Large Deviations.
A survey

S.R.S. Varadhan
Courant Institute, NYU

ICTS, Bangalore
Jan 1, 2010

- The theory of large deviations deals with techniques for estimating exponentially small probabilities.

- The theory of large deviations deals with techniques for estimating exponentially small probabilities.
- It depends heavily on the duality between the functions e^x and $xlogx - x$.

- The theory of large deviations deals with techniques for estimating exponentially small probabilities.
- It depends heavily on the duality between the functions e^x and $xlogx - x$.
- $$P_n(A) = \exp[-n \inf_{x \in A} I(x) + o(n)]$$

  for "nice" sets A.

- More precisely

$$P_n(U) \geq \exp[-n \inf_{x \in U} I(x) + o(n)]$$

for open sets $U \subset X$

- More precisely

$$P_n(U) \geq \exp[-n \inf_{x \in U} I(x) + o(n)]$$

for open sets $U \subset X$

- For closed sets $C \subset X$,

$$P_n(C) \leq \exp[-n \inf_{x \in C} I(x) + o(n)]$$

- Equivalently

$$\int \exp[nF(x)]dP_n = \exp[n \sup_x [F(x) - I(x)] + o(n)]$$

for bounded continuous functions $F(x)$ on X.

- Equivalently

$$\int \exp[nF(x)]dP_n = \exp[n \sup_x [F(x) - I(x)] + o(n)]$$

for bounded continuous functions $F(x)$ on X.

- The non-negative rate function $I(\cdot)$ is assumed to be lower semicontinuous and with compact level sets

$$K_\ell = \{x : I(x) \leq \ell\}$$

- If X_i are i.i.d. random variables with finite exponential moments

$$M(\theta) = E[\exp[\theta X]]$$

then the distribution P_n of $Z_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ satisfies an LDP with rate function

- If X_i are i.i.d. random variables with finite exponential moments

$$M(\theta) = E[\exp[\theta\, X]]$$

then the distribution P_n of $Z_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ satisfies an LDP with rate function

-

$$I(x) = \sup_{\theta}[\theta x - \log M(\theta)]$$

- This is a theorem of Cramér (1937).

- This is a theorem of Cramér (1937).
- One can replace real X_i with independent random variables with values in R^d.

$$M(\theta) = E[\exp[\langle\theta, X\rangle]]$$

and

$$I(x) = \sup_{\theta \in R^d} [\langle\theta, X\rangle - \log M(\theta)]$$

- This is a theorem of Cramér (1937).
- One can replace real X_i with independent random variables with values in R^d.

$$M(\theta) = E[\exp[\langle \theta, X \rangle]]$$

and

$$I(x) = \sup_{\theta \in R^d} [\langle \theta, X \rangle - \log M(\theta)]$$

- More generally $\{X_i\}$ could be i.i.d random variables with values in some X with a common distribution α

$$Z_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$$

- $$Z_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$$
- P_n to be the distribution of Z_n with values in $\mathcal{M}(X)$ and Cramér's theorem morphs into Sanov's theorem.

- $$Z_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{X_i}$$

- P_n to be the distribution of Z_n with values in $\mathcal{M}(X)$ and Cramér's theorem morphs into Sanov's theorem.

- The dual of $\mathcal{M}(X)$ is $C(X)$ and

$$M(f) = E^{\alpha}[\exp[\langle \delta_X, f\rangle]] = \int e^{f(x)} d\alpha(x)$$

and for $\beta \in \mathcal{M}(X)$

- $$I(\beta) = \sup_{f \in C(X)} \left[\int f d\beta - \log \int e^f d\alpha \right]$$

- $$I(\beta) = \sup_{f \in C(X)} \left[\int f d\beta - \log \int e^f d\alpha \right]$$

- $$= h(\beta; \alpha) = h_\alpha(\beta)$$

- $$I(\beta) = \sup_{f \in C(X)} \left[\int f \, d\beta - \log \int e^f \, d\alpha \right]$$

- $$= h(\beta;\alpha) = h_\alpha(\beta)$$

- $h(\beta\,;\alpha) = \infty$ unless $\beta << \alpha$ and $b(x) = \frac{d\beta}{d\alpha}(x)$ is such that $|\log b(x)| \in L_1(\beta)$.

