TTIC/CMSC 31150 Mathematical Toolkit **Spring 2013**

Lecture 3: April 8, 2013

Madhur Tulsiani Scribe: Abhishek Sen

1 Inequalities

We will develop some inequalities which let us bound the probability of a random variable taking a value very far from its expectation.

1.1 Markov's Inequality

This is the most basic inequality we will use. This is useful if the only thing we know about a random variable is its expectation. It will also be useful to derive other inequalities later.

Lemma 1.1 (Markov's Inequality) *Let Z be non-negative variable. Then,*

$$\mathbb{P}[Z \geq k] \leq \frac{\mathbb{E}[Z]}{k}. \tag{1}$$

Proof: This proof assumes that Z is a discrete random variable. However, the result is also valid if Z is a continuous random variable. We have,

$$\mathbb{E}[Z] = \sum_{i<k} \mathbb{P}[Z=i] \cdot i + \sum_{i \geq k} \mathbb{P}[Z=i] \cdot i \geq k \cdot \mathbb{P}[Z \geq k]$$

$$\left(\because \sum_{i<k} \mathbb{P}[Z=i] \cdot i \geq 0 \text{ and } \sum_{i \geq k} \mathbb{P}[Z=i] \cdot i \geq k \cdot \mathbb{P}[Z \geq k] \right)$$

So we have,

$$\mathbb{P}[Z \geq k] \leq \frac{\mathbb{E}[Z]}{k}.$$

∎

1.2 Chebyshev's Inequality

Lemma 1.2 (Chebyshev's inequality) *Let Z be a random variable and let $\mu = \mathbb{E}[Z]$. Then,*

$$\mathbb{P}[|Z-\mu| \geq t] \leq \frac{\mathbb{E}\left[(Z-\mu)^2\right]}{t^2} \tag{2}$$

Proof: Consider the non-negative random variable $(Z-\mu)^2$. Applying Markov's inequality we have

$$\mathbb{P}\left[(Z-\mu)^2 \geq k\right] \leq \frac{\mathbb{E}\left[(Z-\mu)^2\right]}{k}$$

Putting $k = t^2$ we have

$$\mathbb{P}[|Z-\mu| \geq t] \leq \frac{\mathbb{E}\left[(Z-\mu)^2\right]}{t^2}$$

■

Definition 1.3 *We define* ***variance*** *of a random variable X as*

$$\mathsf{Var}\left[X\right] = \mathbb{E}\left[(X-\mathbb{E}\left[X\right])^2\right] = \mathbb{E}\left[X^2\right] - (\mathbb{E}\left[X\right])^2$$

This definition allows us to write Chebyshev's Inequality as,

$$\mathbb{P}\left[|Z-\mu| \geq t\right] \;\; \leq \;\; \frac{\mathsf{Var}\left[Z\right]}{t^2}\,. \tag{3}$$

Definition 1.4 *We define the* ***covariance*** *between two random variables X_i, X_j as*

$$\mathsf{Cov}\left[X_i, X_j\right] = \mathbb{E}\left[X_i X_j\right] - \mathbb{E}\left[X_i\right]\mathbb{E}\left[X_j\right]\,.$$

In the following sections we will illustrate the applications of these inequalities in some problems.

2 Coin Tosses

An unbiased coin is tossed n times. Probability that head shows up in each toss is $\frac{1}{2}$. Let Z be a random variable for the number of heads that have showed up after n tosses. We also have random variables X for i^{th} coin toss, where $X_i = 1$ if head shows up in i^{th} toss and 0 otherwise.
So we have

$$Z = \sum_{i=1}^{n} X_i \quad \text{and} \quad \mathbb{E}\left[Z\right] = \sum_{i=1}^{n} \mathbb{E}\left[X_i\right] = \frac{n}{2}\,.$$

Let us now compare the kind of bounds we get using Markov's and Chebyshev's inequalities.

