left(\frac{-kc_n^2}{n}\right)$$

- Typical value $\approx \sqrt{n}$

  $S_n = O_p(\sqrt{n})$

  $\Leftrightarrow \forall \ \varepsilon > 0, \ \exists \ C$ s.t.

  $\mathbb{P}(S_n \geq C\sqrt{n}) \leq \varepsilon$.

- $\boxed{\mathbb{P}\left(\frac{S_n}{n} - \mu \geq c\right) \leq_{=} \exp(-k_c n)}$ → Prove SLLN!

  Largest k_c possible

  [Exc: Use Borel Cantelli Lemma I]

- $\text{poly}(n) \exp(-k_c n) \leq \mathbb{P}\left(\frac{S_n}{n} - \mu \geq c\right) \leq \text{poly}(n) \exp(-k_c n)$

  Exact rate of decay. Large Deviation Event.

Cramer - Chernoff bound:

X_n iid X, $\mathbb{E} e^{sX} < \infty \quad \forall s \in \mathbb{R}$, $\mu = \mathbb{E}X$.

$\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq an\right)$

- Should decay exponentially for $a \geq \mathbb{E}X$
- $1 - e^{-n}$ if $a < \mathbb{E}X$

$$= \mathbb{P}\left(e^{\lambda \sum_{i=1}^{n} X_i} \geq e^{\lambda a n}\right) \quad \forall \ \lambda > 0$$

$$\leq e^{-\lambda a n} \mathbb{E} e^{\lambda \sum_{i=1}^{n} X_i} = e^{-\lambda a n} \left(\mathbb{E} e^{\lambda X}\right)^n, \quad \text{let } \psi_X(\lambda) = \log \mathbb{E} e^{\lambda X}$$

$$= \exp(-\lambda a n + n \psi_X(\lambda)) = \exp(-n(\lambda a - \psi_X(\lambda)))$$

$$\therefore \boxed{\mathbb{P}\left(\sum_{i=1}^{n} X_i \geq an\right) \leq \exp\left(-n \underbrace{\sup_{\lambda \geq 0} (\lambda a - \psi_X(\lambda))}\right)}$$

Observations: 1. $\psi_X(\lambda)$ — strictly convex, so long as $Var(X) > 0$.

Assume $\psi \in C^2$

$$\cdot \quad \psi_X''(\lambda) = \mathbb{E}\left[X^2 \frac{e^{\lambda X}}{\mathbb{E}e^{\lambda X}}\right] - \left(\mathbb{E}\left[X \frac{e^{\lambda X}}{\mathbb{E}e^{\lambda X}}\right]\right)^2$$

$$= Var_{\mu}(X) > 0$$

2. $\sup_{\lambda \in \mathbb{R}} (\lambda a - \psi(\lambda)) = \psi^*(a)$ ← Convex dual of ψ.

= Convex f$\underline{n}$.

Max. at $\psi'(\lambda^*) = a > \mathbb{E}X$

$\psi'(0) = \mathbb{E}X$

$$\psi'(\lambda) = \mathbb{E}\left[X \frac{e^{\lambda X}}{\mathbb{E}e^{\lambda X}}\right]$$

Since ψ is strictly convex, ψ' is strictly $\uparrow$.

$\therefore \lambda^* > 0$.

$$\therefore \quad \boxed{\mathbb{P}\left(\frac{S_n}{n} \geq a\right) \leq \exp\left(-n\psi^*(a)\right)} \quad a > \mathbb{E}X$$

— Chernoff Cramer bound

Right Tail Chernoff bd

egs. 1. $X \sim N(0,1)$, $\psi_X(\lambda) = \frac{\lambda^2}{2}$,

$$\psi_X^*(a) = \sup_{\lambda \in \mathbb{R}} \left\{\lambda a - \frac{\lambda^2}{2}\right\} = \frac{a^2}{2}$$

$\underline{\underline{\lambda = a}}$

$$\mathbb{P}\left(\frac{S_n}{n} \geq a\right) \leq \exp\left(-n\frac{a^2}{2}\right)$$

Left Tail:

$\mathbb{P}\left(\frac{S_n}{n} \leq a\right)$ for $a < \mathbb{E}X$

$$= \mathbb{P}\left(e^{\lambda S_n} \geq e^{\lambda a n}\right) \quad \text{for } \lambda < 0$$

$$= \exp\left(-n \sup_{\lambda < 0}\left(\lambda a - \psi(\lambda)\right)\right), \quad a < \mathbb{E}X$$

$$= \exp\left(-n \sup_{\lambda \in \mathbb{R}}\left(\lambda a - \psi(\lambda)\right)\right) = \exp\left(-n\psi_X^*(a)\right)$$

$$\boxed{\begin{aligned} \mathbb{P}(S_n \geq an) &\leq \exp(-n\psi^*(a)) \quad \text{if } a > \mathbb{E}X \\ \mathbb{P}(S_n \leq an) &\leq \exp(-n\psi^*(a)) \quad \text{if } a < \mathbb{E}X \end{aligned}}$$

Automatically gives us SLLN

0^{th} result of Large Deviation Theory.

$$\rightarrow \quad \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) = -\psi^*(a)$$

★ Thm (Cramer's theorem): $\mathbb{E}e^{sX} < \infty \;\; \forall\, |s| \leq s_0$, then

Large dev. event

LD rate fn.

