Week 8

Large Deviations for Random Walks

LECTURER: SIVA ATHREYA SCRIBE: VENKAT TRIVIKRAM, SRIVATSA B

Nash Inequality (continued)

Theorem 8.0.1. *Let (Γ, μ) be a weighted graph, and let $\alpha \geq 1$. TFAE.*

(a) (Nash Inequality) (Γ, μ) satisfies (N_α)

(b) (On Diagonal Bounds) There exists $C_H > 0$ such that for every $x \in V$ and $n \geq 0$

$$p_n(x,x) \leq \frac{C_H}{(n \vee 1)^{\alpha/2}}$$

(c) (Off Diagonal Bounds) There exists $C'_H > 0$ such that for every $x, y \in V$ and $n \geq 0$

$$p_n(x,y) \leq \frac{C'_H}{(n \vee 1)^{\alpha/2}}$$

Proof. We provide only a sketch of the proof. From Worksheet 2, $(a) \implies (b)$ holds, and $(c) \implies (b)$ is trivial. First, we show $(b) \implies (c)$. So assume (b). Let $m \geq 0$. If n is even, with $n = 2m$, then for any $x, y \in V$, we have

$$p_{2m}(x,x) \leq \frac{C_H}{(2m \vee 1)^{\alpha/2}} \quad \text{and} \quad p_{2m}(y,y) \leq \frac{C_H}{(2m \vee 1)^{\alpha/2}}$$

As an exercise, show that $p_{2m}(x,y) \leq \sqrt{p_{2m}(x,x)p_{2m}(y,y)}$, and using this, we get $(b) \implies (c)$ with $C'_H = C_H$. If $n = 2m+1$, then, since $p_{2m+1}(x,y) \leq \sqrt{p_{2m}(x,x)p_{2m+2}(y,y)}$ (by a similar exercise), we get

$$p_{2m+1}(x,y) \leq \sqrt{\frac{C_H^2}{(2m \vee 1)^{\alpha/2}(2m+2 \vee 1)^{\alpha/2}}} \leq \frac{C'}{(2m+1 \vee 1)^{\alpha/2}}$$

for some $C' > 0$. To show the last inequality above, use the fact that there exists $C_\alpha > 0$ such that $(2m)^{\alpha/2}(2m+2)^{\alpha/2} \leq C_\alpha (2m+1)^{\alpha/2}$ (details left as exercises). Thus $(b) \implies (c)$.
Now, we show $(c) \implies (a)$. Assuming (c), observe that (by taking supremum over $x \in V$)

$$|P_n f(x)| \leq \sum_{y \in V} p_n(x,y)\,|f(y)|\,\mu_y \implies \|P_n f\|_\infty \leq \frac{C_H}{(n \vee 1)^{\alpha/2}} \|f\|_1$$

$$\text{and} \quad \|P_n f\|_2^2 = \langle P_n f,\ P_n f\rangle = \langle P_{2n} f,\ f\rangle \le \|P_{2n} f\|_\infty \|f\|_1 \le \frac{C_H}{(2n \vee 1)^{\alpha/2}} \|f\|_1^2 \tag{8.1}$$

Now, we make use of the following inequality - (verify!)

$$\mathcal{E}(f,f) \ge \frac{1}{2n}[\|f\|_2^2 - \|P_n f\|_2^2]$$

Using this, and (8.1), we get

$$\mathcal{E}(f,f) \ge \frac{1}{2n}\left[\|f\|_2^2 - \frac{C_H}{(2n \vee 1^{\alpha/2})}\|f\|_1^2\right]$$

WLOG, assume $\|f\|_1 = 1$, and choose smallest possible k such that

$$\frac{C_H}{(2n \vee 1)^{\alpha/2}} \le \frac{\|f\|_2^2}{2} \quad \text{so that} \quad \mathcal{E}(f,f) \ge \frac{1}{4k}\|f\|_2^2$$

Since $k \ge 1$, we have $k^{-\alpha/2} \le C^2 \|f\|_2^2$ for some $C > 0$, and hence $k^{-\alpha/2} \le C\|f\|_2$. Therefore,

$$\mathcal{E}(f,f) \ge \frac{C_2 \|f\|_2^2}{\|f\|_2^{\frac{4}{\alpha}}} = C_2 \|f\|_2^{2-4/\alpha} \implies (N_\alpha)$$

□

Carne-Varopoulos Bound

We begin with a few lemmas and some results involving Chebyshev polynomials.