2.1 Application of Markov's inequality

Using Markov's inequality we have,

$$\mathbb{P}\left[Z \geq \frac{3n}{4}\right] \leq \frac{\mathbb{E}\left[Z\right]}{(3n/4)} \quad \Rightarrow \quad \mathbb{P}\left[Z \geq \frac{3n}{4}\right] \leq \frac{2}{3} \quad \Rightarrow \quad \mathbb{P}\left[Z - \frac{n}{2} \geq \frac{n}{4}\right] \leq \frac{2}{3}\,.$$

2.2 Application of Chebyshev's inequality

We want to show that Chebyshev's inequality gives a stronger bound on probability. For this we need to calculate the variance of Z. We do this calculation below in a way that applies in many other situations as well. We have,

$$\mathsf{Var}[Z] \;=\; \mathbb{E}\left[Z^2\right] - (\mathbb{E}[Z])^2$$

We observe that,

$$\mathbb{E}\left[Z^2\right] \;=\; \mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)^2\right] \;=\; \mathbb{E}\left[\sum_{i,j} X_i X_j\right] \;=\; \sum_{i,j} \mathbb{E}[X_i X_j]\,.$$

Similarly,

$$(\mathbb{E}[Z])^2 \;=\; \left(\mathbb{E}\left[\left(\sum_{i=1}^{n} X_i\right)\right]\right)^2 \;=\; \sum_{i,j} \mathbb{E}[X_i]\,\mathbb{E}[X_j]$$

So we have,

$$\begin{aligned}\mathsf{Var}[Z] &= \sum_{i,j} \mathbb{E}[X_i X_j] - \sum_{i,j} \mathbb{E}[X_i]\,\mathbb{E}[X_j] \\ &= \sum_{i} \left(\mathbb{E}\left[X_i^2\right] - (\mathbb{E}[X_i])^2\right) + \sum_{i \neq j} \left(\mathbb{E}[X_i, X_j] - \mathbb{E}[X_i]\,\mathbb{E}[X_j]\right) \\ &= \sum_{i} \mathsf{Var}[X_i] + \sum_{i \neq j} \mathsf{Cov}[X_i, X_j]\end{aligned}$$

Since the coin tosses are independent, we have $\mathbb{E}[X_i X_j] = \mathbb{E}[X_i]\,\mathbb{E}[X_j]$ and hence $\mathsf{Cov}[X_i, X_j] = 0$. This yields,

$$\mathsf{Var}[Z] \;=\; \sum_{i} \mathsf{Var}[X_i] \quad \text{for independent random variables } X_i\,. \tag{4}$$

Also $\mathsf{Var}[X_i] = \mathbb{E}\left[X_i^2\right] - (\mathbb{E}[X_i])^2 = p - p^2$, where $p = \mathbb{P}[X_i = 1]$ Here $p = \frac{1}{2}$, so $\mathsf{Var}[X_i] = \frac{1}{4}$ and hence, $\mathsf{Var}[Z] = \frac{n}{4}$. Applying Chebyshev's inequality we have,

$$\mathbb{P}\left[\left|Z - \frac{n}{2}\right| \geq t\right] \;\leq\; \frac{n}{4t^2}\,.$$

Setting $t = n/2$ and $t = \sqrt{n}$, gives the following bounds

$$\mathbb{P}\left[\left|Z - \frac{n}{2}\right| \geq \frac{n}{4}\right] \;\leq\; \frac{4}{n} \quad \text{and} \quad \mathbb{P}\left[\left|Z - \frac{n}{2}\right| \geq \sqrt{n}\right] \;\leq\; \frac{1}{4}$$

Thus, Chebyshev's inequality gives a much stronger bound on a deviation of $n/4$ from the mean, and can also bound the probability of deviations as small as $\sqrt{n}$. In particular, it gives a non-trivial bound whenever the deviation is larger than $\sqrt{\mathsf{Var}[Z]}$, a quantity which is referred to as the *standard deviation* of the random variable Z.

3 Max-cut

Given a graph $G = (V, E)$, where $|V| = n$ and $|E| = m$, we want to divide V into disjoint partitions S and $\overline{S}$, so that the number of edge crossings are maximized. For each vertex $i \in V$, we assign i to S with probability $\frac{1}{2}$, and to $\overline{S}$ with probability $\frac{1}{2}$.

Let Z be a random variable representing the number of edges cut and for each edge e we define a random variable X_e which is 1 if e is cut and 0 otherwise.

So we have

$$Z \;=\; \sum_e X_e \quad \text{and} \quad \mathbb{E}\left[Z\right] \;=\; \sum_e \mathbb{E}\left[X_e\right] = \frac{m}{2}$$

We want to show that Z is large with good probability.

