CSE525: Randomized Algorithms and Probabilistic Analysis May 7, 2013

Lecture 9

Lecturer: Anna Karlin *Scribe: Ravi Bhoraskar and Tom Guo*

1 The moment methods

The moment method is used to bound the probability that a random variable fluctuates far from its mean, by using its moments. These are used to prove properties about probabilistic structures, such as random graphs.

1.1 First Moment Method

The first moment method uses the simple fact that if a random variable X has expected value $E(X)$, then it takes some value $\geq E(X)$ and some value $\leq E(X)$.

Now, if $X \geq 0$ is an integer valued function, using the Markov Inequality, we can say that $Pr(X \leq 1) < E(X)$. Thus, if $E(X) << 1$, we can say that $X = 0$ with high probability. This simple application of the Markov inequality is known as the First Moment method.

Note however, that if we are told that $E(X) >> 1$, it is not possible to prove that $Pr(X \geq 1)$ is large using the first moment method. This is because X taking an enormous value with low probability is not ruled out by the first moment method.

1.2 Second Moment Method

The second moment method overcomes the limitations of the first moment method by using $E(X^2)$ instead of $E(X)$.

One version of the Second Moment Method is Chebychev's inequality, which states that

$$\forall \lambda > 0$$
$$Pr(|X - \mu| \geq \lambda\sigma) \leq \frac{1}{\lambda^2}$$

for a non negative integer valued random variable X

Another version states that

$$Pr(X = 0) \leq \frac{Var(X)}{(E(X))^2}$$

Proof

$$\begin{aligned} Pr(X = 0) &\leq Pr(|X - \mu| \geq \mu) \\ &\leq \frac{\sigma^2}{\mu^2} \\ &= \frac{Var(X)}{(E(X))^2} \end{aligned}$$

■

Corollary 1. *If $Var(X) = o(E(X)^2)$, then $Pr(X > 0) = 1 - o(1)$*

Another second moment inequality states the following:

Theorem 2. $Pr(X > 0) \geq \frac{(E[X])^2}{E[X^2]}$.

Proof For non-negative valued random variables X, Y with $E[X^2] > 0$ and $E[Y^2] > 0$. Let $U = \frac{X}{\sqrt{E[X^2]}}, V = \frac{Y}{\sqrt{E[Y^2]}}$. That $2|UV| \leq U^2 + V^2$ implies

$$2|E[UV]| \leq 2E[|UV|] \leq E[U^2] + E[V^2] = 2,$$

which implies $(E[UV])^2 \leq 1$, or equivalently, $(E[XY])^2 \leq E[X^2]E[Y^2]$.
Set $Y = \mathbf{1}_{X>0}$, and we have

$$(E[X])^2 \leq E[X^2]E[\mathbf{1}^2_{X>0}] = E[X^2]Pr(X > 0),$$

from which $Pr(X > 0) \geq \frac{(E[X])^2}{E[X^2]}$ immediately follows. ∎

2 Covariance in a Binomial Random Variable

For a bionomial random variable

$$X = X_1 + X_2 + ... + X_n$$

Claim 3. $Var(X) = \sum_{i=1}^{n} Var(X_i) + \sum_{i}\sum_{j\neq i} Cov(X_i, X_j)$, *where* $Cov(X, Y) = E(XY) - E(X)E(Y)$.

Proof

$$\begin{aligned}
Var(X) &= E[\left(\sum_{i=1}^{n} X_i\right)^2] - \left(E[\sum_{i=1}^{n} X_i]\right)^2 \\
&= \sum_{i=1}^{n} \left(E[X_i^2] - E[X_i]^2\right) + \sum_{i\neq j} E[X_iX_j] - E[X_i]E[X_j] \\
&= \sum_{i=1}^{n} Var(X_i) + \sum_{j\neq i} Cov(X_i, X_j).
\end{aligned}$$