- $$I(\beta) = \sup_{f \in C(X)} [\int f d\beta - \log \int e^f d\alpha]$$
- $$= h(\beta;\alpha) = h_\alpha(\beta)$$
- $h(\beta\,;\alpha) = \infty$ unless $\beta << \alpha$ and $b(x) = \frac{d\beta}{d\alpha}(x)$ is such that $|\log b(x)| \in L_1(\beta)$.
- $$h(\beta\,;\alpha) = \int \log b(x)\, d\beta = \int b(x) \log b(x)\, d\alpha$$

Contraction principle

- $F : X \to Y, Q = PF^{-1}.$

$$dP = dQ\, dP_y$$

Contraction principle

- $F : X \to Y, Q = PF^{-1}$.

$$dP = dQ\, dP_y$$

- If P_n on X satisfies an LDP with rate $I(x)$ and $F : X \to Y$ is a continuous map then $Q_n = P_n F^{-1}$ on Y satisfies an LDP with rate

$$J(y) = \inf_{x : F(x) = y} I(x)$$

We optimize when we project.

- $\bar{X}_n = \int x d\nu_n$. $\nu \to \int x d\nu$ maps $\mathcal{M} \to R$.

- $\bar{X}_n = \int x d\nu_n$. $\nu \to \int x d\nu$ maps $\mathcal{M} \to R$.
- $$\inf_{\beta : \int x \, d\beta(x) = a} h(\beta \, ; \alpha) = I(a)$$

- $\bar{X}_n = \int x d\nu_n$. $\nu \to \int x d\nu$ maps $\mathcal{M} \to R$.
- $$\inf_{\beta : \int x \, d\beta(x) = a} h(\beta\,; \alpha) = I(a)$$
- $$= \sup_{\theta \in R} [a\theta - \log \int e^{\theta x} \, d\alpha(x)]$$

- The next step is to try and calculate $\psi_P(F)$

$$\lim_{n\to\infty} \frac{1}{n} \log E^P[\exp[\sum_{i=1}^{n} F(x_i, x_{i+1}, \ldots, x_{i+k-1})]]$$

- The next step is to try and calculate $\psi_P(F)$

$$\lim_{n\to\infty} \frac{1}{n} \log E^P[\exp[\sum_{i=1}^{n} F(x_i, x_{i+1}, \ldots, x_{i+k-1})]]$$

- F is a function of k variables $(x_1, \ldots, x_k)$

- The next step is to try and calculate $\psi_P(F)$

$$\lim_{n\to\infty} \frac{1}{n} \log E^P[\exp[\sum_{i=1}^{n} F(x_i, x_{i+1}, \ldots, x_{i+k-1})]]$$

- F is a function of k variables $(x_1, \ldots, x_k)$
- P is stationary process with values in X. i.e a shift invariant probability measure on X^∞, i.e $P \in \mathcal{M}_s(X)$.

- The empirical process which looks at all the finite dimensional distributions

$$\frac{1}{n}\sum_{i=1}^{n}\delta_{x_1,x_{i+1},\ldots,x_{i+k-1}}$$

- The empirical process which looks at all the finite dimensional distributions

$$\frac{1}{n}\sum_{i=1}^{n} \delta_{x_1, x_{i+1}, \ldots, x_{i+k-1}}$$

- Start with $(x_1, x_2, \ldots x_n)$ extend it periodically to get a sequence $\omega \in X^\infty$ and consider the orbital measure

$$R_n(x_1, x_2, \ldots x_n) = \frac{1}{n}\sum_{i=1}^{n} \delta_{T^{i-1}\omega}$$

- Let P_n be the distribution of $R_n \in \mathcal{M}(\mathcal{M}_s(X))$ It satisfies an LDP with rate function $I_P(Q)$ and

$$\psi_P(F) = \sup_{Q \in \mathcal{M}_s(X)} [E^Q[F] - I_P(Q)]$$

- Let P_n be the distribution of $R_n \in \mathcal{M}(\mathcal{M}_s(X))$ It satisfies an LDP with rate function $I_P(Q)$ and

$$\psi_P(F) = \sup_{Q \in \mathcal{M}_s(X)} [E^Q[F] - I_P(Q)]$$

- The rate function $I_P(Q)$ is universal and is a version of Kolmogorov-Sinai entropy.