3.1 Application of Markov's Inequality

We define a non-negative random variable $Y = m - Z$. We observe that $\mathbb{E}\left[Y\right] = \frac{m}{2}$. Using the Markov's Inequality, we have

$$\mathbb{P}\left[Z \leq \frac{m}{3}\right] \;=\; \mathbb{P}\left[Y \;\geq\; \frac{2m}{3}\right] \;\leq\; \frac{m/2}{2m/3} \;=\; \frac{3}{4}\,.$$

Hence, Markov's inequality gives that Z is at least $m/3$ with probability at least $1/4$.

3.2 Application of Chebyshev's Inequality

We need to calculate $\mathsf{Var}\left[Z\right]$ to use Chebyshev's Inequality. As before, we can write

$$\mathsf{Var}\left[Z\right] \;=\; \mathsf{Var}\left[\sum_e X_e\right] \;=\; \sum_e \mathsf{Var}\left[X_e\right] + \sum_{e_1 \neq e_2} \mathsf{Cov}\left[X_{e_1}, X_{e_2}\right]\,.$$

We now analyze $\mathsf{Cov}\left[X_{e_1}, X_{e_2}\right]$ for two different edges e_1 and e_2. There are two cases:

- **Case 1: e_1 and e_2 don't share a vertex.** In this case the variables X_{e_1} and X_{e_2} are independent and hence $\mathsf{Cov}\left[X_{e_1}, X_{e_2}\right] = 0$.

- **Case 2: e_1 and e_2 share one vertex.** If we assign labels s or $\overline{s}$ to each of the vertices of e_1 and e_2, there are a total of $2^3 = 8$ assignments, since there are 3 vertices (one vertex is shared by the two edges). Of these 8 assignments, only 2 assignments correspond to both the edges being cut. Thus, we still have,

  $$\mathbb{E}\left[X_{e_1} X_{e_2}\right] \;=\; \frac{1}{4} \;=\; \mathbb{E}\left[X_{e_1}\right] \cdot \mathbb{E}\left[X_{e_2}\right]\,.$$

  Hence, in this case we *also* have that X_{e_1} and X_{e_2} are independent and $\mathsf{Cov}\left[X_{e_1}, X_{e_2}\right] = 0$.

This gives us,

$$\mathsf{Var}[Z] \;=\; \sum_e \mathsf{Var}[X_e] \;=\; \frac{m}{4}\,.$$

Now applying Chebyshev's Inequality, we have

$$\mathbb{P}\left[\left|Z - \frac{m}{2}\right| \geq t\right] \;\leq\; \frac{\mathsf{Var}[Z]}{t^2} = \frac{m}{4t^2}\,.$$

Taking $t = m/6$ and $t = \sqrt{m}$ gives,

$$\mathbb{P}\left[\left|Z - \frac{m}{2}\right| \geq \frac{m}{6}\right] \leq \frac{9}{m} \quad \text{and} \quad \mathbb{P}\left[\left|Z - \frac{m}{2}\right| \geq \sqrt{m}\right] \leq \frac{1}{4}\,.$$

Thus, again we get a much better bound on the probability that the cut is at least $m/3$, and can even say that the cut is at least $(m/2) - \sqrt{m}$ with probability at least 3/4.

4 Weak Law of Large Numbers

Suppose we have a finite space $\mathcal{X}$ and some function $f : \mathcal{X} \to \mathbb{R}$. We can estimate $\mathbb{E}_{x\in\mathcal{X}}[f(x)]$ by picking k independent samples $x_1, \ldots, x_k \in \mathcal{X}$ and computing the *empirical average*

$$\overline{\mu}_k \;=\; \frac{1}{k} \cdot \sum_{i=1}^{k} f(x_i)\,.$$

The weak law of large numbers says that empirical average converges to the true average as k tends to infinity.

Lemma 4.1 (Weak Law of Large Numbers) *Let μ_k be the empirical average corresponding to k independent random samples from $\mathcal{X}$. Then, for any $\varepsilon > 0$,*

$$\lim_{k\to\infty} \mathbb{P}\left[\left|\overline{\mu}_k - \mathbb{E}[f]\right| \geq \varepsilon\right] \;=\; 0\,.$$

Proof: Note that $\overline{\mu}_k$ is also a random variable, and

$$\mathbb{E}[\overline{\mu}_k] \;=\; \mathop{\mathbb{E}}_{x_1,\ldots,x_k}\left[\frac{1}{k} \cdot \sum_{i=1}^{k} f(x_i)\right] \;=\; \frac{1}{k} \cdot \sum_{i=1}^{k} \mathop{\mathbb{E}}_{x_i}[f(x_i)] \;=\; \mathbb{E}[f]$$