∎

Often, the X_i's are indicator random variables, with $Pr(X_i = 1) = p_i$ and $Pr(X_i = 0) = 1 - p_i$. In this case,

$$\begin{aligned}
Var(X_i) &= p_i(1 - p_i) \\
&\leq p_i \\
&= E(X_i)
\end{aligned}$$

$$\implies Var(X) \leq E(X) + \sum_{i\neq j} Cov(X_i, X_j)$$

3 Random Graphs

3.1 4-cliques in random graphs

Definition 4. *Let $G_{n,p}$ be the random graph with n vertices and each edge independently present with probability p.*

Claim 5. *$G_{n,p}$ almost surely does not have a clique with 4 or more vertices for $p = o(n^{-2/3})$ and almost surely does for $p = \omega(n^{-2/3})$.*

Proof There are $\binom{n}{k}$ possible 4-cliques, which we enumerate as $C_1, C_2, \ldots, C_{\binom{n}{4}}$. Let $X_i, i = 1, 2, \ldots, \binom{n}{4}$ be indicator random variables for each of the cliques. Then,

$$X = \sum_{i=1}^{\binom{n}{4}} X_i$$

is a random variable for the number of 4-cliques with

$$E[X] = \binom{n}{4} p^6 \approx \frac{n^4 p^6}{4!}.$$

If $p = o(n^{-2/3})$, then by the first moment method $E[X] = o(1)$.
If $p = \omega(n^{-2/3})$, then $E[X] \to \infty$ as $n \to \infty$. If $Var(X) = o(E[X^2])$, then by the second moment method,

$$\limsup_{n\to\infty} Pr(X = 0) \leq \limsup_{n\to\infty} \frac{Var(X)}{E(X^2)} \leq 0.$$

Hence it suffices to show that $Var(X) = o(E[X^2])$. For pairs of possible cliques C_i, C_j, we have the following cases:
Case 1: $C_i \cap C_j = \emptyset$. $Cov(X_i, X_j) = 0$.
Case 2: $|C_i \cap C_j| = 1$. $Cov(X_i, X_j) = 0$.
Case 3: $|C_i \cap C_j| = 2$. $Cov(X_i, X_j) \leq E[X_i X_j] = p^{11}$.
Case 4: $|C_i \cap C_j| = 3$. $Cov(X_i, X_j) \leq E[X_i X_j] = p^9$.

$$\begin{aligned} Var(X) &\leq E[X] + \sum_{1\leq i\neq j\leq m} Cov(X_i, X_j) \\ &\leq \binom{n}{k} p^6 + \underbrace{\binom{n}{6}\binom{6}{2,2,2} p^{11}}_{\text{case 3}} + \underbrace{\binom{n}{5}\binom{5}{3,1,1} p^9}_{\text{case 4}} \\ &\leq c\left(n^4p^4 + n^6p^11 + n^5p^9\right) = o((E[X])^2). \end{aligned}$$

since $(E[X])^2 = \left(\binom{n}{4}p^6\right)^2 = \theta(n^8p^{12})$, $p = \omega(n^{-2/3})$. ■

3.2 Evaluation of random graphs

Theorem 6. *The following hold almost surely, or in other words, with probability going to 1 as $n \to \infty$.*

- *If $p = o\left(\frac{1}{n}\right)$, then the graph has no cycles.*
- *If $p \leq \frac{c}{n}, c < 1$, then the longest connected component has size $\theta(\log n)$.*
- *If $p = \frac{1}{n}$, then the longest connected component has size $\theta(n^{2/3})$.*
- *If $p \geq \frac{c}{n}$, then the longest connected component has size $\theta(n)$.*
- *If $p = \frac{\log}{n}$, then the graph is connected and has a Hamiltonian cycle.*

3.3 First order theory of random graphs

Let us define a language where variables $(x, y, z, \ldots)$ represent vertices, $x = y$ denotes equality of x and y, $x\ y$ denotes adjacency, and $\wedge, \vee, \neg, \forall, \exists$ denote logical and, logical or, logical not, universal quantifier, and existential quantifier respectively. Call this the first-order language. Below are examples:

- G contains a triangle: $\exists x \exists y \exists z\ x \sim y \wedge x \sim z \wedge y \sim z$.
- no isolated point: $\forall x \exists y (x\ y)$.
- diameter equals 2: $\forall x \forall y \exists z\ ((x = y) \vee (x \sim y) \vee (x \sim z \wedge y \sim z))$.

