Large Deviation Theory

The theory of large deviation deals with the proba. of rare events that're exponentially small as function of some parameter.

(1) For i.i.d seq.

Consider on $(\Omega, \mathcal{F}, \mathbb{P})$. $(X_n)_{n\geq 1}$. i.i.d.

s.t. $\forall \lambda \in \mathbb{R}^1$. $\mathbb{E}(e^{\lambda X_1}) < \infty$

Rmk: $\forall k \in \mathbb{R}^1$. $\mathbb{E}(|X_1|^k) < \infty$. Since we have $|x|^k \underset{k}{\lesssim} e^{-x} + e^{x}$.

By CLT: $\lim\limits_{n\to\infty} \mathbb{P}\left(\frac{\sum_1^n X_k}{n} \geq t\right) = (2\pi)^{-\frac{1}{2}} \int_t^\infty e^{-\frac{x^2}{2}} dx$.

Q: Does $\mathbb{P}\left(\frac{\sum_1^n X_k}{n} \geq t\right) \Big/ (2\pi)^{-\frac{1}{2}} \int_t^\infty e^{-\frac{x^2}{2}} \xrightarrow{t\to\infty} 1$?

A: It holds under some suitable conditions when $t \ll n^{\frac{1}{2}}$. But when $t \sim \sqrt{n}$ it won't hold (e.g. $|X_1| \leq C$. $t > Cn^{\frac{1}{2}}$)

Thm. For $A \in \mathcal{F}$. $\lim\limits_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\bar{X}_n \in A)$

$= -\inf\limits_{A} h(x)$. $h(x) = \sup\limits_{\theta} (\theta x - \psi(\theta))$.

$\psi(\theta) = \log M(\theta)$. $M(\theta) = \mathbb{E}(e^{\theta X_1})$

① Upper bound:

First consider $A = (l, \infty)$. $l > \mathbb{E}(X_1) \triangleq m$.

$$\mathbb{P}(\overline{X}_n \geq l) \overset{\text{cheby.}}{\leq} e^{-\theta n l}\, \mathbb{E}(e^{\theta S_n})$$

$$= e^{-\theta n l}\, M(\theta)^n \quad \text{. with } \theta \geq 0$$

$$\Rightarrow \frac{1}{n} \log \mathbb{P}(\overline{X}_n \geq l) \leq -\sup_{\theta \geq 0} (\theta l - \psi(\theta))$$

Rmk: For $\theta < 0$. By Jesen inequality:

$\psi(\theta) \geq m\theta \geq l\theta$. But 0 is trivial lower bdd $\Rightarrow$ Replace RHS by $-h(l)$.

About $h(l)$:

Def: i) Legendre-Fenchel transform $\mathcal{L}$ is defined on functions $\psi(x)$. by $\mathcal{L}\psi(l) =: \sup_{\theta} (l\theta - \psi(\theta))$.

ii) $\mathbb{P}_\lambda$ is p.m. on Ω. $\mathbb{P}_\lambda(A) =:$

$$\mathbb{E}(I_A e^{\lambda X_1}) / \mathbb{E}(e^{\lambda X_1}).$$

Prop. i) ψ is convex $\Rightarrow$ $\mathcal{L}\psi$ is convex.

ii) $\mathcal{L}^2 = id$. on covex functions.

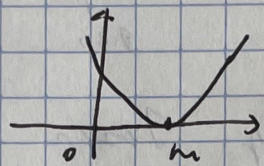
Note that $M(\theta)$ is convex:

$\psi'(\theta) = \mathbb{E}_\theta(X_1)$, $\psi''(\theta) = \mathbb{E}_\theta(X_1^2) - \mathbb{E}_\theta(X_1)^2 \geq 0$.

So $h(\lambda)$ is convex. $h(\lambda) \geq 0$.

Rmk: Note $h(m) = 0$, $h'(m) = 0$. m is the point where h attains its min

convex $\Rightarrow$



$h \uparrow$ on $\lambda \geq m$

$h \downarrow$ on $\lambda \leq m$.

By Rmk, we can also extend to set $A = (-\infty, a)$. For general set A, we can set $(-\infty, a] \cup [b, \infty)$ contains A, and $a, b \in \bar{A}$, with conti of h (by convex)

$$\Rightarrow \overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}(\bar{X}_n \in A) \leq \overline{\lim_{n\to\infty}} \frac{1}{n} \log \mathbb{P}(\bar{X}_n \leq a \text{ or } \geq b)$$

$$\leq -h(a) \wedge h(b)$$

$$\overset{\text{conti}}{=} -\inf_{(-\infty,a]\cup[b,\infty)} h(\lambda) \overset{\text{mono.}}{=} -\inf_A h$$

Rmk: We always assume A is closed for the upper bound.

② Lower bound:

Consider $\mu_\lambda^N(A) := \mathbb{E}\left(I_{(X_1 \ldots X_N) \in A}\, e^{\sum_1^N (\lambda X_i - \psi(\lambda))} \right)$

For $A = G$. open set of $\mathbb{R}^1$.

Set $J_\theta(\lambda) = \lambda\theta - \psi(\lambda)$. $J_\theta'(\lambda) = 0 \Rightarrow \theta = \psi'(\lambda)$

Assumption: $\lim\limits_{\lambda \to \infty} \psi'(\lambda) = +\infty$.

Rmk: By monotone. $\lim\limits_{\lambda \to \infty} \psi'(\lambda) = p$ exists, if $p < \infty$. then: for $\lambda > p$. $h(\lambda) = \infty$.

$$\Rightarrow h(\theta) = \theta (\psi')^{-1}(\theta) - \psi((\psi')^{-1}(\theta))$$

$$h'(\theta) = (\psi')^{-1}(\theta).$$

For $\forall \alpha \in G$. $B(\alpha, \delta)$ is nbd of α in G.

$$\mathbb{P}(\bar{X}_n \in G) \geq \mathbb{P}(\bar{X}_n \in B(\alpha, \delta))$$

$$\geq e^{-n(\lambda(\alpha+\delta) - \psi(\lambda))} M_\lambda^n (B(\alpha, \delta))$$

Choose $\hat{\lambda} = (\psi')^{-1}(\alpha)$. with CLT: we have.

$M_{\hat{\lambda}}^n (B(\alpha, \delta)) \xrightarrow{n \to \infty} 1$. by $\mathbb{E}_{\hat{\lambda}}(X_1) = \psi'(\hat{\lambda}) = \alpha$

$$\Rightarrow \frac{1}{n} \log \mathbb{P}(\bar{X}_n \in G) \geq -(\psi')^{-1}(\alpha)(\alpha+\delta) - \psi((\psi')^{-1}(\alpha))$$

$$\xrightarrow{\delta \to 0} -h(\alpha). \quad \forall \alpha \in G.$$

For general set A: consider $\bar{A}$ and $\operatorname{int} A$. with conti. of $h(\lambda)$.