Applying Chebyshev's Inequality to $\overline{\mu}_k$, we have

$$\mathbb{P}\left[\left|\overline{\mu}_k - \mathbb{E}[f]\right| \geq \varepsilon\right] \;\leq\; \frac{\mathsf{Var}[\overline{\mu}_k]}{\varepsilon^2}\,.$$

It remains to compute $\mathsf{Var}[\overline{\mu}_k]$. Since $\overline{\mu}_k$ is the sum of k independent random variables, we have

$$\mathsf{Var}[\overline{\mu}_k] \;=\; \mathsf{Var}\left[\sum_{i=1}^{k} \frac{1}{k} \cdot f(x_i)\right] \;=\; \sum_{i=1}^{k} \mathsf{Var}\left[\frac{1}{k} \cdot f(x_i)\right] \;=\; k \cdot \frac{1}{k^2} \cdot \mathsf{Var}[f] \;=\; \frac{1}{k} \cdot \mathsf{Var}[f]\,.$$

Substituting the value of $\mathsf{Var}[\overline{\mu}_k]$ in the above inequality, we get

$$\mathbb{P}\left[\left|\overline{\mu}_k - \mathbb{E}[f]\right| \geq \varepsilon\right] \;\leq\; \frac{\mathsf{Var}[k]}{k \cdot \varepsilon^2}\,.$$

Since $\mathcal{X}$ is a finite space, $\mathsf{Var}[f]$ is a finite quantity. Hence, $\lim_{k\to\infty} \mathbb{P}\left[\left|\overline{\mu}_k - \mathbb{E}[f]\right| \geq \varepsilon\right] = 0$. ■

5 Threshold Phenomena in Random Graphs

We consider a model of Random Graphs by Erdős and Rényi [ER60]. To generate a random graph with n vertices, for every pair of vertices $(i, j) \in V$, we put an edge independently with probability p. This model is denoted by $G_{n,p}$.

Let G be a random $G_{n,p}$ graph and let H be any fixed graph (on some constant number of vertices independent of n). We will be interested in understanding the probability that G contains a copy of H. We start by computing this when H is K_4, the clique on 4 vertices.

Definition 5.1 *We define k-**clique** to be a fully connected graph with k vertices.*

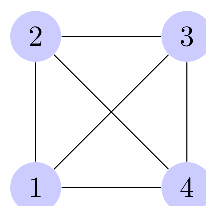

Figure 1: 4-Clique

As a convention, we will count a permutation of a copy of K_4 as the *same* copy. We define the random variable

$$Z \;=\; \text{number of copies of } K_4 \text{ in } G \;=\; \sum_C X_C\,.$$

where C ranges over all subsets of V of size 4 and the random variable X_C is defined as

$$X_C = \begin{cases} 1 & \text{if all pair of vertices in the set } C \text{ have an edge in between them} \\ 0 & \text{otherwise} \end{cases}$$

We have $\mathbb{E}[X_C] = p^6$, since the probability of connecting all 4 vertices (using 6 edges) in the 4-tuple is p^6. So we have the expectation of Z :

$$\mathbb{E}[Z] \;=\; \sum_C \mathbb{E}[X_C] \;=\; \binom{n}{4} \cdot p^6$$

We observe that,

$$\mathbb{E}[Z] \to 0 \;\text{ when } p \ll n^{-2/3} \qquad \text{and} \qquad \mathbb{E}[Z] \to \infty \;\text{ when } p \gg n^{-2/3}\,.$$

Here, by $p \ll n^{-2/3}$, we mean that $\lim_{n\to\infty}(p/n^{-2/3}) = 0$. $p \gg n^{-2/3}$ is defined similarly. We will prove that there is in fact a threshold phenomenon in the probability that G contains a copy of K_4. When $p \ll n^{-2/3}$, the probability that a random graph G generated according to model $G_{n,p}$ contains a copy of K_4, goes to 0 as $n \to \infty$. On the other hand, when $p \gg n^{-2/3}$, this probability goes to 1.