- Let $p(dx_1|\omega)$ and $q(dx_1|\omega)$ be the conditional distributions of x_1 given the past $\{x_i : i \leq 0\}$ under P and Q respectively. Then

$$I_P(Q) = E^Q[h(q(\cdot|\omega)\,;p(\cdot|\omega))]$$

- Has one problem.

- Has one problem.
- Does not make sense!

- Has one problem.
- Does not make sense!
- The problem is that $p(\cdot|\omega)$ is only defined a.e. P and we need to integrate with respect to Q.

- Has one problem.
- Does not make sense!
- The problem is that $p(\cdot|\omega)$ is only defined a.e. P and we need to integrate with respect to Q.
- Put assumptions on P so that $p(\cdot|\omega)$ has a nice everywhere defined version. Markov will do it.

- if F is only a function of one variable $F(x_1)$ one can contract

$$I_P(\beta) = \inf_{\substack{Q \in \mathcal{M}_s(X) \\ Q x_1^{-1} = \beta}} I_P(Q)$$

- if F is only a function of one variable $F(x_1)$ one can contract

$$I_P(\beta) = \inf_{\substack{Q \in \mathcal{M}_s(X) \\ Q x_1^{-1} = \beta}} I_P(Q)$$

- In the Markov case

$$I_\pi(\beta) = \sup_{u>0} \int \log \frac{u(x)}{(\pi u)(x)} \, d\beta(x)$$

- if F is only a function of one variable $F(x_1)$ one can contract

$$I_P(\beta) = \inf_{\substack{Q \in \mathcal{M}_s(X) \\ Q x_1^{-1} = \beta}} I_P(Q)$$

- In the Markov case

$$I_\pi(\beta) = \sup_{u>0} \int \log \frac{u(x)}{(\pi u)(x)} \, d\beta(x)$$

- Controls the large deviations of $\frac{1}{n} \sum \delta_{x_i}$

- if F is only a function of one variable $F(x_1)$ one can contract

$$I_P(\beta) = \inf_{\substack{Q \in \mathcal{M}_s(X) \\ Q x_1^{-1} = \beta}} I_P(Q)$$

- In the Markov case

$$I_\pi(\beta) = \sup_{u>0} \int \log \frac{u(x)}{(\pi u)(x)} \, d\beta(x)$$

- Controls the large deviations of $\frac{1}{n} \sum \delta_{x_i}$
- If $\pi(x, dy) = \alpha(dy)$ then we are back in the Sanov case.

- We now turn to a more general problem. Calculate

$$J = \lim_{n\to\infty} \frac{1}{n} E^P[\exp[\sum_{i=1}^{n} a_i F(x_i, x_{i+1}, \ldots, x_{i+k-1})]]$$

for a given sequence $\{a_i : i \geq 1\}$.

- When will it exist?

- When will it exist?
- What will it be?

- When will it exist?
- What will it be?
- What is it good for?

- We note that when P is a product measure and F is a function of one variable, we need to have the limit

$$\log J = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(a_i)$$

- We note that when P is a product measure and F is a function of one variable, we need to have the limit

$$\log J = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} f(a_i)$$

- where

$$f(a) = \log E[\exp[aF(x)]]$$

- We note that when P is a product measure and F is a function of one variable, we need to have the limit

$$\log J = \lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} f(a_i)$$

- where

$$f(a) = \log E[\exp[aF(x)]]$$

- This requires the empirical distribution $\frac{1}{n}\sum_{i=1}^{n}\delta_{a_i}$ to have a limit.

- More generally we can assume that all the finite dimensional joint distributions $\frac{1}{n}\sum_{i=1}^{n}\delta_{a_i,a_{i+1},\ldots,a_{i+k-1}}$ have limits.

- More generally we can assume that all the finite dimensional joint distributions $\frac{1}{n}\sum_{i=1}^{n}\delta_{a_i,a_{i+1},\dots,a_{i+k-1}}$ have limits.
- The empirical process $R_n(a_1,\dots,a_n)$ has a limit $\nu \in \mathcal{M}_s(X)$

- More generally we can assume that all the finite dimensional joint distributions $\frac{1}{n}\sum_{i=1}^{n}\delta_{a_i,a_{i+1},\ldots,a_{i+k-1}}$ have limits.
- The empirical process $R_n(a_1,\ldots,a_n)$ has a limit $\nu \in \mathcal{M}_s(X)$
- It looks like a sample from ν.

