Rare Events

SLLN: $\{X_n\}_{n=1}^{\infty}$ i.i.d. L^1, $S_n = X_1 + \cdots + X_n$

$$\Rightarrow \quad \frac{S_n}{n} \to_{p} \mathbb{E}[X_1] \quad \cancel{a.s.}$$

If $X_n \in L^2$, Chebyshev shows

$$\mathbb{P}\left(\left| \frac{S_n}{n} - \mathbb{E}[X_1] \right| > \varepsilon \right) = O\left(\frac{1}{n}\right)$$

So $\left\{ \left| \frac{S_n}{n} - \mathbb{E}[X_1] \right| > \varepsilon \right\}$ is a rare event as $n \to \infty$.

$\forall t \neq 0, p > 0$

$e^{t|x|} > |x|^p$

To understand rare events without pathologies creeping in, assume good regularity for $\{X_n\}$.

Def: A random variable X is exponentially integrable $\subseteq L^p \quad \forall p \geq 1$ if $\mathbb{E}[e^{t|X|}] < \infty$ for all $t \in \mathbb{R}$.

Moments and Cumulants

If X is exponentially integrable, its **moment generating function**

$$M_X(t) = \mathbb{E}[e^{tX}]$$

[Lec 10.1]

is finite for all $t \in \mathbb{R}$; and analytic:

$$M_X(t) = \sum_{k=0}^{\infty} \mathbb{E}[X^k] \frac{t^k}{k!} = 1 + \mathbb{E}[X]t + \frac{1}{2}\mathbb{E}[X^2]t^2 + \cdots$$

$\therefore$ $\Psi_X(t) := \log M_X(t)$ is analytic for all $t \in \mathbb{R}$

Cumulant generating function

$$\Psi_X(t) = \sum_{k \geq 1} c_k(X) \frac{t^k}{k!}$$

cumulants

$c_k(X)$ is a deg k poly. in $\mathbb{E}[X^n]$: $n \leq k$.

$$c_k(X) = \Psi_X^{(k)}(0)$$

$$c_1(X) = \frac{M_X'(0)}{M_X(0)} = \frac{\mathbb{E}[X]}{1}$$

$$c_2(X) = \frac{M_X''(0) - M_X'(0)^2}{M_X(0)^2} = \frac{\mathbb{E}[X^2] - \mathbb{E}[X]^2}{1} = \mathrm{Var}[X]$$

How rare is it for $\left|\frac{S_n}{n} - \mathbb{E}[X_1]\right| > \varepsilon$?

Let's assume X_n is exponentially integrable.

We'll also center $\mathring{X}_n = X_n - \mathbb{E}[X_n]$, $\mathring{S}_n = \sum_{k=1}^{n} \mathring{X}_k = S_n - n\mathbb{E}[X_1]$

$$
\begin{aligned}
\mathbb{P}\left(\left|\frac{S_n}{n} - \mathbb{E}[X_1]\right| > \varepsilon\right) &= \mathbb{P}(|\mathring{S}_n| > n\varepsilon) \\
&\leq \mathbb{P}(\mathring{S}_n > n\varepsilon) \\
&= \mathbb{P}(e^{t\mathring{S}_n} > e^{tn\varepsilon}) \qquad \forall t > 0 \\
\text{Markov} \rightarrow &\leq e^{-tn\varepsilon}\, \mathbb{E}[e^{t\mathring{S}_n}]
\end{aligned}
$$

$$
\left.\begin{aligned}
&\mathring{S}_n = \mathring{X}_1 + \cdots + \mathring{X}_n \ \text{indep} \\
&\therefore e^{t\mathring{S}_n} = e^{t\mathring{X}_1} \cdots e^{t\mathring{X}_n} \ \text{indep.}
\end{aligned}\right\}
\qquad \overset{\shortparallel}{e^{-tn\varepsilon}\, \mathbb{E}[e^{t\mathring{X}_1}] \cdots \mathbb{E}[e^{t\mathring{X}_n}]} = e^{-tn\varepsilon} M_{\mathring{X}_1}(t)^n
$$

$$
\begin{aligned}
\therefore \mathbb{P}(\mathring{S}_n > n\varepsilon) &\leq \left(e^{-t\varepsilon} M_{\mathring{X}_1}(t)\right)^n = \left(e^{-t\varepsilon} e^{\Psi_{\mathring{X}_1}(t)}\right)^n \\
&= e^{-n(t\varepsilon - \Psi_{\mathring{X}_1}(t))}
\end{aligned}
$$

for any $t > 0$.

