On Large Deviation Analysis of Sampling from Typical Sets

D.Krithivasan and S.S.Pradhan

University of Michigan

ITA 2007, San Diego

Acknowledgements

The authors would like to thank

- ITA organizing committee for the invitation
- NSF for financial support
- Prof. Prakash Ishwar of Boston University for helpful discussions

Presentation Outline

Outline

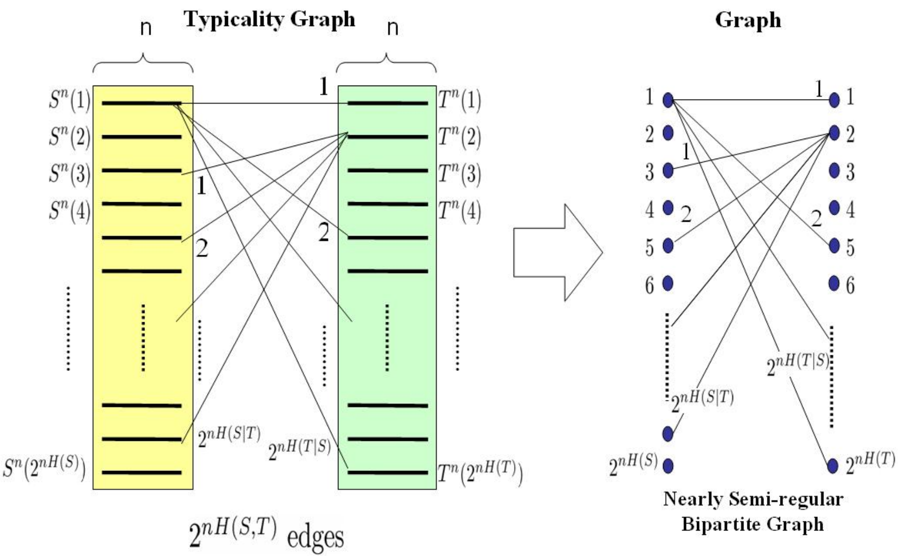
Typicality - An Overview

- Sequence s^n is *typical* with respect to a distribution $p(s)$ if its empirical histogram is close to $p(s)$
- Typical set $A_\epsilon^{(n)}(S)$ is the set of all n-length typical sequences
- Properties of typical sequences
  - $|A_\epsilon^{(n)}(S)| \approx 2^{nH(S)}$
  - S_i is drawn i.i.d $\sim p(s)$. Then $Pr(S^n \in A_\epsilon^{(n)}(S)) \to 1$ as $n \to \infty$
  - All typical sequences nearly equally likely

Typicality Contd.

- Typicality can be extended to pairs of sources (S, T) with distribution $p(s, t)$
- Roughly $2^{nH(S,T)}$ jointly typical sequences
- All such sequences are equally likely
- For every typical s^n, nearly $2^{nH(T|S)}$ typical t^n sequences such that (s^n, t^n) is jointly typical
- Joint typicality captured by a bipartite graph called the typicality graph

Typicality Graph



Random Sampling of the Typicality Graph

- Induce a random sampling of the typicality graph
- $\theta_1 = 2^{nR_1}$ (respectively $\theta_2 = 2^{nR_2}$) sequences are sampled from the typical set of S (respectively T) independently with replacement
- Will assume WLOG that $R_1 \geq R_2$
- We study the probability that this random graph
  - has no edges
  - has number of edges is significantly smaller than expected

Applications

- Correlated sources viewed through typicality graphs has applications in transmitting these sources through multiuser channels
- Partial characterizations of bipartite graphs that can be reliably transmitted over multiple-access channels and broadcast channels are available.

Outline

No Jointly Typical Sequences

- X, Y are correlated finite-alphabet random variables with distribution $p(x,y)$
- Given : $\epsilon > 0$, positive real numbers R_1 and R_2 such that $R_1 + R_2 > I(X;Y)$
- Pick 2^{nR_1} sequences $\mathscr{C}_X$ from typical set $A_\epsilon^{(n)}(X)$
- Pick 2^{nR_2} sequences $\mathscr{C}_Y$ from typical set $A_\epsilon^{(n)}(Y)$
- U - number of jointly typical sequences in this collection

Theorem

$$\lim_{n\to\infty} \frac{1}{n} \log\log \frac{1}{P(U=0)} \geq \min(R_2, R_1 + R_2 - I(X;Y))$$

Result holds with equality for $R_2 \leq R_1 < I(X;Y)$.