Lemma 8.0.1. *Let $\{S_n\}_{n\ge 0}$ denote the simple symmetric random walk on $\mathbb{Z}$ with $S_0 = 0$. Then*

(a)

$$\mathbf{P}(S_n \ge D) \le \exp\left(-\frac{D^2}{2n}\right)$$

(b)

$$\mathbf{E}[\lambda^{S_n}] = \sum_{r\in\mathbb{Z}} \lambda^r \mathbf{P}(S_n = r) = 2^{-n}\sum_{r=0}^{n}\binom{n}{r}\left(\frac{1}{\lambda}\right)^{2n-r}$$

Proof. (a) was given in Worksheet 2, and (b) is trivial using results from Week 1. □

Definition 8.0.1. *(Chebyshev Polynomials) For $-1 \le t \le 1$, define*

$$H_k(t) := \frac{1}{2}(t + i\sqrt{1-t^2})^k + \frac{1}{2}(t - i\sqrt{1-t^2})^k$$

Lemma 8.0.2. *For each $k \ge 0$, we have*

(a) H_k is a real polynomial of degree k.

(b) $t^n = \sum_{k\in\mathbb{Z}} \mathbf{P}(S_n = k) H_{|k|}(t)$

Proof. To show (a), fix $t \in [-1,1]$ and set $s = \sqrt{1-t^2}$. Observe that

$$H_k(t) = \frac{1}{2}\sum_{r=0}^{k}\binom{k}{r}t^{k-r}[(is)^r + (-is)^r] = \frac{1}{2}\sum_{r=0}^{k/2}\binom{k}{2r}t^{k-2r}\psi(s)$$

where ψ is some real function of s.
To show (b) set $z_1 = t + is$ and $z_2 = t - is$ so that $|z_1| = |z_2| = 1$ and $z1z2 = 1$. Then,

$$H_k(t) = \frac{1}{2}(z_1^k + z_2^k) = H_{-k}(t) \implies |H_k(t) \leq 1|$$

Now, observe that $t = (z_1 + z_2)/2$, so that

$$t^n = \sum_{k=0}^{n}\frac{1}{2^n}\binom{n}{k}z_1^k z_2^{n-k} = \sum_{k=0}^{n}\frac{1}{2^n}\binom{n}{k}z_1^{2k-n} = \frac{1}{2^n}\sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)z_1^r$$

Repeating the same arguments above, we get

$$t^n = \frac{1}{2^n}\sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)z_1^r = \frac{1}{2^n}\sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)z_2^r$$

$$\implies t^n = \frac{1}{2^n}\sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)\left(\frac{z_1^r + z_2^r}{2}\right) = \sum_{r\in\mathbb{Z}}\mathbf{P}(S_n = r)H_{|r|}(t)$$

□

Theorem 8.0.2. *(Carne-Varopoulos bound) Let (Γ, μ) be a weighted graph. Then, for every $x, y \in V$ and $n \geq 1$*

$$p_n(x,y) \leq \frac{2}{\sqrt{\mu_x \mu_y}}\exp\left(-\frac{d(x,y)^2}{2n}\right)$$

Proof. Proved in Worksheet 2. □

Large Deviations for Random Walks

Let $\{\xi_i\}_{i\geq 1}$ be IID $\mathbb{Z}$ valued random variables such that $\mathbf{E}[\xi_1] = \mu$ and $\mathrm{Var}[\xi_1] < \infty$. Define $S_0 = 0$ and $S_n = \sum_{i=1}^n \xi_i$. Then, the strong law of large numbers (SLLN) and the central limit theorem (CLT) respectively state that

$$\mathbf{P}\left(\lim_{n\to\infty}\frac{S_n}{n} = \mu\right) = 1 \quad \text{and} \quad \frac{S_n - n\mu}{\sqrt{n}} \xrightarrow{d} \mathcal{N}(0,1)$$

Thus, the CLT loosely states that $S_n \approx n\mu + \sqrt{n}Z$, where $Z \sim \mathcal{N}(0,1)$.

As an exercise, show that for every $\epsilon > 0$, $\mathbf{P}(A_n^\epsilon) \to 0$ as $n \to \infty$, where $A_n^\epsilon = \{S_n \geq n(\mu+\epsilon)\}$. What is the rate of decay of $\mathbf{P}(A_n^\epsilon)$ (as $n \to \infty$)?
(Hint: $\mathbf{P}(S_n \geq n(\mu+\epsilon)) \approx \mathbf{P}(\xi_i > \mu + \epsilon \ \forall\ 1 \leq i \leq n) = [\mathbf{P}(\xi_1 > \mu+\epsilon)]^n \approx e^{-Cn}$ for some $C > 0$)

