Lecture 31: Basic probability theory

Nitin Saxena *

IIT Kanpur

Theorem 1 (Chernoff, 1952). *Let X be a random variable which takes value 1 with probability p and 0 otherwise. Let $X_1, X_2, \cdots, X_n$ correspond to random variable X measured n times (the experiment is independently repeated n times). Define $S = \sum_{i=1}^{n} X_i$, then (for any $\delta \in (0,1)$)*

$$P(\, S < (1-\delta) \cdot n \cdot E[X]\,) \leq e^{-nE[X]\delta^2/2}\,.$$

Note 1. We have taken a very special form of random variable X, but it can be generalized.

Proof. (This proof is taken from John Canny's lecture notes, `http://www.cs.berkeley.edu/~jfc/cs174/lecs/lec10/lec10.pdf`.)

The proof of Chernoff bound follows by looking at the random variable e^{-tS}, where t is a parameter and will be optimized later. Define $u := E[S] = nE[X]$, so

$$P(S < (1-\delta)u) = P(e^{-tS} > e^{-t(1-\delta)u})\,.$$

We can apply Markov's inequality for e^{-tS},

$$P(S < (1-\delta)u) \leq \frac{E[e^{-tS}]}{e^{-t(1-\delta)u}}\,.$$

But e^{-tS} is the product of e^{-tX_i}, where X_i are independent. So,

$$P(S < (1-\delta)u) \leq \frac{\Pi_{i=1}^{n} E[e^{-tX_i}]}{e^{-t(1-\delta)u}}. \tag{1}$$

Exercise 1. Show that $E[e^{-tX_i}] = 1 - p(1-e^{-t}) \leq e^{p(e^{-t}-1)}$.

Use the inequality $1 - x \leq e^{-x}$.

The above exercise implies that $\Pi_{i=1}^{n} E[e^{-tX_i}] \leq e^{u(e^{-t}-1)}$. From Eq. 1, we get

$$P(S < (1-\delta)u) \leq e^{u(e^{-t}+t(1-\delta)-1)}\,.$$

Exercise 2. Show that the bound on the right is minimized for $t = \ln \frac{1}{1-\delta}$.

Putting the best t, we get

$$P(S < (1-\delta)u) \;\leq\; \left(\frac{e^{-\delta u}}{(1-\delta)^{u(1-\delta)}}\right)\,.$$

Using the Taylor expansion of $\ln(1-\delta)$,

$$P(S < (1-\delta)u) \;\leq\; e^{-u\delta^2/2}\,.$$

Hence proved.

□

Exercise 3. Similarly, show that $P(\, S > (1+\delta) \cdot n \cdot E[X]\,) \leq e^{-nE[X]\delta^2/3}$.

* Edited from Rajat Mittal's notes.

1 Probabilistic methods

Now we will see examples of probabilistic methods. This is used to prove the existence of a *good* structure using probability theory. We will define a probability distribution over the set of structures. Then we prove that the good event happens with positive probability, which implies that a good structure exists.

These ideas are best illustrated with the help of applications.

1.1 Ramsey numbers

Previously in class we proved that if we color the edges of K_6 using blue or red, then either there is a blue K_3 or a red K_3 as a subgraph. Here K_n is the complete graph (every pair of vertices are connected) on n vertices.

We can generalize the above concept and ask, are there complete graphs for which any 2-coloring (of the edges) gives rise to either a blue K_k or a red K_ℓ. It has been shown that there always exists n, s.t., any two-coloring of K_n will have a monochromatic blue K_k or a monochromatic red K_ℓ. The smallest such number n is called the *Ramsey number* $R(k, \ell)$.

It has been a big open question to find out the bounds on $R(k, \ell)$. We will use probabilistic method to give a *lower bound* on the diagonal Ramsey number $R(k, k)$.

Call an edge coloring of K_n *good*, if there are no monochromatic K_k's.

The idea would be to randomly color the edges of the graph K_n. If there is a positive probability (over the random coloring) that none of the K_k subgraphs are monochromatic red or blue, then there exist a coloring which is good.

We color every edge either red or blue independently with probability $1/2$. There are in total $\binom{n}{k}$ subgraphs K_k for a K_n.

Exercise 4. A particular subgraph K_k is monochromatic with probability $2^{1-\binom{k}{2}}$.

There are $\binom{k}{2}$ edges and the first one could be of any color.

We have already proved that,

$$P(\cup_{i=1}^m E_i) \leq \sum_{i=1}^m P(E_i) \quad \text{[Union bound]} \ .$$

So the total probability that some K_k is monochromatic is at most $\binom{n}{k} \cdot 2^{1-\binom{k}{2}}$. If this probability is less than 1, then there is a positive probability that none of the K_k's are monochromatic.

Since the probability was over random coloring, there exists a good coloring (such that no K_k is monochromatic).

Theorem 2. *If* $\binom{n}{k} \cdot 2^{1-\binom{k}{2}} < 1$ *then* $R(k, k) > n$.

To get an explicit lower bound, you can check that $n = \lceil 2^{k/2} \rceil$ will satisfy the above equation.

The essential argument in the above proof is that the number of colorings are much higher than the total number of graphs which have monochromatic K_k.

A counting argument for the above theorem can also be constructed. Actually, in all our applications, a counting argument can always be given. But the probabilistic argument in general is much simpler and easier to construct.

1.2 Probabilistic algorithm for construction

One of the important thing to notice in a probabilistic method of proofs is that the proofs are *non-constructive*. In the previous example, we were only able to show the existence of a good coloring. This proof does not construct the required coloring and hence is called non-constructive.

But suppose we choose n to be $\frac{1}{2}\left\lceil 2^{k/2}\right\rceil$. Then the probability of having a monochromatic K_k is very small. This shows that most of the random colorings will be good colorings.

This suggests a randomized algorithm. We take K_n and color the edges randomly. Because of the argument above, with high probability we will get a good coloring.

1.3 Sum-free subsets

Let us take another example. Given a set of integers S, $S+S$ is defined as the subset of integers which contain all possible sums of pair of elements in S,

$$S+S := \{t : t = s_1 + s_2,\ s_1, s_2 \in S\}.$$

A set S is called *sum-free* if S does not contain any element of $S+S$.

Exercise 5. Construct a set of 10 elements which is sum-free. Construct a set of n elements which is sum-free.

Eg. you can take a sequence of rapidly growing integers.

Using probabilistic method, we will show that every subset of integers contains a large sum-free subset.

Theorem 3. *For any subset S of n non-zero integers, There exists a subset of S which is sum-free and has size more than $n/3$.*

References

1. N. Alon and J. H. Spencer. The Probabilistic Method. *Wiley*, 2008.
2. H. Tijms Understanding Probability. *Cambridge University Press*, 2012.
3. D. Stirzaker. Elementary Probability. *Cambridge University Press*, 2003.
4. U. Schöning. Gems of Theoretical Computer Science. *Springer-Verlag*, 1998.