Large Deviations

Probability Theory

Carlos Martins (63404)

Rafael Cabral (81292)

Instituto Superior Técnico

21 December 2017

Table of Contents

Weak Law of Large Numbers

Let $X_1, \ X_2, \ \ldots$ be i.i.d r.v with mean $\mu < \infty$ and let $S_n = X_1 + \cdots + X_n$ then for every $\epsilon > 0$

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) \to 0,$$

as $n \to \infty$.

Question

How quickly does this convergence to zero occur?

We can try to use Chebyshev inequality which states

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \mu\right| > \epsilon\right) \leq \frac{\mathbb{V}\text{ar}[X_1]}{n\epsilon^2}.$$

This suggest a “decay rate” of order $\frac{1}{n}$ if we treat Var (X_1) and ϵ as a constant.

Goal

The goal of the *large* deviation theory is to show that in many interesting cases the decay rate is in fact *exponential*: $e^{n\gamma(a)}$. The exponent $\gamma(a) < 0$ and -$\gamma(a)$ is called the *large deviations rate function*, and in many cases it can be computed explicitly or numerically.

Table of Contents

Let $X_1,\ X_2,\ \dots$ be i.i.d and let $S_n = X_1 + \cdots + X_n$. Fix a value $a > \mu$, where $\mu = \mathbb{E}[X]$. We consider now probability that the average of $X_1, \dots, X_n$ exceeds a.

Chernoff Upper Bound

Fix a positive parameter $\theta > 0$. We have

$$\mathbb{P}\left(\frac{S_n}{n} \geq a\right) \leq \frac{(\mathbb{E}[e^{\theta X_1}])^n}{e^{\theta n a}} \quad (*)$$

Thus we obtain an upper bound:

$$\mathbb{P}\left(\frac{\sum_{1\leq i\leq n} X_i}{n} \geq a\right) \leq \left(\frac{\mathbb{E}[e^{\theta X_1}]}{e^{\theta a}}\right)^n$$

• The bound is useful only if the ratio $\mathbb{E}[e^{\theta X_1}]/e^{\theta a}$ is less than unity.
• $\mathbb{E}[e^{\theta X_1}] = M(\theta)$ is moment generating function of X_1
• We need $\mathbb{E}[e^{\theta X_1}]$ to be at least finite.
• If we could show that this ratio is less than unity, we would be done - exponentially fast decay of the probability would be established.

Similarly, suppose we want to estimate

$$\mathbb{P}\left(\frac{S_n}{n} \leq a\right) ,$$

for some $a < \mu$. Fixing now a negative $\theta < 0$, we obtain

$$\mathbb{P}\left(\frac{S_n}{n} \leq a\right) \leq \left(\frac{M(\theta)}{e^{\theta a}}\right)^n (*)$$

Now we need to find a negative θ such that $M(\theta) < e^{\theta a}$. In particular, we need to focus on θ for which the moment generating function is finite.

Table of Contents

We've seen previously that the Chernoff Upper Bound suggests an exponential "decay rate" of $\mathbb{P}(S_n \geq na)$

Rate function

Let $X_1, X_2, \ldots$ be i.i.d and let $S_n = X_1 + \cdots + X_n$. We want to conclude that if the moment-generating function $M(\theta) < \infty$ for some $\theta > 0$, $\mathbb{P}(S_n \geq na) \to 0$ exponentially rapidly and we will identify:

$$\gamma(a) = \lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq na)$$

Our first step is to prove that the limit exists. -$\gamma(a)$ is also known as the rate function.

We will use an observation that will be useful several times below: Let $\pi_n = \mathbb{P}(S_n \geq na)$. Then

$$\pi_{m+n} \geq \pi_m \pi_n \quad (*)$$

Letting $\gamma_n = \log \pi_n$ we have that $\gamma_{m+n} \geq \gamma_m + \gamma_n$

Lemma 1

If $\gamma_{m+n} \geq \gamma_m + \gamma_n$ *then as* $n \to \infty$, $\gamma_n/n \to \sup_m \gamma_m/m$. (*)

- Lemma 1 implies that:

$$\lim_{n\to\infty} \frac{1}{n} \log \mathbb{P}(S_n \geq na) = \lim_{n\to\infty} \frac{\gamma_n}{n} = \gamma(a) \text{ exists and it's } \leq 0.$$

- It follows that:

$$\mathbb{P}(S_n \geq na) \leq e^{n\gamma(a)}$$

Exercise 1

The following are equivalent:

(a) $\gamma(a) = -\infty$,

(b) $\mathbb{P}(X_1 \geq a) = 0$

(c) $\mathbb{P}(S_n \geq na) = 0$ for all n.