Theorem 8.0.3. *Let $\{\xi_i\}_{i\geq 0}$ be IID random variables with $\mathbf{P}(\xi_1 = 0) = \mathbf{P}(\xi_1 = 1) = 1/2$. Then, for every $a > 1/2$,*

$$\lim_{n\to\infty} \frac{1}{n} \log[\mathbf{P}(S_n \geq an)] = -I(a)$$

where

$$I(z) = \begin{cases} \log 2 + a\log a + (1-a)\log a & \text{if } 0 \leq z \leq 1 \\ \infty & \text{otherwise} \end{cases}$$

Observations:

(1) Minima of $I(z)$ is achieved at $z = 1/2$, and the graph increases from $[1/2, 1]$. This implies rate of exponential decay increases as $1/2 \to a \to 1$.

(2) Symmetry of the function $I(\cdot)$ around $1/2$ suggests that for $a < 1/2$, (Requires a proof)

$$\frac{1}{n} \log[\mathbf{P}(S_n \geq an)] \to -I(a)$$

(3) The theorem implies SLLN. The idea of the proof makes use of the following inequality

$$\mathbf{P}(S_n > (1/2+\delta)n) \leq \exp\{-I(n(1/2+\delta))\}$$

Proof.
If, $a > 1$ then, since S_n can be atmost n, $\mathbf{P}(S_n > an) = 0$ so the result follows. Now, consider $\frac{1}{2} < a \leq 1$, then

$$\mathbf{P}(S_n > an) = \sum_{an<k\leq n} \mathbf{P}(S_n = k) = \sum_{an<k\leq n} \binom{n}{k} \frac{1}{2^n} = \frac{1}{2^n} \sum_{an<k\leq n} \binom{n}{k}$$

Let, $Q_n(a) = \max\limits_{an<k\leq n} \binom{n}{k}$. So, we have,

$$2^{-n}\, Q_n(a) \leq \mathbf{P}(S_n > an) \leq 2^{-n}\, Q_n(a)\, (n+1) \tag{8.2}$$

First equality follows from the fact that one summand in the $\sum\limits_{an<k\leq n} \binom{n}{k}$ attains maximum and the second equality follows since, each summand of $\sum\limits_{0\leq k\leq n} \binom{n}{k}$ is $\leq Q_n(a)$.

Claim:

For, $\frac{1}{2} < a < 1$,

$$\frac{1}{n} \log Q_n(a) \xrightarrow[n\to\infty]{} -a\log a - (1-a)\log(1-a)$$

Now, from (8.2),

$$-\log 2 + \frac{1}{n}\log Q_n(a) \leq \frac{1}{n}\log \mathbf{P}(S_n > an) \leq -\log 2 + \frac{1}{n}\log Q_n(a) + \frac{1}{n}\log(n+1) \tag{8.3}$$

assuming the claim as LHS and RHS of (8.3) goes to $-I(a)$, the result follows. We now prove the claim.

Proof of claim:

Since, $a > \frac{1}{2}$, $\max_{an<k\leq n} \binom{n}{k} = \binom{n}{\lceil an \rceil}$. Now, from stirling's approximation

$$\binom{n}{\lceil an \rceil} = \frac{n!}{\lceil an \rceil!(n-\lceil an \rceil)!} \sim \frac{n^n e^{-n}\sqrt{2\pi n}}{\lceil an \rceil^{\lceil an \rceil} e^{-\lceil an \rceil}\sqrt{2\pi \lceil an \rceil}} \cdot \frac{1}{(n-\lceil an \rceil)^{n-\lceil an \rceil} e^{n-\lceil an \rceil}\sqrt{2\pi(n-\lceil an \rceil)}}$$

For, $a > \frac{1}{2}, a < 1$; $\lceil an \rceil \to \infty$ and $n - \lceil an \rceil \to \infty$ as $n \to \infty$ (Check!) and

$$\frac{1}{n}\log Q_n(a) \sim \frac{1}{n}\left[(n+\frac{1}{2})\log n - (\lceil an \rceil + \frac{1}{2})\log\lceil an \rceil - (n - \lceil an \rceil + \frac{1}{2})\log(n-\lceil an \rceil) - \log(\sqrt{2\pi})\right]$$
$$= \log n + \frac{1}{2n}\log n - \frac{\lceil an \rceil}{n}\log\lceil an \rceil - \frac{1}{2n}\log\lceil an \rceil - \frac{1}{n}\log\sqrt{2\pi} - \frac{n-\lceil an \rceil}{n}\log(n-\lceil an \rceil) - \frac{1}{2}\log(n-\lceil an \rceil)$$