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq an) = -\psi^*(a) \qquad \forall a > \mathbb{E}X$$

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \leq an) = -\psi^*(a) \qquad \forall a < \mathbb{E}X.$$

eg: 1. If $X \sim N(0,1)$,

$S_n = \sum_{i=1}^{n} X_i$, X_i iid $N(0,1)$

$$\left(\frac{1}{a} - \frac{1}{a^3}\right) e^{-\frac{a^2}{2}} \leq \mathbb{P}(X \geq a) \leq \frac{1}{a} e^{-\frac{a^2}{2}}$$

$$\mathbb{P}(S_n \geq na) = \mathbb{P}\left(\frac{S_n}{\sqrt{n}} \geq \sqrt{n}\, a\right), \quad \frac{S_n}{\sqrt{n}} \sim N(0,1)$$

$$\left(\frac{1}{\sqrt{n}\,a} - \frac{1}{(\sqrt{n}\,a)^3}\right) e^{-\frac{na^2}{2}} \leq \cdot \leq \frac{1}{\sqrt{n}\,a} e^{-\frac{na^2}{2}}$$

☆ Ex: Ber(p), Poisson(λ), Exp(λ), ...

For these, try to the Cramer-Chernoff bds.

☆ $\mathbb{P}\left(\frac{S_n}{n} \in [a,\infty)\right) \rightsquigarrow \mathbb{P}\left(\frac{S_n}{n} \in [a,b)\right)$

$$= \mathbb{P}\left(\frac{S_n}{n} \in [a,\infty)\right) - \mathbb{P}\left(\frac{S_n}{n} \in [b,\infty)\right)$$

$$\approx (\cdot)\, e^{-n\psi^*(a)} - (\cdot)\, e^{-n\psi^*(b)}$$

$\psi^*(a) < \psi^*(b) \longrightarrow$ Rate of decay $= -\psi^*(a)$

- $\frac{1}{n}\log\left(\mathbb{P}\left(\frac{S_n}{n} \in A\right)\right) = -\inf_{x \in A} \psi^*(x) \longrightarrow$ Large Deviation Principle.

  sufficiently regular

⊛ Concentration Inequalities for $X_1 + \cdots + X_n$:

1. Chebychev's Ineq: Variance bounds $\Longrightarrow$ Conc. ineq.

2. Exponential Conc. Ineq. → Hoeffding ⟶ Depends on sG norm
   → Bernstein ⟶ Depends on Variance

3. Cramer Chernoff method ⟶ More direct exp. conc. ✓

II. $\mathbb{P}\left(\left| f(X_1, X_2, \ldots) - \mathbb{E} f(X_1, \ldots)\right| \geq c_n\right) = o(1)$

eg: $f = \frac{X_1 + \cdots + X_n}{n}$ ✓, $\frac{X_1^2 + \cdots + X_n^2}{n}$ ✓, ... what others

"$f(X_1, \ldots, X_n)$ conc around its mean if

→ i) $X_1, \ldots, X_n$ are We~~akly~~ in dependent

ii) f dependends weakly on each of $X_1, \ldots, X_n$."

§. Martingales:

Def$^{\underline{n}}$: $X_0, X_1, X_2, \ldots$ is a collection of random variables. Then a collection $(M_t)_{t \in \mathbb{N} \cup \{0\}}$ is called a Martingale wrt $X_0, X_1, X_2, \ldots$ if

✓ i) $X_0, \ldots, X_t$ completely determines M_t ← Amount of money gained

✓ ii) $\mathbb{E}|M_t| < \infty$ and

iii) $\mathbb{E}[\underbrace{M_{t+1}}\,|\, X_0, X_1, \ldots, X_t] = \underbrace{M_t}$

eg. 1. $X_1, \ldots$ iid X, $\mathbb{E}X = \mu$,

Then $M_n = (X_1 + \cdots + X_n) - n\mu$

$$\mathbb{E}[M_{t+1} \mid X_1, \ldots, X_t] = \mathbb{E}[X_1 + \cdots + X_t + X_{t+1} - (t+1)\mu \mid X_1 + \cdots + X_t]$$

$$= (X_1 + \cdots + X_t) + \mathbb{E}[X_{t+1}] - (n+1)\mu$$

$$= M_t$$

$\therefore$ This a martingale.