- If $\{x_i\}$ is a Markov process with positive transition probabilities on a finite set X, then for every $\{a_i\}$ such that $R_n(a_1, \ldots, a_n) \to \nu$ the limit $J = J(\nu)$ exists and is a continuous linear function of ν

- If $\{x_i\}$ is a Markov process with positive transition probabilities on a finite set X, then for every $\{a_i\}$ such that $R_n(a_1, \ldots, a_n) \to \nu$ the limit $J = J(\nu)$ exists and is a continuous linear function of ν
- Sub-additive argument will do it.

- If $\{x_i\}$ is a Markov process with positive transition probabilities on a finite set X, then for every $\{a_i\}$ such that $R_n(a_1, \ldots, a_n) \to \nu$ the limit $J = J(\nu)$ exists and is a continuous linear function of ν
- Sub-additive argument will do it.
- What is it?

- Let $\mathcal{M}_\nu$ be the set of stationary process Q with values in $R \times X$ such that the marginal on R^∞ is ν.

- Let $\mathcal{M}_\nu$ be the set of stationary process Q with values in $R \times X$ such that the marginal on R^∞ is ν.
- Let $Q_0 = \nu \times P$, i.e $\{a_i\}$ has ν for its distribution and while $\{x_i\}$ is distributed according to P, the two components are independent.

- The Kolomogorov-Sinai entropy $H(Q; Q_0)$

- The Kolomogorov-Sinai entropy $H(Q; Q_0)$
-

$$E^Q[h(q(dx_1, da_1|\omega)\,; q_0(dx_1|\omega) \times \nu(da_1|\omega))$$

is well defined for every ν, provided $Q \in \mathcal{M}_\nu$.

- The Kolomogorov-Sinai entropy $H(Q;Q_0)$
-

$$E^Q[h(q(dx_1, da_1|\omega)\,; q_0(dx_1|\omega) \times \nu(da_1|\omega))$$

is well defined for every ν, provided $Q \in \mathcal{M}_\nu$.
- $q_0(dx_1|\omega) = q_0(dx_1|\omega_1), \nu(da_1|\omega) = \nu(da_1|\omega_2)$

- The Kolomogorov-Sinai entropy $H(Q;Q_0)$
-
$$E^Q[h(q(dx_1,da_1|\omega)\,;q_0(dx_1|\omega)\times\nu(da_1|\omega))$$
is well defined for every ν, provided $Q\in\mathcal{M}_\nu$.
- $q_0(dx_1|\omega)=q_0(dx_1|\omega_1),\nu(da_1|\omega)=\nu(da_1|\omega_2)$
- Then $J(\nu)$ is equal to
$$\sup_{Q\in\mathcal{M}_\nu}\ [E^Q[a_1F(x_1,\ldots,x_k)]-H(Q;Q_0)]$$

- Define $H_n(a_1, \ldots, a_n)$

$$\log E^P[\exp[\sum_{i=1}^n a_i F(x_i, x_{i+1}, \ldots, x_{i+k-1}]]$$

- Define $H_n(a_1, \ldots, a_n)$

$$\log E^P[\exp[\sum_{i=1}^{n} a_i F(x_i, x_{i+1}, \ldots, x_{i+k-1}]]$$

-

$$\begin{aligned} |H_{n+m}(a_1, \ldots, a_{n+m}) - H_n(a_1, \ldots, a_n) \\ - H_m(a_1, \ldots, a_m)| \leq C \end{aligned}$$

uniformly in n, m and $\{a_i\}$.

$$\left|H_{nk}(a_1, \ldots, a_{nk}) - \sum_{i=1}^{k} H_n(a_{(i-1)k+1}, \ldots, a_{ik})\right| \le Cn$$

- $$\left|H_{nk}(a_1,\ldots,a_{nk}) - \sum_{i=1}^{k} H_n(a_{(i-1)k+1},\ldots,a_{ik})\right| \leq Cn$$

- Partitioning of a block of size kn into blocks of size k allows some freedom as to the location and one can average over this collection

$$\limsup_{n\to\infty} \left|\frac{1}{n}H_n(a_1,\ldots,a_n) - \int \frac{1}{k}H_k(a_1,\ldots,a_k)\,dR_n\right| \le \frac{C}{k}$$