Few Jointly Typical Sequences

- Same assumptions as before
- Choose any $\gamma > 0$
- Result gives bound on probability that number of jointly typical sequences is exponentially smaller than expected
- Let A_n be the event $\frac{\mathbb{E}(U)-U}{\mathbb{E}(U)} \geq e^{-n\gamma}$

Theorem

$$\lim_{n\to\infty} \frac{1}{n} \log\log \frac{1}{P(A_n)} \geq \begin{cases} R_1 + R_2 - I(X;Y) - \gamma & \textit{if } R_1 < I(X;Y) \\ R_2 - \gamma & \textit{if } R_1 \geq I(X;Y) \end{cases}$$

Outline

Suen's Correlation Inequalities

- Suen's Inequalities deal with sums of possibly dependent random variables
- Uses Dependency graphs
    - Vertex i represents the indicator random variable I_i
    - Vertices i and j are connected if random variables I_i and I_j are dependent

Suen's Inequalities contd.

- $I_i,\ i \in \mathscr{I}$ - Bernoulli random variable of parameter p_i
- Corresponding dependency graph L has vertex set $\mathscr{I}$ and edge set $E(L)$
- Write $i \sim j$ if $(i,j) \in E(L)$
- $X = \sum_i I_i$

Theorem

Suen's Inequality I

$$P(X=0) \leq \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Suen's Inequalities contd.

- I_i, $i \in \mathscr{I}$ - Bernoulli random variable of parameter p_i
- Corresponding dependency graph L has vertex set $\mathscr{I}$ and edge set $E(L)$
- Write $i \sim j$ if $(i,j) \in E(L)$
- $X = \sum_i I_i$

Theorem

Suen's Inequality I

$$P(X=0) \le \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Definitions

$$\lambda \triangleq \mathbb{E}(X) = \sum_i p_i$$

Suen's Inequalities contd.

- I_i, $i \in \mathscr{I}$ - Bernoulli random variable of parameter p_i
- Corresponding dependency graph L has vertex set $\mathscr{I}$ and edge set $E(L)$
- Write $i \sim j$ if $(i,j) \in E(L)$
- $X = \sum_i I_i$

Theorem

Suen's Inequality I

$$P(X=0) \leq \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Definitions

$$\Delta \triangleq \frac{1}{2}\sum_i \sum_{j \sim i} \mathbb{E}(I_i I_j)$$

Suen's Inequalities contd.

- $I_i,\ i \in \mathscr{I}$ - Bernoulli random variable of parameter p_i
- Corresponding dependency graph L has vertex set $\mathscr{I}$ and edge set $E(L)$
- Write $i \sim j$ if $(i,j) \in E(L)$
- $X = \sum_i I_i$

Theorem

Suen's Inequality I

$$P(X=0) \le \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Definitions

$$\delta \triangleq \max_i \sum_{k \sim i} p_k$$

- Under same assumptions, we can also derive upper bounds for the lower tail of X

Theorem

Suen's Inequality II For $0 \le a \le 1$, *we have*

$$P(X \le a\lambda) \le \exp\left\{-\min\left((1-a)^2 \frac{\lambda^2}{8\Delta + 2\lambda}, (1-a)\frac{\lambda}{6\delta}\right)\right\}$$

Lovász Local Lemma

- Bound on probability that none of the events in a given collection $\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_n$ occurs
- L is the dependency graph for events $\mathscr{E}_1, \mathscr{E}_2, \ldots, \mathscr{E}_n$

Theorem

Lovász Local Lemma: Suppose there exists $x_i \in [0,1]$ for $1 \le i \le n$ such that

$$P(\mathscr{E}_i) \le x_i \prod_{(i,j) \in E(L)} (1 - x_j)$$

Then, we have the lower bound

$$P(\cap_{i=1}^{n} \overline{\mathscr{E}_i}) \ge \prod_{i=1}^{n} (1 - x_i)$$

Outline

Proof Idea

- For $1 \le i \le \theta_1$, $1 \le j \le \theta_2$, define Indicator random variables

$$U_{ij} = \begin{cases} 1 & (X^n(i), Y^n(j)) \in A_\epsilon^{(n)}(X,Y) \\ 0 & \text{else} \end{cases}$$