- **Property:** $\mathbb{E}M_t = \mathbb{E}M_0$, t a fixed time.

Proof: $\mathbb{E}M_{t+1} = \mathbb{E}M_t = \cdots$ (use Tower law).

2. $M_n = S_n^2 - n\sigma^2$ if $\operatorname{Var} X = \sigma^2 < \infty$ $\longrightarrow$ Also defines a martingale.

Def$^{\underline{n}}$: 1. Supermartingale : $\mathbb{E}[M_{t+1} \mid \overbrace{X_0, \ldots, X_t}^{\mathcal{F}_t}] \leq M_t$

Submartingale : $\mathbb{E}[M_{t+1} \mid \mathcal{F}_t] \geq M_t$

Sup : $\mathbb{E}M_0 \leq \mathbb{E}M_1 \leq \cdots$

Sub : $\mathbb{E}M_0 \geq \cdots$

Def$^{\underline{n}}$: 1. Stopping time : A stopping time $X_0, X_1, X_2, \ldots$ is a random variable τ s.t. $\{\tau \leq t\}$ is completely determined by $X_0, \ldots, X_t$.

2. Predictable process: $(A_t)_{t\geq 0}$ s.t. A_t is compl. det. by $X_0, X_1, \dots, X_{t-1}$.

eg. Let $\overset{f(X_0, X_1, \dots)}{Y}$ be a rv depending upon $X_0, X_1, X_2, \dots$, assume $\mathbb{E}|Y| < \infty$.
Then, $M_t = \mathbb{E}[Y \mid X_0, \dots, X_t]$ is a martingale.

$$\mathbb{E}[M_t \mid X_0, \dots, X_{t-1}] = \mathbb{E}\Big[\overbrace{\mathbb{E}[Y \mid X_0, \dots, X_t]}\,\Big|\, X_0, \dots, X_{t-1}\Big]$$

$$\overset{\text{Tower Law}}{=} \mathbb{E}[Y \mid X_0, \dots, X_{t-1}] = M_{t-1}.$$

(Doob's Martingale OR Information Exposure Martingale)

★ **Property:** $M_0, M_1, \dots$ are martingales wrt $X_0, X_1, \dots$, then

- $\mathrm{Cov}(M_k - M_l, M_s - M_t) = 0$ if

so long as $[k, l] \cap [s, t] = \phi$. [Exercise]

Idea: $\left(-f(X_1, X_2, \dots, X_n) + \mathbb{E}f(X_1, X_2, \dots, X_n)\right)$

$$= \left(\mathbb{E}f(\underset{\sim}{X}) - \mathbb{E}[f(\underset{\sim}{X}) \mid X_1]\right) + \left(\mathbb{E}[f(\underset{\sim}{X}) \mid X_1] - \mathbb{E}[f(\underset{\sim}{X}) \mid X_1, X_2]\right)$$
$$+ \dots + \left(\mathbb{E}[f(\underset{\sim}{X}) \mid X_1, \dots, X_{n-1}] - f(\underset{\sim}{X})\right)$$

◎ $f - \mathbb{E}f$ = Sum of uncorrelated random variables.
(using Doob's Martingale).

- $\mathbb{E}e^{X_1 + X_2} = \mathbb{E}e^{X_1}\,\mathbb{E}e^{X_2}$

→ We would get around these difficulties using monotonicity of Submartingale.

Thm: (Optional Stopping Theorem):

- (M_t) is a submartingale. Then
$$\mathbb{E} M_t \geq \mathbb{E} M_0$$

Real q$^{\underline{n}}$: $\boxed{\mathbb{E} M_\tau \geq \mathbb{E} M_0 \ ? \quad \forall \text{ sub } M}$ → Requires some cond$^{\underline{n}}$s

i) τ is bounded ✓
ii) .
iii) .
iv) $M_t \geq 0$

→ Let (M_t) be a M, then M_t & $-M_t$ are both sub M
$$\mathbb{E}[-M_\tau] \geq \mathbb{E}[-M_0], \quad \mathbb{E} M_\tau \geq \mathbb{E} M_0$$
$$\Rightarrow \boxed{\mathbb{E} M_\tau = \mathbb{E} M_0}$$

☆

Week 1

Week 2

Week 3 -

Week 4 - 9/06, 12/06

[- 16/06 19/06
-] → No classes

-Weeks 5 - ☆30/06 . . . .
- -
. -

Week 8 -