MTH742P: Advanced Combinatorics	21/11/2016

Lecture 8: Linearity of expectation

Lecturer: Heng Guo

1 Linearity of expectation

Now let us see some extensions of the basic method.

Theorem 1 (Linearity of expectation)**.** *Let $X_1, \cdots, X_n$ be random variables and $X = c_1X_1 + \cdots + c_nX_n$, where c_i's are constants. Then*

$$\mathbb{E}\, X = c_1\, \mathbb{E}\, X_1 + \cdots + c_n\, \mathbb{E}\, X_n.$$

Proof. We prove it by induction. The base case of $n = 1$ is trivial.

For the inductive step, it is sufficient to show that $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$ for two random variables X and Y. Indeed,

$$\begin{aligned}
\mathbb{E}[X + Y] &= \sum_{x,y} \Pr(X = x, Y = y)(x + y) \\
&= \sum_{x,y} \Pr(X = x, Y = y)x + \sum_{x,y} \Pr(X = x, Y = y)y \\
&= \sum_{x} x \sum_{y} \Pr(X = x, Y = y) + \sum_{y} y \sum_{x} \Pr(X = x, Y = y) \\
&= \sum_{x} x \Pr(X = x) + \sum_{y} y \Pr(Y = y) \\
&= \mathbb{E}\, X + \mathbb{E}\, Y,
\end{aligned}$$

where the summation over x and y are over all possible values of them. To finish the proof, let $X' = \sum_{i=1}^{n-1} c_iX_i$ and $Y' = c_nX_n$ and apply the equation above. □

To apply the result, we often use the fact that there must exist a point in the probability space such that $X \geq \mathbb{E}\, X$ or $X \leq \mathbb{E}\, X$.

Theorem 1 is rather simple but it turns out to be surprisingly strong, mainly because that we have no requirement on the dependence X_i.

Let us do a warm up. Let σ be a random permutation on $\{1, 2, \cdots, n\}$, uniformly chosen. Let $X(\sigma)$ be the number of fixed points; that is $i = \sigma(i)$. Define X_i to be the indicator variable of the event $i = \sigma(i)$. Then $X(\sigma) = \sum_{i=1}^{n} X_i$. It is easy to see that

$$\mathbb{E}\, X_i = \Pr(i = \sigma(i)) = \frac{(n-1)!}{n!} = \frac{1}{n}.$$

Thus by Theorem 1,

$$\mathbb{E}\, X = \sum_{i=1}^{n} X_i = 1.$$

1.1 Hamiltonian paths in tournaments

The following result is often considered the first use of the probabilistic method. It is due to Szele (1943).

Theorem 2. *There is a tournament T with n players and at least $\frac{2n!}{2^n}$ Hamiltonian paths.*

Proof of Theorem 2. We still randomize every edge uniformly to get a random tournament. Let X be the number of Hamiltonian paths in such a tournament. For each permutation σ, let X_σ be the indicator variable for the event that σ gives a Hamiltonian path. It is easy to see that X_σ is 1 if and only if $(\sigma(i), \sigma(i+1))$ is oriented this way for every $1 \le i < n$. Each individual orientation happens with probability $1/2$ and they are independent. As there are $n-1$ edges on a Hamiltonian path, we have that

$$\mathbb{E}\, X_\sigma = \Pr(X_\sigma = 1) \cdot 1 + \Pr(X_\sigma = 0) \cdot 0 = \Pr(X_\sigma = 1) = \frac{1}{2^{n-1}}.$$

In addition, there are $n!$ permutations. Therefore by Theorem 1,

$$\mathbb{E}\, X = \sum_{\sigma} \mathbb{E}\, X_\sigma = \frac{2n!}{2^n}.$$

It implies that there must exist one tournament with at least $\mathbb{E}\, X$ Hamiltonian paths. □

Note that by Stirling's approximation, $n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n$, implying that $\frac{2n!}{2^n}$ grows (super) exponentially in n.

Szele also conjectured that the maximum possible number of Hamiltonian paths is at most $\frac{n!}{(2-o(1))^n}$. This is proved by Alon in 1990.

As shown in the examples, the basic argument involving expectations goes as follows. We want to construct an object that maximizes (or minimizes) a certain parameter X.

1. First define a random object.

2. Calculate the expectation of $\mathbb{E}\, X$. To do so, we usually decompose X into atomic indicator variables X_i such that $X = \sum_{i=1}^{N} X_i$ and apply linearity of expectation.

3. We conclude that there must exist an object with $X \ge \mathbb{E}\, X$ (or $X \le \mathbb{E}\, X$).

2 Splitting Graphs

Theorem 3. *Let $G = (V, E)$ be a graph with n vertices and m edges. Then G contains a bipartite subgraph with at least $m/2$ edges.*

Proof. For every $v \in V$, we choose it with probability $1/2$ and independently. This yields a random subset $L \subset V$ and let $R = V \setminus L$. Let X denote the number of "crossing" edges which are between L and R.

