# Rare Events

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

$$\Rightarrow \quad \frac{S_n}{n} \to \mathbb{E}[X_1] \quad \text{a.s.}$$

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 $n \to \infty$.

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

Def: A random variable $X$ is exponentially integrable 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**

[Lec 10.1]

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

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

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

**Cumulant generating function**

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

**cumulants**

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

$$c_1(X) = \frac{M_X'(0)}{M_X(0)}$$

$$c_2(X) = \frac{M_X''(0) - M_X'(0)^2}{M_X(0)^2}$$

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 $\mathring{X}_n = X_n - \mathbb{E}[X_n]$, $\mathring{S}_n = \sum_{k=1}^{n} \mathring{X}_k$

$$\mathbb{P}\left( \left| \frac{S_n}{n} - \mathbb{E}[X_1] \right| > \varepsilon \right) = \mathbb{P}( |\mathring{S}_n| > n\varepsilon)$$
$$\leq \mathbb{P}( \mathring{S}_n > n\varepsilon)$$

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

for any $t > 0$.

# Legendre Transform

Let $\psi : I \to \mathbb{R}$ 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. $\psi(t) = \frac{1}{2}\sigma^2 t^2$

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

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

If $\psi = \Phi_X^\circ$ 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: $\Phi_X^\circ$ ]

$\Psi_X$ is convex for any exponentially integrable $X$.

Now for centered random variables,

$\Psi_{\mathring{X}}'(0) = \mathbb{E}[\mathring{X}] = 0.$

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

$\therefore\ \Psi_{\mathring{X}}^*(\eta) = \sup_{t \geq 0} (\eta t - \Psi_{\mathring{X}}(t)) \geq 0$ for $\eta \geq 0$. ///

Thus, our SLLN calculation shows that

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

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

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

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

**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 $\left\{ \frac{S_n}{n} - \mathbb{E}[X_1] > \varepsilon \right\} = \{ \mathring{S}_n > n\varepsilon \}$ 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 $n$.