Legendre Transform

Let $\psi : \mathbb{R} \to \mathbb{R}$ be a convex function.

The Legendre transform $\psi^* : \mathbb{R} \to \mathbb{R}$ is

$$\psi^*(\eta) := \sup_{t \in \mathbb{R}} (\eta t - \psi(t))$$

Eg. $\psi(t) = \frac{1}{2}\sigma^2 t^2 = \Phi_Z(t) \qquad Z \stackrel{d}{=} N(0, \sigma^2)$

$$\sup_{t \in \mathbb{R}} (\eta t - \psi(t)) = \max_{t \in \mathbb{R}} (\eta t - \tfrac{1}{2}\sigma^2 t^2) = \frac{1}{2\sigma^2}\eta^2 = \psi^*(\eta)$$

@ $t = \eta / \sigma^2$

Prop: ψ^* is a convex function, and $\psi^{**} = \psi$.

If $\psi = \Phi_{\mathring{X}}$ for some exponentially integrable X, then

$$\psi^*(\eta) = \sup_{t \geq 0} (\eta t - \psi(t)) \geq 0 \quad \text{for } \eta \geq 0.$$

Pf. [Re: $\Phi_{\mathring{X}}$]

$$\Phi_{\mathring{X}}''(t) = \frac{M_{\mathring{X}}''(t) M_{\mathring{X}}(t) - M_{\mathring{X}}'(t)^2}{M_{\mathring{X}}(t)^2} = \frac{\mathbb{E}[\mathring{X}^2 e^{t\mathring{X}}]}{M_{\mathring{X}}(t)} - \left(\frac{\mathbb{E}[\mathring{X} e^{t\mathring{X}}]}{M_X(t)}\right)^2$$

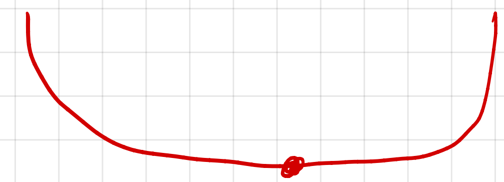
Let $d\mathbb{P}_t = \frac{e^{t\mathring{X}}}{M_{\mathring{X}}(t)} d\mathbb{P}$

$$= \int \mathring{X}^2 d\mathbb{P}_t - \left(\int \mathring{X} d\mathbb{P}_t\right)^2$$

$$= Var_{\mathbb{P}_t}(\mathring{X}) \geq 0.$$

$\Psi_{\mathring{X}}$ is convex for any exponentially integrable $\mathring{X}$.

Now for centered random variables,



$$\Psi'_{\mathring{X}}(0) = \mathbb{E}[\mathring{X}] = 0,$$

$$\Psi_{\mathring{X}}(0) = 0.$$

$\therefore\ \eta t - \Psi_{\mathring{X}}(t) \le 0$ for $\eta \ge 0$, $t \le 0$. But @ $t = 0$, $\eta \cdot 0 - \Psi_{\mathring{X}}(0) = 0$.

$$\therefore\ \Psi^*_{\mathring{X}}(\eta) = \sup_{t \ge 0} \left(\eta t - \Psi_{\mathring{X}}(t)\right) \ge 0 \text{ for } \eta \ge 0.$$

///

Thus, our SLLN calculation shows that

$$\mathbb{P}(\mathring{S}_n > n\varepsilon) \le e^{-n(t\varepsilon - \Psi_{\mathring{X}_1}(t))} \quad \forall t \in \mathbb{R}$$

$$\Rightarrow \quad \le e^{-n \sup_{t \in \mathbb{R}} (\varepsilon t - \Psi_{\mathring{X}_1}(t))}$$

$$= e^{-n \sup_{t \ge 0} (\varepsilon t - \Psi_{\mathring{X}_1}(t))} = e^{-n \Psi^*_{\mathring{X}_1}(\varepsilon)}$$

$\ne 0$, > 0 $\forall$ large $\varepsilon > 0$.

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

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 $\varepsilon > 0$,

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

That's exactly how rare the event $\{\frac{S_n}{n} - \mathbb{E}[X_1] > \varepsilon\} = \{\mathring{S}_n > n\varepsilon\}$ is:
it is exponentially small in n, with exponential rate $-\Psi^*_{\mathring{X}_1}(\varepsilon)$.

Being so rare, Cramér referred to such events as large deviations.
In modern parlance, we would say

The family of probability measures $\{\mu_{\mathring{S}_n/n}\}_{n=1}^{\infty}$
satisfies a Large Deviations Principle (LDP)
with rate function $\Psi^*_{\mathring{X}_1}$ and (exponential) rate n.