We shall decompose $X = \sum_{uv \in E} X_{uv}$, where X_{uv} is the indicator variable of the event that the edge uv is a crossing edge. For any such edge uv, it is crossing in two cases: $u \in L$, $v \in R$ or $u \in R$, $v \in L$. Thus,

$$\begin{aligned}\Pr(X_{uv} = 1) &= \Pr(u \in L,\ v \in R) + \Pr(u \in R,\ v \in L)\\ &= 1/4 + 1/4 = 1/2.\end{aligned}$$

In other words,

$$\mathbb{E}\, X_{uv} = \Pr(X_{uv} = 1) \cdot 1 + \Pr(X_{uv} = 0) \cdot 0 = 1/2.$$

By Theorem 1,

$$\mathbb{E}\, X = \sum_{uv \in E} X_{uv} = \frac{m}{2}.$$

Hence, there must exist some L such that the number of crossing edges is at least $m/2$. $\square$

If G is a complete graph K_{2n}, then the largest bipartite subgraph is to split V evenly. The number of edges, in this case, is $m' = n^2$ whereas $m = \frac{2n(2n-1)}{2}$. Thus $\frac{m'}{m} = \frac{n}{2n-1} > 1/2$. We can actually also achieve this slightly better bound by considering a more subtle probability space.

Theorem 4. *If $G = (V, E)$ is a graph with $2n$ vertices and m edges, then G contains a bipartite subgraph with at least $\frac{n}{2n-1} \cdot m$ edges.*

If $G = (V, E)$ is a graph with $2n + 1$ vertices and m edges, then G contains a bipartite subgraph with at least $\frac{n+1}{2n+1} \cdot m$ edges.

Proof. Suppose G has $2n$ vertices. Instead of choosing every vertex independently and uniformly at random, we choose L from all subsets of size n uniformly at random. In other words, we choose every possible L with probability $\frac{1}{\binom{2n}{n}}$. Define X in the same way as in the proof of Theorem 3. We do the same decomposition as well; that is $X = \sum_{uv \in E} X_{uv}$.

The difference is the calculation of $\Pr(X_{uv} = 1)$. In fact, the probability that $u \in L$ for any vertex u is still

$$\Pr(u \in L) = \frac{\binom{2n-1}{n-1}}{\binom{2n}{n}} = \frac{(2n-1)!}{(n-1)!n!} \cdot \frac{n!n!}{2n!} = \frac{n}{2n} = \frac{1}{2}.$$

However, the probability that $u \in L$ and $v \in R$ is changed. In fact,

$$\Pr(u \in L,\ v \in R) = \frac{\binom{2n-2}{n-1}}{\binom{2n}{n}} = \frac{(2n-2)!}{(n-1)!(n-1)!} \cdot \frac{n!n!}{2n!} = \frac{n^2}{2n(2n-1)} = \frac{n}{2(2n-1)}.$$

Here the term $\binom{2n-2}{n-1}$ is the number of subsets of vertices which contains u but not v. As before, there are two possibilities for the edge uv to be crossing. Thus,

$$\begin{aligned}\Pr(X_{uv} = 1) &= \Pr(u \in L,\ v \in R) + \Pr(v \in R,\ u \in L) \\ &= \frac{n}{2n-1}.\end{aligned}$$

Therefore, by Theorem 1, we have that

$$\mathbb{E}\, X = \sum_{uv \in E} \mathbb{E}\, X_{uv} = \frac{n}{2n-1} \cdot m.$$

Hence, there must exist some X that is larger than $\mathbb{E}\, X$. This finishes the proof of the $2n$ case.

For the case of $2n+1$, we still choose L as a random subset of size n. In this case, $R = V \setminus L$ has $n+1$ many vertices. The only difference from the above is that for any $uv \in E$,

$$\Pr(u \in L,\ v \in R) = \frac{\binom{2n-1}{n-1}}{\binom{2n+1}{n}} = \frac{(2n-1)!}{(n-1)!n!} \cdot \frac{n!(n+1)!}{(2n+1)!} = \frac{n(n+1)}{2n(2n+1)} = \frac{n+1}{2(2n+1)}.$$

Hence,

$$\begin{aligned}\Pr(X_{uv} = 1) &= \Pr(u \in L,\ v \in R) + \Pr(v \in R,\ u \in L) \\ &= \frac{n+1}{2n+1}.\end{aligned}$$

The rest of the argument is exactly the same. □

3 Unbalancing Lights

The next theorem has an amusing interpretation. Suppose we have an $n \times n$ array of lights, either switched on or off. Suppose that we have a switch for each row (or each column) so that it toggles all lights in this row (or column). The question is, given any initial configuration, what is the maximum number of lights on by pulling these switches for rows and columns?

We will formalize this problem as follows. We associate each light at (i, j) with $a_{ij} = \pm 1$, where, say, $+1$ means that the light is on. If we pull a switch of row i, then we set $x_i = -1$ and otherwise let $x_i = +1$. Similarly, if we pull a switch of column j, then we set $y_j = -1$ and otherwise let $x_i = +1$. The final state of a light is $a_{ij} x_i y_j$. The quantity

$$S := \sum_{i \in [n]} \sum_{j \in [n]} a_{ij} x_i y_j \tag{1}$$

is equal to the difference between lights on and off. Thus, to maximize the number of lights on, it is equivalent to maximize S.