- $A_n(x_1, x_2, \ldots x_n)$ is almost additive if

- $A_n(x_1, x_2, \dots x_n)$ is almost additive if
-

$$|A_{n+m}(x_1,x_2, \dots x_n) - A_n(x_1, x_2, \dots x_n) \\ - A_m(x_{n+1}, x_2, \dots x_{n+m})| \leq C$$

- $A_n(x_1, x_2, \ldots x_n)$ is almost additive if
- $$|A_{n+m}(x_1,x_2, \ldots x_n) - A_n(x_1, x_2, \ldots x_n) - A_m(x_{n+1}, x_2, \ldots x_{n+m})| \leq C$$
- Then for every ν typical $\{x_i\}$ the limit

$$\lim_{n\to\infty} \frac{1}{n} A_n(x_1, x_2, \ldots x_n) = \mathcal{A}(\nu)$$

exists and is a continuous linear functional on $\mathcal{M}_s(X)$

$$\mathcal{H}(\nu) = \lim_{n\to\infty} \frac{1}{n} H_n(a_1, a_2, \ldots, a_n)$$
$$= \lim_{k\to\infty} \int H_k(a_1, a_2, \ldots, a_k) d\nu$$

exists and depends (linearly and continuously) on ν.

- A slight variant is the following theorem.

- A slight variant is the following theorem.
- If $H_n((a_1, x_1), \ldots, (a_n, x_n))$ is an almost additive sequence on $\{(A \times X)^n\}$ with $\frac{H_n}{n} \to \mathcal{H}(\cdot)$,

- A slight variant is the following theorem.
- If $H_n((a_1, x_1), \ldots, (a_n, x_n))$ is an almost additive sequence on $\{(A \times X)^n\}$ with $\frac{H_n}{n} \to \mathcal{H}(\cdot)$,
- $K_n(a_1, \ldots, a_n)$ defined as

$$\log E^P[\exp[H_n((a_1, x_1), \ldots, (a_n, x_n))]]$$

is almost additive.

- $\frac{K_n}{n}$ converges to $\mathcal{K}(\cdot)$ where

- $\frac{K_n}{n}$ converges to $\mathcal{K}(\cdot)$ where
-

$$\mathcal{K}(\nu) = \sup_{Q \in \mathcal{M}_\nu} \left[\mathcal{H}(Q) - H_{\nu \times P}(Q) \right]$$

- An application.

- An application.
- Calculate

$$J = \lim_{n\to\infty} \frac{1}{n} \log E^P[\exp[\sum_{i=1}^{n} U(x_i)V(x_{2i})]]$$

- An application.
- Calculate

$$J = \lim_{n\to\infty} \frac{1}{n} \log E^P[\exp[\sum_{i=1}^{n} U(x_i)V(x_{2i})]]$$

- To begin with consider the following simpler problem. $\{x_i\}$ are i.i.d and U, V are bounded. Let n be large say $2^k\, n$ for large k.

- We split the sum

$$S_{2^k n} = S_n^1 + S_n^2 + \cdots + S_n^k$$

- We split the sum

$$S_{2^k n} = S_n^1 + S_n^2 + \cdots + S_n^k$$

-

$$S_n^1 = \sum_{j=1}^{n} U(x_j)V(x_{2j})$$

- We split the sum

$$S_{2^k n} = S_n^1 + S_n^2 + \cdots + S_n^k$$

- $$S_n^1 = \sum_{j=1}^{n} U(x_j)V(x_{2j})$$

- For $i \geq 2$

$$S_n^i = \sum_{j=2^{i-1}n+1}^{2^i n} U(x_j)V(x_{2j})$$

- We denote by $\mathcal{F}_n$ the σ-field generated by $\{x_i : i \leq n\}$.

