Theorem: (Cramér's Large Deviation Theorem)

Let $\{X_n\}_{n\geq 1}^{\infty}$ be iid exponentially integrable random variables.
Let $\mathring{X}_n = X_n - \mathbb{E}[X_n]$, and $\mathring{S}_n = \mathring{X}_1 + \cdots + \mathring{X}_n$. For any $\varepsilon > 0$,

$$\boxed{\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\mathring{S}_n > n\varepsilon) = -\Psi^*_{\mathring{X}_1}(\varepsilon) \leq 0 .}$$

We have already proved that

$$\limsup_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\mathring{S}_n > n\varepsilon) \leq -\Psi^*_{\mathring{X}_1}(\varepsilon) .$$

Thus, it suffices to prove that

$$\liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\mathring{S}_n > n\varepsilon) \geq -\Psi^*_{\mathring{X}_1}(\varepsilon) .$$

For the sake of readability, set $\Psi(t) := \Psi_{\mathring{X}_1}(t)$ $\quad M(t) := M_{\mathring{X}_1}(t)$

$\Psi^*(\varepsilon) := \Psi^*_{\mathring{X}_1}(\varepsilon)$

Observations about Ψ^*:

- Since $\Psi = \log M$ is smooth, and convex,

$$\Psi^*(\eta) = \sup_{t \geq 0} \left(\eta t - \Psi(t) \right) = \eta t_\eta - \Psi(t_\eta)$$ for some maximizer $t_\eta > 0$.

$$0 = \frac{d}{dt}\left(\eta t - \Psi(t)\right)\Big|_{t = t_\eta} = \eta - \Psi'(t_\eta)$$

$$\Psi'(t_\eta) = \eta$$

Useful trick:

Given exponentially integrable Y (w.r.t. $(\Omega, \mathcal{F}, \mathbb{P})$),
for each t define a new probability measure $\mathbb{P}_t$ on $(\Omega, \mathcal{F})$:

$$d\mathbb{P}_t = \underset{\overset{\vee}{0}}{\frac{e^{tY}}{M_Y(t)}} \, d\mathbb{P}$$

$$\int_\Omega d\mathbb{P}_t = \int \frac{e^{tY}}{M(t)} \, d\mathbb{P} = \frac{1}{M(t)} \underbrace{\int e^{tY} \, d\mathbb{P}}_{\mathbb{E}[e^{tY}] = M(t)} = 1.$$

In the proof, we will use independent samples from $\mathbb{P}_t$ with $t = t_\eta$ (maximizer defining $\Psi^*_{X_1}(\eta)$)

Prop. For all $\varepsilon > 0$,

$$\liminf_{n\to\infty} \frac{1}{n}\log \mathbb{P}(\mathring{S}_n > n\varepsilon) \geq -\Psi^*(\varepsilon).$$

$$= \Psi(t_\varepsilon) - \varepsilon t_\varepsilon$$

Pf. Any maximizer t_ε of $t \mapsto \varepsilon t - \Psi(t)$ is a solution of

$$\Psi'(t) = \varepsilon.$$

$$\|$$

$$\frac{d}{dt}\log M(t) = \frac{M'(t)}{M(t)} = \frac{\mathbb{E}[\mathring{X}_1 e^{t\mathring{X}_1}]}{M(t)} = \int \mathring{X}_1 \frac{e^{t\mathring{X}_1}}{M(t)}\, d\mathbb{P}$$

$$\|$$

I.e. $t = t_\varepsilon$ is characterized by the requirement that $\mathbb{E}_{\mathbb{P}_t}[\mathring{X}_1] = \varepsilon$.

Let $\{W_n\}_{n=1}^{\infty}$ be i.i.d. random variables with

$$W_n \overset{d}{=} \mathring{X}_1 * \mathbb{P}_t = \mathrm{Law}_{\mathbb{P}_t}(\mathring{X}_1)$$

$$\mu_{W_n}(B) = \mathbb{P}_t(\mathring{X}_1 \in B) = \int_{\mathring{X}_1^{-1}(B)} \frac{e^{t\mathring{X}_1}}{M(t)}\, d\mathbb{P}$$

$$= \int_B \frac{e^{tx}}{M(t)} \mu_{\mathring{X}_1}(dx)$$

$$\boxed{\begin{aligned} &\mathbb{E}_{\mathbb{P}_t}[\mathring{X}_1] = \varepsilon \\ &\Rightarrow \quad \mathbb{E}_{\mathbb{P}}[W_n] = \varepsilon \end{aligned}}$$

$$\therefore \mu_{W_n}(dx) = \frac{e^{tx}}{M(t)} \mu_{\mathring{X}_1}(dx) \qquad \therefore \mu_{\mathring{X}_1}(dx) = M(t)\, e^{-tx} \mu_{W_n}(dx)$$