Theorem 5. *Let $a_{ij} = \pm 1$ for $1 \le i, j \le n$. There exists $x_i, y_j = \pm 1$, $1 \le i, j \le n$ such that*

$$S \ge \left(\sqrt{\frac{2}{\pi}} + o(1) \right) n^{3/2},$$

where S is defined in (1).

It is easy to see that if we choose x_i and y_j uniformly at random, then $\mathbb{E}\, S = 0$. Thus, we need to be more clever when picking x_i and y_j.

Proof. We will choose $y_1, \cdots, y_n = \pm 1$ independently and uniformly at random. Let

$$R_i := \sum_{j=1}^{n} a_{ij} y_j.$$

Notice that y_j is uniformly at random. Regardless of a_{ij} being $+1$ or -1, $a_{ij} y_j$ is uniformly at random. We still have $\mathbb{E}\, R_i = 0$. The reason for this is that R_i can be positive and negative where these values cancel out. However, recall that we still have the freedom to choose x_i so that it has the same sign as R_i. Thus, we should calculate the expectation of $|R_i|$.

Let X_t be a uniform ± 1 and $X = \sum_{t=1}^{n} X_t$. Thus $\mathbb{E}\,|X| = \mathbb{E}\,|R_i|$ for any $i \in [n]$. We claim that

$$\mathbb{E}\,|X| = n 2^{1-n} \binom{n-1}{\lfloor (n-1)/2 \rfloor}.$$

This is because, if the number of -1 is $k < n/2$ from all X_i, then $X = n - k - k = n - 2k$.

If the number of $+1$ is $k < n/2$, we also have $|X| = |k - (n-k)| = n - 2k$. Thus,

$$\begin{aligned}
\mathbb{E}\,|X| &= \frac{2}{2^n} \sum_{k=0}^{\lfloor (n-1)/2 \rfloor} \binom{n}{k}(n-2k) \\
&= 2^{1-n} \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} \left(\binom{n}{k}(n-k) - \binom{n}{k}k \right) + 2^{1-n}n \\
&= 2^{1-n} \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} \left(\frac{n!}{k!(n-k)!} \cdot (n-k) - \frac{n!}{k!(n-k)!} \cdot k \right) + 2^{1-n}n \\
&= 2^{1-n} \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} \left(\frac{n!}{k!(n-k-1)!} - \frac{n!}{(k-1)!(n-k)!} \right) + 2^{1-n}n \\
&= 2^{1-n} \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} \left(n\binom{n-1}{k} - n\binom{n-1}{k-1} \right) + 2^{1-n}n \\
&= n2^{1-n} \sum_{k=1}^{\lfloor (n-1)/2 \rfloor} \left(\binom{n-1}{k} - \binom{n-1}{k-1} \right) + 2^{1-n}n \\
&= n2^{1-n} \left(\binom{n-1}{\lfloor (n-1)/2 \rfloor} - 1 \right) + 2^{1-n}n \\
&= n2^{1-n} \binom{n-1}{\lfloor (n-1)/2 \rfloor}.
\end{aligned}$$

The claim holds.

Using the claim and Stirling's approximation ($n! = (1+o(1))\sqrt{2\pi n}\left(\frac{n}{e}\right)^n$), we have that

$$\begin{aligned}
\mathbb{E}\,|R_i| = \mathbb{E}\,|X| &= n2^{1-n} \binom{n-1}{\lfloor (n-1)/2 \rfloor} \\
&= \frac{2n}{2^n} \cdot \frac{(n-1)!}{\lfloor (n-1)/2 \rfloor! \lceil (n-1)/2 \rceil!} \\
&= (1+o(1)) \cdot \frac{2n}{2^n} \cdot \frac{\sqrt{2\pi n}\left(\frac{n-1}{e}\right)^{n-1}}{\pi n \left(\frac{n-1}{2e}\right)^{n-1}} \\
&= \left(\sqrt{\frac{2}{\pi}} + o(1) \right) \sqrt{n}.
\end{aligned}$$

Now let $R = \sum_{i=1}^n |R_i|$. We apply the linearity of expectation, Theorem 1,

$$\mathbb{E}\,R = \sum_{i=1}^n \mathbb{E}\,R_i = \left(\sqrt{\frac{2}{\pi}} + o(1) \right) n^{3/2}.$$

Thus, there must exists $y_1, \cdots, y_n = \pm 1$ such that R is at least this value. Finally, we pick x_i as the same sign as R_i. In this case,

$$
\begin{aligned}
S &= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i y_j = \sum_{i=1}^{n} x_i \sum_{j=1}^{n} a_{ij} y_j \\
&= \sum_{i=1}^{n} x_i R_i = \sum_{i=1}^{n} |R_i| \\
&= R \geq \left(\sqrt{\frac{2}{\pi}} + o(1) \right) n^{3/2}.
\end{aligned}
$$

This finishes the proof. $\square$

In fact, the order of magnitude $n^{3/2}$ cannot be improved in the bound above.