- We denote by $\mathcal{F}_n$ the σ-field generated by $\{x_i : i \leq n\}$.
- S_n^i is measurable w.r.t $\mathcal{F}_{2^{i+1}n}$

$$E[e^{S_{2^k n}} | \mathcal{F}_{2^k n}] = e^{S_{2^{k-1}n}} E[e^{S_n^k} | \mathcal{F}_{2^k n}]$$

$$= e^{S_{2^{k-1}n}} e^{\sum_{j=2^{k-1}n+1}^{2^k n} \psi_1(x_j)}$$

- We denote by $\mathcal{F}_n$ the σ-field generated by $\{x_i : i \leq n\}$.
- S_n^i is measurable w.r.t $\mathcal{F}_{2^{i+1}n}$

$$E[e^{S_{2^k n}}|\mathcal{F}_{2^k n}] = e^{S_{2^{k-1}n}} E[e^{S_n^k}|\mathcal{F}_{2^k n}]$$
$$= e^{S_{2^{k-1}n}} e^{\sum_{j=2^{k-1}n+1}^{2^k n} \psi_1(x_j)}$$

- where $\psi_1(x) = \log E^y[e^{U(x)V(y)}]$. The odd j's do not appear again and hence $2^{k-2}n$ terms separate out.

- We can remove $2^{k-2}n$ factors of

$$E[e^{\psi_1(x)}] = e^{c_1}$$

- We can remove $2^{k-2}n$ factors of

$$E[e^{\psi_1(x)}] = e^{c_1}$$

- Left with

$$e^{S_{2^{k-1}n}} e^{\sum_{j=2^{k-2}n+1}^{2^{k-1}n} \psi_1(x_{2j})}$$

- We can remove $2^{k-2}n$ factors of

$$E[e^{\psi_1(x)}] = e^{c_1}$$

- Left with

$$e^{S_{2^{k-1}n}} e^{\sum_{j=2^{k-2}n+1}^{2^{k-1}n} \psi_1(x_{2j})}$$

- Now condition with respect to $\mathcal{F}_{2^{k-1}n}$. We get

$$e^{S_{2^{k-2}n}} E\left[e^{\sum_{j=2^{k-2}n+1}^{2^{k-1}n} [U(x_j)V(x_{2j})+\psi_1(x_{2j})]} \middle| \mathcal{F}_{2^{k-1}n}\right]$$

- If we define

$$\psi_2(x) = \log E[e^{U(x)V(y)+\psi_1(y)}]$$

- If we define

$$\psi_2(x) = \log E[e^{U(x)V(y)+\psi_1(y)}]$$

- This reduces to

$$e^{S_{2^{k-2}n}} e^{\sum_{j=2^{k-2}N+1}^{2^{k-1}n} e^{\psi_2(x_j)}}$$

- If we define

$$\psi_2(x) = \log E[e^{U(x)V(y)+\psi_1(y)}]$$

- This reduces to

$$e^{S_{2^{k-2}n}} e^{\sum_{j=2^{k-2}N+1}^{2^{k-1}n} e^{\psi_2(x_j)}}$$

- Again the odd ones stand alone. Can factor $2^{k-3}n$ factors of

$$E[e^{\psi_2(x)}] = E[e^{U(x)V(y)+\psi_1(y)}] = e^{c_2}$$

- Left with

$$e^{S_{2^{k-2}n}} e^{\sum_{j=2^{k-3}n+1}^{2^{k-2}n} e^{\psi_2(x_{2j})}}$$

- Left with

$$e^{S_{2^{k-2}n}} e^{\sum_{j=2^{k-3}n+1}^{2^{k-2}n} e^{\psi_2(x_{2j})}}$$

- Continuing recursively

$$\psi_k(x) = \log E[e^{U(x)\,V(y)+\psi_{k-1}(y)}]$$

- Left with

$$e^{S_{2^{k-2}n}} e^{\sum_{j=2^{k-3}n+1}^{2^{k-2}n} e^{\psi_2(x_{2j})}}$$

- Continuing recursively

$$\psi_k(x) = \log E[e^{U(x)\,V(y)+\psi_{k-1}(y)}]$$

- $$E[e^{\psi_k(x)}] = E[e^{U(x)V(y)+\psi_{k-1}(y)}] = e^{c_k}$$

  We have some terms left over from $1 \leq j \leq n$. But for k large they can be ignored. Hence the limit equals

$$\lim_{n\to\infty} \frac{1}{n} \log E[\exp[\sum_{i=1}^{n} U(x_i)V(x_{2i})]]$$

$$= \sum_{j=1}^{\infty} \frac{c_j}{2^{j+1}}$$

- $$\lim_{n\to\infty}\frac{1}{n}\log E[\exp[\sum_{i=1}^{n}U(x_i)V(x_{2i})]] = \sum_{j=1}^{\infty}\frac{c_j}{2^{j+1}}$$

- Now we turn to the more general case where $\{x_i\}$ is a finite state space Markov chain that is mixing.

