Lecture 11

Lecturer: Nati Linial *Scribe: Ofer Neiman*

In the previous lecture we presented how the linearity of expectation is used in the *probabilistic method* to prove the existence of certain properties. In this lecture we show another important technique, known as the second moment method. We use the second moment of a random variable, i.e., $\mathbb{E}X^2$, to show the existence of K_4, a clique with 4 vertices, for a random graph in $G(n, p)$

1 Second Moment Method

Let $X : \Omega \to \mathbb{R}$ be a random variable. We are interested in the distribution of the values of X, however, very rarely we can learn all the information about the distribution. Usually it is easy to calculate first moments, e.g., $\mathbb{E}X, \mathbb{E}X^2$, and we show that from this limited information we can derive interesting properties.

Definition 1 *We say that an event A happens* (asymptotically) almost always *if* $\Pr[A] = 1 - o(1)$.

In the previous lecture we have seen that if X is a non-negative random variable and $\mathbb{E}X = o(1)$ then $X = 0$ almost always. We use the second moment method to derive a similar statement: If $\mathbb{E}X \to \infty$ and $Var(X)$ is not too big then $\Pr[X = 0] = o(1)$. To show this we use Chebychev's inequality: $\Pr\left[|X - \mathbb{E}X| > c\sigma\right] \le \frac{1}{c^2}$, and the fact that $\{\omega : X(\omega) = 0\} \subseteq \{\omega : |X(\omega) - \mathbb{E}X| \le \mathbb{E}X\}$ to conclude

$$\Pr[X = 0] \le \frac{Var(X)}{\mathbb{E}^2 X} \tag{1}$$

2 Existence of K_4 in $G(n, p)$

K_4 is a clique of 4 vertices. $G(n, p)$ is a probability space of all the graphs with $|V| = n$ vertices such that every edge is present with probability p independent of the other edges. We use the second moment method to prove the following:

Theorem 2 *The threshold function for the existence of K_4 in $G(n, p)$ is $p = n^{-2/3}$*

1. *If $p = o(n^{-2/3})$ then G does not contain K_4 almost always.*
2. *If $p = \omega(n^{-2/3})$ then G contains K_4 almost always.*

Proof

1. The first part is proved using the linearity of the expectation: Let $X : G(n,p) \to \mathbb{N}_+$ be a random variable counting the number of copies of K_4 in G, and we represent it as a sum of simple variables:

$$X = \sum_{|T|=4, T\subseteq[n]} X_T,$$

where X_T is an indicator to the event that T is a clique. $\mathbb{E}X = \binom{n}{4}p^6 \leq n^4p^6 = o(1)$ therefore $X = 0$ almost always, which implies that most of the graphs G with n vertices do not contain K_4.

2. To prove the second part we note that $\mathbb{E}X > \binom{n-4}{4}p^6 \to \infty$, thus to prove that $\Pr[X = 0] = o(1)$ we use the estimation in equation (**??**). To estimate $Var(X)$ we compute $\mathbb{E}X^2 - \mathbb{E}^2X$ while noting that for every two sets $T \neq S \subset [n]$ of size 4 the random variables X_T, X_S are independent unless $|S \cap T| = 2, 3$:

$$Var(X) \sum_{|T|=4, T\subseteq[n]} (\mathbb{E}X_T - \mathbb{E}^2X_T) + \underbrace{2\sum_{S\neq T}(\mathbb{E}(X_SX_T) - \mathbb{E}X_S\mathbb{E}X_T)}_{\text{vanishes unless } |S\cap T|=2,3}$$

- The term $\sum_{|T|=4,T\subseteq[n]}(\mathbb{E}X_T - \mathbb{E}^2X_T)$ is less than $\mathbb{E}X$
- $\sum_{|S\cap T|=2}\mathbb{E}(X_SX_T) = \binom{n}{4}\cdot\binom{4}{2}\cdot\binom{n-4}{2}p^{11}$, where $\binom{n}{4}$ are the number of possibilities to choose the set T, $\binom{4}{2}$ are the possibilities to choose 2 elements of S in the 4 elements of T, $\binom{n-4}{2}$ are the possible ways to choose the remaining elements of S and p^{11} is the probability that all 11 edges appear between the vertices in T and S.
- Similarly $\sum_{|S\cap T|=3}\mathbb{E}(X_SX_T) = \binom{n}{4}\cdot\binom{4}{3}\cdot\binom{n-4}{1}p^9$.
- $\mathbb{E}X_T = \mathbb{E}X_S = p^6$.

Therefore $Var(X) \leq \mathbb{E}X + n^6p^{11} + n^5p^9 = o(n^8p^{12}) = o(\mathbb{E}^2X)$.

■