Hypothesis H1

$M(\theta) = \mathbb{E}[\exp(\theta X_i)] < \infty$ for some $\theta > 0$

Chernoff Upper Bound

- Let $\theta_+ = \sup\{\theta : (\theta) < \infty\}$, $\theta_- = \inf\{\theta : (\theta) < \infty\}$ and note that $M(\theta) < \infty$ for $\theta \in (\theta_-, \ \theta_+)$.
- (H1) implies that $\mu = \mathbb{E}[X] \in (-\infty, \ \infty)$.

We've seen previously that:

$$\mathbb{P}(S_n \geq na) \leq \left(\frac{M(\theta)}{e^{\theta a}}\right)^n = \exp(-n\{a\theta - \kappa(\theta)\})$$

letting $\kappa(\theta) = \log M(\theta)$

The previous inequality is useful when:

$$\left(\frac{M(\theta)}{e^{\theta a}}\right)^n < 1 \Leftrightarrow a\theta - \kappa(\theta) > 0$$

Lemma 2

If $a > \mu$ and $\theta > 0$ is small then $a\theta - \kappa(\theta) > 0$. (*)

Optimizing Upper Bound

Having found an upper bound on $\mathbb{P}(S_n \geq na)$, it is natural to optimize it by finding the maximum of $\theta a - \kappa(\theta)$:

$$\frac{d}{d\theta}\{\theta a - \log M(\theta)\} = a - M'(\theta)/M(\theta)$$

so (assuming things are nice) the maximum occurs when $a = M'(\theta)/M(\theta)$.

To be more precise, let's define the distribution function:

$$F_\theta(x) = \frac{1}{\varphi(\theta)} \int_{-\infty}^{x} e^{\theta y} dF(y)$$

whenever $M(\theta) < \infty$.

We've seen in the proof of Lemma 2 that for $\theta \in (0,\ \theta_+)$:

$$M'(\theta) = \int x e^{\theta x} dF(x)$$

Mean of F_θ

So if $\theta \in (\theta_-,\ \theta_+)$, F_θ is a distribution function with mean

$$\int x dF_\theta(x) = \frac{M'(\theta)}{M(\theta)} (*)$$

Repeating the proof in Lemma 2, it is easy to see that if $\theta \in (\theta_-,\ \theta_+)$ then

$$M''(\theta) = \int_{-\infty}^{\infty} x^2 e^{\theta x} dF(x)$$

So we have

$$\frac{d}{d\theta}\frac{M'(\theta)}{M(\theta)} = \frac{M''(\theta)}{M(\theta)} - \left(\frac{M'(\theta)}{M(\theta)}\right)^2 = \int x^2 dF_\theta(x) - \left(\int x dF_\theta(x)\right)^2 \geq 0$$

since the last expression is the variance of F_θ.

Hypothesis H2

The distribution F is not a point mass at μ ($\implies Var(X_i) \neq 0$)

We know that:
• $M'(\theta)/M(\theta)$ is strictly increasing and $a\theta - \log M(\theta)$ is concave. (*)
• $M'(0)/M(0) = \mu$ (*)
This shows that for each $a > \mu$ there is at most one $\theta_a \geq 0$ that solves $a = M'(\theta_a)/M(\theta_a)$, and this value of θ maximizes $a\theta - \log M(\theta)$.

Before discussing the existance of θ_a we will consider the following:

Example

$M(\theta)$ and $M'(\theta)/M(\theta)$ for:

1. Normal Distribution
   $M(\theta) = \exp\big(\theta^2/2\big)$ e $M'(\theta)/M(\theta) = \theta$
2. Exponential distribution with parameter $\lambda > \theta$
   $M(\theta) = \lambda/(\lambda - \theta)$ e $M'(\theta)/M(\theta) = 1/(\lambda - \theta)$
3. Coin flips
   $M(\theta) = (e^{\theta} + e^{-\theta})/2$ e
   $M'(\theta)/M(\theta) = (e^{\theta} - e^{-\theta})/(e^{\theta} + e^{-\theta})$
4. Perverted Exponential
   $M'(\theta)/M(\theta) < 2$