G is Hamiltonian, G is connected are not first order theory properties. That is, they cannot be expressed using the first order language. Given that A is a first-order property, let

$$Pr(G(n,p)| = A) \equiv Pr(\text{random graph from } G(n,p) \text{ has property } A).$$

Theorem 7 (Fagin; Glebskii, Kogan, Liagonkii, and Talanov)**.** *For all fixed p, $0 < p < 1$, and any first order graph property,*

$$\lim_{n\to\infty} Pr(G(n,p)| = A) = 0 \text{ or } 1.$$

Theorem 8 (Shelah and Spencer)**.** *For all irrational $\alpha \in (0,1)$, setting $p = p(n) = n^{-\alpha}$ and for any first order property A*

$$\lim_{n\to\infty} Pr(G(n,p)| = A) = 0 \text{ or } 1.$$

3.4 Percolation on a tree

Theorem 9. *Given the random complete binary tree of depth n where each edge is present with probability p, let X_i be a random indicator variable for the reachability of the ith leaf from the root and $X = \sum_{i=1}^{2^n} X_i$, the number of leaves reachable from the root. Then, if $p \geq \frac{1}{2}$*

$$\frac{2}{n+2} \leq Pr(X > 0).$$

Proof We have

$$\lim_{n\to\infty} E[X] = \lim_{n\to\infty} 2^n \cdot p^n = \begin{cases} \infty & p > \frac{1}{2} \\ 1 & p = \frac{1}{2} \\ 0 & p < \frac{1}{2} \end{cases}.$$

By the second moment method,

$$Pr(X > 0) \geq \frac{(E[X])^2}{E[X^2]}.$$

We have

$$X^2 = \left(\sum_{i=1}^{2^n} X_i\right)^2 = \sum_{i=1}^{2^n} X_i^2 + \sum_{1 \leq i \neq j \leq 2^n} X_i X_j.$$

Let $k(i,j)$ denote the depth of the node from which the paths from the root to the ith and jth leafs diverge. Then,

$$E[X_i X_j] = p^{2n-k(i,j)}$$

and for $p > \frac{1}{2}$,

$$\begin{aligned}
\sum_{1 \le i \ne j \le 2^n} X_i X_j &= \sum_{1 \le i \ne j \le 2^n} p^{2n-k(i,j)} \\
&= \sum_{k=0}^{n-1} 2^k 2^{n-k} 2^{n-k-1} p^{2n-k} \\
&= \frac{1}{2} \sum_{k=0}^{n} 2^{2n} \frac{1}{2^k} p^{2n} \frac{1}{p^k} \\
&= \frac{1}{2} (2p)^{2n} \sum_{k=0}^{n-1} \frac{1}{(2p)^k} \\
&= \frac{1}{2} (2p)^{2n} \cdot \frac{1 - \frac{1}{(2p)^n}}{1 - \frac{1}{2p}} \\
&\le \frac{1}{2} (2p)^{2n} \frac{2p}{2p-1} \\
&= \frac{p}{2p-1} (2p)^{2n}.
\end{aligned}$$

Hence,

$$E[X^2] \le \frac{p}{2p-1} (2p)^{2n} [1 + o(1)],$$

from which follows

$$Pr(X > 0) \ge \frac{(E[X])^2}{E[X^2]} \ge \frac{2p-1}{p} (1 - o(1)) > c$$

for $p > \frac{1}{2}$. For $p = \frac{1}{2}$,

$$\sum_{1 \le i \ne j \le 2^n} X_i X_j = \sum_{k=0}^{n-1} 2^k 2^{n-k} 2^{n-k-1} p^{2n-k} = \frac{n}{2}$$

and

$$E[X^2] = 1 + \frac{n}{2}, (E[X])^2 = 1,$$

from which follows

$$Pr(X > 0) \ge \frac{(E[X])^2}{E[X^2]} \ge \frac{2}{n+2},$$

which proves our lower bound. ■