Then for measurable $f: \mathbb{R}^n \to [0, \infty]$,

$$\mathbb{E}[f(\mathring{X}_1, \dots, \mathring{X}_n)] = \int_{\mathbb{R}^n} f(x_1, \dots, x_n) \mu_{\mathring{X}_1}(dx_1) \cdots \mu_{\mathring{X}_1}(dx_n)$$

$$= \int_{\mathbb{R}^n} f(x_1, \dots, x_n)\, M(t)^n e^{-t(x_1 + \dots + x_n)} \mu_{W_1}(dx_1) \cdots \mu_{W_1}(dx_n)$$

$$= M(t)^n\, \mathbb{E}[f(W_1, \dots, W_n)\, e^{-t(W_1 + \dots + W_n)}]$$

Let $T_n = W_1 + \cdots + W_n$.

$\therefore$ We've shown that, for measurable $f : \mathbb{R}^n \to [0,\infty]$,

$$\mathbb{E}[f(\mathring{X}_1, \ldots, \mathring{X}_n)] = M(t)^n \mathbb{E}[f(W_1, \ldots, W_n) e^{-tT_n}]$$

Apply this with $f(x_1, \ldots, x_n) = \mathbb{1}_{x_1 + \cdots + x_n > n\varepsilon}$.

$$\mathbb{P}(\mathring{S}_n > n\varepsilon) = M(t)^n \mathbb{E}[e^{-tT_n} \underbrace{\mathbb{1}_{T_n \geq n\varepsilon}}_{\geq \mathbb{1}_{n\varepsilon \leq T_n \leq n(\varepsilon+\delta)}}]$$

$\delta > 0$

$$\geq M(t)^n e^{-tn(\varepsilon+\delta)} \mathbb{P}(n\varepsilon \leq T_n \leq n(\varepsilon+\delta))$$

$$= e^{n(\Phi(t) - t\varepsilon)} e^{-nt\delta} \mathbb{P}(n\varepsilon \leq T_n \leq n(\varepsilon+\delta))$$

$t = t_\varepsilon$

$$= e^{-n\Phi^*(\varepsilon)} e^{-nt\delta} \mathbb{P}(n\varepsilon \leq T_n \leq n(\varepsilon+\delta))$$

$$\mathbb{P}(\mathring{S}_n > n\varepsilon) \geq e^{-n\Psi^*(\varepsilon)} e^{-nt_\varepsilon \delta} \mathbb{P}(n\varepsilon \leq T_n \leq n(\varepsilon+\delta))$$

$$\therefore \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(\mathring{S}_n > n\varepsilon) \geq -\Psi^*(\varepsilon) - t_\varepsilon \delta$$
$$+ \liminf_{n\to\infty} \frac{1}{n} \log \mathbb{P}(n\varepsilon \leq T_n \leq n(\varepsilon+\delta)) \quad (\star)$$

for <u>any</u> $\delta > 0$.

To complete the proof, we want to let $\delta \downarrow 0$. This means we have to show that $(\star)$ is well-behaved as $n \to \infty$.

Note: $\mathbb{E}[W_n] = \varepsilon$, so $\mathbb{E}[T_n] = \mathbb{E}[W_1 + \cdots + W_n] = n\varepsilon$.

$$\therefore \mathbb{P}(n\varepsilon \leq T_n \leq n(\varepsilon+\delta)) = \mathbb{P}(0 \leq T_n - n\varepsilon \leq n\delta)$$
$$= \mathbb{P}(0 \leq \mathring{T}_n \leq n\delta) = \mathbb{P}\left(0 \leq \frac{\mathring{T}_n}{\sqrt{n \operatorname{Var} W_1}} \leq \sqrt{\frac{n}{\operatorname{Var} W_1}}\,\delta\right)$$

$$\sqrt{\frac{n}{\operatorname{Var} W_1}}\,\delta \searrow \infty$$

* Flash Forward *

CLT: $\mathbb{P}\left(\mathring{T}_n / \sqrt{n \operatorname{Var}(W_1)} \geq 0\right) \xrightarrow{n\to\infty} \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\,dx = \frac{1}{2}$

///