Observation 1

- Recall $\theta_+ = \sup\{\theta : \varphi(\theta) < \infty\}$
- In Example 1 (Normal Distribution) $\theta_+ = \infty$ and $M'(\theta)/M(\theta) \uparrow \infty$ as $\theta \uparrow \theta_+$
- In Example 2 (Exponential distribution) $\theta_+ = \lambda$ and $M'(\theta)/M(\theta) \uparrow \infty$ as $\theta \uparrow \theta_+$

So we can find θ that solves $a = M'(\theta)/M(\theta)$ for any $a > \mu$.

Observation 2

In Example 3 (Coin Flips) $M'(\theta)/M(\theta) \uparrow 1$ as $\theta \to \infty$, but in this case F and hence F_θ only have support on $\{-1, 1\}$.

Observation 3

Example 4 (Perverted Exponential) presents a problem since we cannot solve $a = M'(\theta)/M(\theta)$ when $a > 2$.
Theorem 2 will cover this problem case, but first we will treat the cases in which we can solve the equation:

Exercise 2

Let $x_o = \sup\{x : F(x) < 1\}$. Show that if $x_o < \infty$ then:

- $M(\theta) < \infty$ for all $\theta > 0$ (*)
- $M'(\theta)/M(\theta) \to x_o$ as $\theta \uparrow \infty$. (*)

Theorem 1

Suppose in addition to $(H1)$ *and* $(H2)$ *that there is a* $\theta_a \in (0,\ \theta_+)$ *so that* $a = M'(\theta_a)/M(\theta_a)$. *Then, as* $n \to \infty$,

$$n^{-1} \log \mathbb{P}(S_n \geq na) \to -a\theta_a + \log\ M(\theta_a)(*)$$

In the proof of this theorem we will use the following lemma:

Lemma 3

$$\frac{dF^n}{dF_\lambda^n} = e^{-\lambda x} M(\lambda)^n$$

.

Reminder

To get a feel for what the answers look like, we consider our examples. Important information:

• $\kappa(\theta) = \log M(\theta)$ • $\kappa'(\theta) = M'(\theta)/M(\theta)$ • θ_a solves $\kappa'(\theta_a) = a$
• $\gamma(a) = \lim_{n\to\infty}(1/n)\log \mathbb{P}(S_n \geq na) \; = -a\theta_a + \kappa(\theta_a)$

Example

$\gamma(a)$ for:

1. Normal Distribution
$$\gamma(a) = -a^2/2$$
2. Exponential distribution with parameter $\lambda > \theta$
$$\gamma(a) - a + 1 + \log a$$
3. Coin flips
$$\gamma(a) = -\{(1+a)\log(1+a) + (1-a)\log(1-a)\}/2$$

Turning now to the problematic values for which we cannot solve $a = M'(\theta_a)/M(\theta_a)$:

• We begin by observing that if $x_o = \sup\{x : F(x) < 1\}$ and F is not a point mass at x_o then:
$M'(\theta)/M(\theta) \uparrow x_0$ as $\theta \uparrow \infty$ but $M'(\theta)/M(\theta) < x_0$ for all $\theta < \infty$.

Exercise 3

• However, the result for $a = x_o$ is trivial:

$$\frac{1}{n} \log \mathbb{P}(S_n \geq n x_o) = \log \mathbb{P}(X_i = x_o) \text{ for all } n \text{ (*)}$$

Exercise 4

Show that as $a \uparrow x_o,\ \gamma(a) \downarrow \log \mathbb{P}(X_i = x_o)$. (*)

When $x_o = \infty,\ M'(\theta)/M(\theta) \uparrow \infty$ as $\theta \uparrow \infty$, so the only case that remains is covered by

Theorem 2

Suppose $x_o = \infty,\ \theta + < \infty$, *and* $\varphi'(\theta)/\varphi(\theta)$ *increases to a finite limit* a_0 *as* $\theta \uparrow \theta_+ \cdot$ *If* $a_0 \le a < \infty$

$$n^{-1} \log \mathbb{P}(S_n \ge na) \to -a\theta_+ + \log M(\theta_+) \quad (*)$$

i. e., $\gamma(a)$ *is linear for* $a \ge a_0$.