- Let

$$f_n = \frac{1}{n} \log E^P[\exp[\sum_{i=1}^{n} U(x_i)V(x_{2j})]]$$

- Let

$$f_n = \frac{1}{n} \log E^P[\exp[\sum_{i=1}^{n} U(x_i)V(x_{2j})]]$$

- Replace n by $2^k\, n$. We can pretend x_{2j} for $j \geq 2^{k-1}n + 1$ are independent of what happened before i.e. x_j for $j \leq 2^k n$.

$$f_{2^k n} = \frac{1}{2^k n} \log E^{P\times P}[\exp[\sum_{i=1}^{2^{k-1}n} U(x_i)V(x_{2j})$$

$$+ \sum_{i=2^{k-1}n+1}^{2^k n} U(x_i)V(y_{2j})]]$$

$$= \frac{1}{2^k n} \log E^{P}[\exp[\sum_{i=1}^{2^{k-1}n} U(x_i)V(x_{2j})$$

$$+ H_{2^{k-1}n}(U(x_{2^{k-1}n+1}), \ldots, U(x_{2^k n}))]]$$

- Now we peel off $j \leq 2^{k-2}n$ and pretend the rest is independent.

- Now we peel off $j \leq 2^{k-2}n$ and pretend the rest is independent.
- $$f_{2^k n} \simeq \frac{1}{2^k n} \log E^{P \times P}[\exp[\sum_{i=1}^{2^{k-2}n} U(x_i)V(x_{2j}) \\ + \sum_{i=2^{k-2}n+1}^{2^{k-1}n} U(x_i)V(y_{2j})]] \\ + H_{2^{k-1}n}(U(y_{2^{k-1}n+1}), \ldots, U(y_{2^k n}))]]$$

- The induction step is

- The induction step is
-

$$H^{r+1}_{2^{k-r-1}n}(a_1,\dots,a_{2^{k-r-1}n})$$

$$= \log E^P\Bigg[\exp\Big[\sum_{i=2^{k-r-1}+1}^{2^{k-r}} a_i V(x_{2i})$$

$$+ H^r_{2^{k-r}n}(U(x_{2^{k-r}n+1}),\dots,U(x_{2^{k-r+1}n}))\Big]\Bigg]$$

- The induction step is
-

$$
\begin{aligned}
H^{r+1}_{2^{k-r-1}n}&(a_1, \ldots, a_{2^{k-r-1}n}) \\
&= \log E^P \Big[\exp[\sum_{i=2^{k-r-1}+1}^{2^{k-r}} a_i V(x_{2i}) \\
&+ H^r_{2^{k-r}n}(U(x_{2^{k-r}n+1}), \ldots, U(x_{2^{k-r+1}n}))] \Big]
\end{aligned}
$$

- This leads to a map

$$\mathcal{H}_r(\cdot) \longrightarrow \mathcal{H}_{r+1}(\cdot)$$

- Starting from $\mathcal{H}_0(\cdot) \equiv 0$. After many iterations $\mathcal{H}_r(\nu)$ will be nearly a constant. Or we drop the first n terms and just calculate

- Starting from $\mathcal{H}_0(\cdot) \equiv 0$. After many iterations $\mathcal{H}_r(\nu)$ will be nearly a constant. Or we drop the first n terms and just calculate
-

$$\lim_{n\to\infty} \frac{1}{2^k n} \log E^P[\exp[H_n^k(U(x_{n+1}, \ldots, x_{2n}))]]$$

- Starting from $\mathcal{H}_0(\cdot) \equiv 0$. After many iterations $\mathcal{H}_r(\nu)$ will be nearly a constant. Or we drop the first n terms and just calculate
- $$\lim_{n\to\infty} \frac{1}{2^k n} \log E^P[\exp[H_n^k(U(x_{n+1}, \ldots, x_{2n}))]]$$
- Remark. One can extend this to the calculation of

$$\lim_{n\to\infty} \frac{1}{n} \log E^P[\exp[\sum_{i=1}^n f(x_i, x_{2i}, \ldots, x_{ki})]]$$

Last Slide

THE END