22C:296 Seminar on Randomization

Scribe: Darren Schipper

October 14, 2003

In this class we will talk about the following.

1. $\frac{1}{2}$-way to *Turan's Theorem.*

2. All the way to Turan's Theorem

3. Introduction to the *second moment method*.

Consider two additional examples of the first moment method.

Question: Given a tree T with n vertices, what is the size of the largest independent set in T?
Answer: A largest independent set in T contains $\geqslant \frac{n}{2}$ vertices.
Proof Repeatedly pick a leaf, and throw it in the independent set being constructed, and delete the leaf and its unique neighbor from the tree. In each iteration two vertices are deleted from T and one vertex is added to the independent set being constructed. This shows that the algorithm will pick $\geqslant \frac{n}{2}$ vertices. □

Question: Given a planar graph with n vertices, what is the size of the largest independent set in the graph?
Answer: $\frac{n}{6}$. This follows in a similar way from the fact that any planar graph has a vertex with at most 5 neighbors.

To find the common theme in the preceding examples, consider the definition of a *d-degenrate graph.* This is a graph in which the vertices can be ordered as $v_1, v_2, \ldots v_n$ such that each v_i has at most d neighbors in the subgraph G$[\{v_1, v_2, \ldots v_{i-1}\}]$. For example, a tree is a 1-degenerate graph. Also, we have that a planar graph is a 5-degenerate graph, because by Euler's Theorem, any n-vertex planar graph has $\leqslant 3n-6$ edges. Hence, there exists a vertex of degree $\leqslant 5$, which implies that there exists an ordering of the vertices of any n-vertex planar graph as $v_1, v_2, \ldots v_n$ such that each v_i has $\leqslant 5$ neighbors in G$[\{v_1, v_2, \ldots v_{i-1}\}]$. All of this discussion points to the following result.

Theorem 1 *Any n-vertex d-degenerate graph has an independent set of size* $\geqslant \frac{n}{d+1}$

Seeking generalization, we observe that any n-vertex d-degenerate graph has $\leqslant nd$ edges, and ask whether the foregoing theorem holds for *any* n-vertex graph with $\leqslant nd$ edges.

1 $\frac{1}{2}$ -way to Turan's Theorem

Theorem 2 *Let G=(V,E) be a graph with n vertices and $\frac{nd}{2}$ edges. Then $\alpha(G) \geqslant \frac{n}{2d}$.*

For example, if $d = 2$, this implies that we are talking about graphs with n edges. In this case, the theorem implies the existence of an independent set of size $\geqslant \frac{n}{4}$.

Proof Let S be a random subset of vertices chosen as follows: visit each v and independently throw it in S with probability p (p to be fixed later).

Let X_v be an indicator random variable defined as $X_v = 1$ if $v \in S$ and $X_v = 0$ otherwise. Note that $\text{Prob}[X_v = 1] = p$ and hence $E[X_v] = p$. Let $X = \sum_{v \in V} X_v$. This implies that $E[X] = \sum_{v \in V} E[X_v] = np$. Now what is the expected number of edges in the induced subgraph $G[S]$? For any edge $e \in E(G)$, let Y_e be the indicator variable defined as $Y_e = 1$ if $e \in G[S]$ and $Y_e = 0$ otherwise. Note that $\text{Prob}[Y_e = 1] = p^2$ and hence $E[Y_e] = p^2$. Let $Y = \sum_{e \in E} Y_e$. This implies that $E[Y] = \sum_{e \in E} E[Y_e] = p^2 \cdot \frac{nd}{2}$. Hence, $E[X - Y] = np - np^2 \cdot \frac{d}{2}$. Now set $p = \frac{1}{d}$. Then $E[X - Y] = \frac{n}{d} - \frac{n}{2d} = \frac{n}{2d}$. This implies that there exists a subset $S \subseteq V(G)$ such that

$$(\#\,\text{vertices in}\,S) - (\#\,\text{edges in}\,G[S]) \geqslant \frac{n}{2d}.$$

So far we have only used the First Moment Method. We now employ the *Alteration Method* as follows: for each edge in $G[S]$, delete from S one of its endpoints. Let the resulting set be S^*. Hence, $|S^*| \geqslant \frac{n}{2d}$. Also, note that S^* is an independent set. □

2 All the way to Turan's Theorem

Theorem 3 *Let G=(V,E) be a graph with n vertices and $\frac{nd}{2}$ edges. Then $\alpha(G) \geqslant \frac{n}{d+1}$.*

Proof Choose a permutation of V uniformly at random. Let $<$ ("less than") be the total order on V induced by this permutation. Let $I = \{v \in V |\, \text{for every edge}(v, w), v < w\}$. Note that I is an independent set. Let X_v be an indicator variable such that $X_v = 1$ if $v \in I$ and $X_v = 0$ otherwise. Let d_v be the degree of v. Then $\text{Prob}[X_v = 1] \equiv \text{Prob}[v\,\text{is the}''\text{smallest}''\,\text{vertex in its neighborhood}] = \frac{1}{d_v+1}$. Let $X = \sum_{v \in V} X_v$. Clearly, $X = |I|$. Also note that $E[X] = \sum_{v \in V} E[X_v] = \sum_{v \in V} \frac{1}{d_v+1}$. This implies that there is an independent set I such that $|I| \geqslant \sum_{v \in V} \frac{1}{d_v+1}$. Now, note that $\sum_{v \in V} d_v = 2 \cdot \frac{nd}{2} = nd$. We have $|I| \geqslant \sum_{v \in V} \frac{1}{d_v+1}$. We minimize $\sum_{v \in V} \frac{1}{d_v+1}$ maintaining $\sum_{v \in V} d_v = nd$ to get $d_v = d$. Hence, $|I| \geqslant \sum_{v \in V} \frac{1}{d+1} = \frac{n}{d+1}$ □

3 The Second Moment Method

$E[X^2]$ is the *second moment* of a random variable X. The *variance* of a random variable X, denoted $\text{var}[X]$is defined as:

$$\text{var}[X] = E[(X - E[X])^2].$$

This can be simplified to

$$\begin{aligned} var[X] &= E[X^2 - 2 \cdot X \cdot E[X] + E[X]^2] \\ &= E[X^2] - 2 \cdot E[X]^2 + E[X]^2 \\ &= E[X^2] - E[X]^2 \end{aligned}$$

Chebyshev's Inequality

$\text{Prob}[|X - E[X]| \geqslant t] \leqslant \frac{\text{var}[X]}{t^2}$ (this is an example of a concentration result)

Proof $|X - E[X]| \geqslant t \equiv (X - E[X])^2 \geqslant t^2$

$\text{Prob}[|X - E[X]| \geqslant t] = \text{Prob}[(X - E[X])^2 \geqslant t^2] \leqslant \frac{E[(X-E[X])^2]}{t^2} = \frac{\text{var}[X]}{t^2}$ The last inequality above is by applying Markov's inequality. □