- Number of jointly typical sequences $U = \sum_i \sum_j U_{ij}$
- Use Suen's Inequality on this family of $\theta_1\theta_2$ indicator random variables

Dependency Graph

- Indicator random variable U_{ij} represented by vertex (i,j)
- U_{11} depends on U_{i1}, $2 \le i \le \theta_1$ and U_{1j}, $2 \le j \le \theta_2$
- Dependency graph is a regular graph
- Each vertex (i,j) is connected to exactly $\theta_1 + \theta_2 - 2$ vertices

Derivation of the Upper Bound

$$P(X = 0) \le \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Derivation of the Upper Bound

$$P(X=0) \le \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Bounds

Derivation of the Upper Bound

$$P(X=0) \leq \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Bounds

$$\lambda \quad \geq \quad \theta_1 \theta_2 2^{-n(I+\epsilon_1)}$$

Derivation of the Upper Bound

$$P(X=0) \leq \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Bounds

$$\lambda \geq \theta_1\theta_2 2^{-n(I+\epsilon_1)}$$

$$\Delta \leq \frac{1}{2}\theta_1\theta_2(\theta_1+\theta_2-2)2^{-2n(I-2\epsilon_2)}$$

Derivation of the Upper Bound

$$P(X=0) \leq \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Bounds

$$\lambda \geq \theta_1\theta_2 2^{-n(I+\epsilon_1)}$$

$$\Delta \leq \frac{1}{2}\theta_1\theta_2(\theta_1+\theta_2-2)2^{-2n(I-2\epsilon_2)}$$

$$\delta \leq (\theta_1+\theta_2-2)2^{-n(I-\epsilon_1)}$$

Derivation of the Upper Bound

$$P(X=0) \leq \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Bounds

$$\lambda \geq \theta_1\theta_2 2^{-n(I+\epsilon_1)}$$

$$\Delta \leq \frac{1}{2}\theta_1\theta_2(\theta_1+\theta_2-2)2^{-2n(I-2\epsilon_2)}$$

$$\delta \leq (\theta_1+\theta_2-2)2^{-n(I-\epsilon_1)}$$

$$\frac{\lambda^2}{8\Delta} \geq \frac{1}{8}2^{nR_2}$$

Derivation of the Upper Bound

$$P(X=0) \leq \exp\left\{-\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta}\right)\right\}$$

Bounds

$$\lambda \geq \theta_1\theta_2 2^{-n(I+\epsilon_1)}$$

$$\Delta \leq \frac{1}{2}\theta_1\theta_2(\theta_1+\theta_2-2)2^{-2n(I-2\epsilon_2)}$$

$$\delta \leq (\theta_1+\theta_2-2)2^{-n(I-\epsilon_1)}$$

$$\frac{\lambda^2}{8\Delta} \geq \frac{1}{8}2^{nR_2}$$

$$\frac{\lambda}{2} \geq \frac{1}{2}2^{n(R_1+R_2-I-\epsilon_1)}$$

Derivation of the Upper Bound

$$P(X=0) \le \exp\left\{ -\min\left(\frac{\lambda^2}{8\Delta}, \frac{\lambda}{2}, \frac{\lambda}{6\delta} \right) \right\}$$

Bounds

$$\lambda \ge \theta_1\theta_2 2^{-n(I+\epsilon_1)}$$

$$\Delta \le \frac{1}{2}\theta_1\theta_2(\theta_1+\theta_2-2)2^{-2n(I-2\epsilon_2)}$$

$$\delta \le (\theta_1+\theta_2-2)2^{-n(I-\epsilon_1)}$$

$$\frac{\lambda^2}{8\Delta} \ge \frac{1}{8}2^{nR_2}$$

$$\frac{\lambda}{2} \ge \frac{1}{2}2^{n(R_1+R_2-I-\epsilon_1)}$$

$$\frac{\lambda}{6\delta} \ge \frac{1}{12}2^{n(R_2-2\epsilon_1)}$$

Tail Estimates and Lower Bound

- Substitution gives main result for probability of non-existence of jointly typical sequences
- Using Suen's Inequality II gives bounds on tail estimates
- Lower bound derived using Lovász local lemma and coincides with the upper bound for $R_2 \leq R_1 < I(X;Y)$
- All results can be extended to the case of more than 2 random variables