the second, fourth, fifth and seventh summand of the above equation tends to 0 as n tends to ∞ and from the exercise (?) we have that

$$\frac{\lceil an \rceil}{n}\log\frac{\lceil an \rceil}{n} \xrightarrow[n\to\infty]{} a\log a \quad \text{and} \quad \frac{n-\lceil an \rceil}{n}\log\frac{n-\lceil an \rceil}{n} \xrightarrow[n\to\infty]{} (1-a)\log(1-a)$$

which proves the claim. $\square$

Cramer, 1930's

$\{\xi_i\}_{i\geq 1}$ i.i.d random variables with $\mathbf{E}[\xi_i] = \mu < \infty$, $\mathbf{E}[e^{r\xi_i}] < \infty$, $\forall\, r \in \mathbb{R}$. For any $a > \mu$,

$$\lim_{n\to\infty}\frac{1}{n}\log\mathbf{P}(S_n > an) = -I(a)$$

where, $I(a) = \sup_{z\in\mathbb{R}}[za - \mathbf{E}[e^{z\xi}]]$

Sanov, 1961 (`Level 2 of LDP`)

$\mathbf{P}(S_n > an) = \mathbf{P}\circ S_n^{-1}((an,\infty)) := \mu_n((an,\infty))$

$$-\frac{1}{n}\log\mu_n((an,\infty)) \xrightarrow[n\to\infty]{} \infty$$

Varadhan's LDP setup, 1960's

Let, $X_n : \Omega \to \mathbb{R}$ be a random variable of $(\Omega, \mathcal{F}, \mathbf{P})$. A be an event, $\mathbf{P}_n(A) := \mathbf{P}(S_n \in A)$, then $\mathbf{P}(\cdot)$ is a probability on $\mathbb{R}$.

A sequence $\{\mathcal{P}_n\}_{n\geq 1}$ of probaility measures on $\mathbb{R}$ (can be any metric space (X,d)) is said to satisfy large deviation principle with rate n and rate function $I : \mathbb{R} \to [0,\infty)\cup\{\infty\}$, if

1. $I \not\equiv \infty$, I is lower-semi continuous and has compact level sets.
2. $\overline{\lim\limits_{n\to\infty}} \frac{1}{n} \log \mathbf{P}_n(\mathcal{C}) \leq -I(\mathcal{C}) \, \forall$ closed sets $\mathcal{C}$
3. $\underline{\lim\limits_{n\to\infty}} \frac{1}{n} \log \mathbf{P}_n(\mathcal{O}) \geq -I(\mathcal{O}) \, \forall$ open sets $\mathcal{O}$

where, $A \subseteq \mathbb{R}$, $I(A) = \inf\limits_{y\in A} I(y)$.

Theorem 8.0.4. *$\{\mathcal{P}_n\}_{n\geq 1}$ satisfied LDP with rate n then , $I(\cdot)$ is unique.*

Theorem 8.0.5 (Varadhan's lemma)**.** *If, $\{\mathcal{P}_n\}_{n\geq 1}$ satisfies LDP with rate n and rate function $I(\cdot)$, let $F_n(x) = \mathbf{P}_n((-\infty, n])$ for some continuous and bounded above function $F : \mathbb{R} \to \mathbb{R}$, we have*

$$\int e^{nF(x)} dF_n(x) \xrightarrow[n\to\infty]{} \sup_{x\in\mathbb{R}} [F(x) - I(x)]$$

Applications

For, $\theta \in S^1$, $t \in \mathbb{R}$, $u : S^1 \times \mathbb{R}_+ \to \mathbb{R}$,

$$\begin{aligned} \frac{\partial u}{\partial t} &= \frac{1}{2}\frac{\partial^2 u}{\partial \theta^2} + V(\theta)u \\ u(0,\theta) &= 1 \end{aligned}$$

then,

$$\frac{1}{t} \log u(t,\theta) \xrightarrow[t\to\infty]{} \lambda_1 = \sup_{f\in\cdots} \left\{ \int V(\theta) f(\theta) d\theta - \frac{1}{8} \int \frac{(f'(\theta)^2)}{f(\theta)} d\theta \right\}$$

we can represent this as follows,

$$u(t,\theta) = \mathbf{E}\, e^{\int_0^t V(\theta_s)ds}, \ \{\theta_s\} - \text{brownian motion on } S^1$$

Exercises

1. For any $a \in \mathbb{R}$, show that,

$$\frac{\lceil an \rceil}{n} \xrightarrow[n\to\infty]{} a \ \text{ and } \ \frac{n - \lceil an \rceil}{n} \xrightarrow[n\to\infty]{} 1 - a$$