Outline

Fully Connected Graph

- Pick M typical sequences from $A_\epsilon^{(n)}(X)$
- Pick N typical sequences from $A_\epsilon^{(n)}(Y)$
- We investigate probability that all MN pairs are jointly typical
- Call this event FC

Fully Connected Graph contd.

Theorem

$$\lim_{n\to\infty} -\frac{1}{n}\log P(FC) \geq (M+N-1)I(X;Y) + \min_{\mathscr{P}} (N-1)I(Y;X_2,\ldots X_M \mid X_1) + A(X_1;\ldots;X_M \mid Y)$$

Fully Connected Graph contd.

Theorem

$$\lim_{n\to\infty} -\frac{1}{n}\log P(FC) \geq (M+N-1)I(X;Y) + \min_{\mathscr{P}} (N-1)I(Y;X_2,\ldots X_M \mid X_1) + A(X_1;\ldots;X_M \mid Y)$$

- X_i, $1 \leq i \leq M$ are random variables of distribution P_X

Fully Connected Graph contd.

Theorem

$$\lim_{n\to\infty} -\frac{1}{n}\log P(FC) \geq (M+N-1)I(X;Y) + \min_{\mathscr{P}} (N-1)I(Y;X_2,\ldots X_M \mid X_1) + A(X_1;\ldots;X_M \mid Y)$$

- X_i, $1 \leq i \leq M$ are random variables of distribution P_X
- $\mathscr{P}$ - family of conditional distributions $P_{X_1,\ldots,X_M|Y}$

Fully Connected Graph contd.

Theorem

$$\lim_{n\to\infty} -\frac{1}{n}\log P(FC) \geq (M+N-1)I(X;Y) + \min_{\mathscr{P}} (N-1)I(Y;X_2,\ldots X_M \mid X_1) + A(X_1;\ldots;X_M \mid Y)$$

- X_i, $1 \leq i \leq M$ are random variables of distribution P_X
- $\mathscr{P}$ - family of conditional distributions $P_{X_1,\ldots,X_M|Y}$
- $A(X_1,\ldots,X_M) \triangleq \sum_{i=1}^{M} H(X_i) - H(X_1,\ldots,X_M)$

Fully Connected Graph - Example

- Lets take $M = N = 2$
- $$\lim_{n\to\infty} -\frac{1}{n}\log P(FC) \geq 3I(X;Y) + \min_{\mathscr{P}} I(Y;X_2 \mid X_1) + I(X_1;X_2 \mid Y)$$

Fully Connected Graph - Example

- Lets take $M = N = 2$
- $$\lim_{n\to\infty} -\frac{1}{n}\log P(FC) \geq 3I(X;Y) + \min_{\mathscr{P}} I(Y;X_2 \mid X_1) + I(X_1;X_2 \mid Y)$$
- X_1, X_2 are random variables of distribution P_X

Fully Connected Graph - Example

- Lets take $M = N = 2$
- $$\lim_{n\to\infty} -\frac{1}{n} \log P(FC) \geq 3I(X;Y) + \min_{\mathscr{P}} I(Y;X_2 \mid X_1) + I(X_1;X_2 \mid Y)$$
- X_1, X_2 are random variables of distribution P_X
- $\mathscr{P}$ - family of conditional distributions $P_{X_1,X_2|Y}$

Outline

Conclusions

- Joint typicality can be characterized by typicality graph
- Studied asymptotic properties of samples taken from it
- Derived bounds on the probabilities of the following events
  - Typicality graph has no edges
  - Typicality graph has significantly fewer edges than expected
- Derived bounds for the probability that the typicality graph is completely connected
- Results have applications in certain frameworks of transmitting correlated sources over multiuser channels