Theorem 5.2 *Let G be a random $G_{n,p}$ graph. We have that:*

- *If $p \ll n^{-2/3}$, then $\mathbb{P}[G$ contains a copy of $K_4] \to 0$ as $n \to \infty$.*
- *If $p \gg n^{-2/3}$, then $\mathbb{P}[G$ contains a copy of $K_4] \to 1$ as $n \to \infty$.*

Proof: As above, we define the random variable Z,

$$Z \;=\; \text{number of copies of } K_4 \text{ in } G \;=\; \sum_C X_C\,.$$

The case when $p \ll n^{-2/3}$ can be easily handled by Markov's inequality. We get that,

$$\mathbb{P}[Z > 0] \;=\; \mathbb{P}[Z \;\geq\; 1] \;\leq\; \frac{\mathbb{E}[Z]}{1}\,.$$

Since $\mathbb{E}[Z] \to 0$ as $n \to \infty$ when $p \ll n^{-2/3}$, we get that $\mathbb{P}[G \text{ contains a copy of } K_4] \to 0$.
When $p \gg n^{-2/3}$, we want to show that $\mathbb{P}[Z > 0] \to 1$ i.e., $\mathbb{P}[Z = 0] \to 0$. We use Chebyshev's Inequality to prove this. We first compute the variance of Z.

$$\mathsf{Var}[Z] \;=\; \mathsf{Var}\left[\sum_C X_C\right] \;=\; \sum_C \mathsf{Var}[X_C] + \sum_{C \neq D} \mathsf{Cov}[X_C, X_D]$$

Since $\mathbb{E}[X_C] = p^6$, we have $\mathsf{Var}[X_C] = p^6 - p^{12}$. Also, for two distinct sets C and D, we consider four different cases depending on the number of vertices they share.

- **Case 1:** $|C \cap D| = 0$. Since no vertex is shared, X_C and X_D are independent and hence $\mathsf{Cov}[X_C, X_D] = 0$.
- **Case 2:** $|C \cap D| = 1$. Since the variables X_C and X_D depend on *pairs* of vertices in the sets C and D, and the two sets do not share any pair, we still have $\mathsf{Cov}[X_C, X_D] = 0$.
- **Case 3:** $|C \cap D| = 2$. Since C and D share a pair of vertices, there are 11 pairs which must all have edges between them in G, for both X_C and X_D to be 1. Thus, we have $\mathbb{E}[X_C X_D] = p^{11}$ and
$$\mathsf{Cov}[X_C, X_D] \;=\; \mathbb{E}[X_C X_D] - \mathbb{E}[X_C] \cdot \mathbb{E}[X_D] \;=\; p^{11} - p^{12}\,.$$
- **Case 4:** $|C \cap D| = 3$. in this case C and D share 3 pairs and thus there are 9 distinct pairs of vertices which must all have edges between them for both X_C and X_D to be 1. Thus,
$$\mathsf{Cov}[X_C, X_D] \;=\; \mathbb{E}[X_C X_D] - \mathbb{E}[X_C] \cdot \mathbb{E}[X_D] \;=\; p^9 - p^{12}\,.$$

Also, there are $\binom{n}{6} \cdot \binom{6}{4}$ pairs C and D such that $|C \cap D| = 2$, and $\binom{n}{5} \cdot \binom{5}{4}$ pairs such that $|C \cap D| = 3$. Combining the above cases we have,

$$\begin{aligned}
\mathsf{Var}[Z] &= \sum_C var(X_C) + \sum_{C \neq D} Cov(X_C, X_D) \\
&= \binom{n}{4} \cdot p^6(1 - p^6) + \binom{n}{6} \cdot \binom{6}{4} \cdot (p^{11} - p^{12}) + \binom{n}{5} \cdot \binom{5}{4} \cdot (p^9 - p^{12}) \\
&= O(n^4 p^6) + O(n^6 p^{11}) + O(n^5 p^9)\,.
\end{aligned}$$

Applying Chebyshev's inequality gives

$$\begin{aligned}\mathbb{P}[Z=0] \;\leq\; \mathbb{P}[|Z-\mathbb{E}[Z]| \geq \mathbb{E}[Z]] \;&\leq\; \frac{\mathsf{Var}[Z]}{(\mathbb{E}[Z])^2} \\ &=\; \frac{1}{\binom{n}{4}^2 \cdot p^{12}} \cdot \left(O(n^4p^6) + O(n^6p^{11}) + O(n^5p^9)\right) \\ &=\; O\left(\frac{1}{n^4p^6}\right) + O\left(\frac{1}{n^2p}\right) + O\left(\frac{1}{n^3p^3}\right).\end{aligned}$$

For $p \gg n^{-2/3}$, all the terms on the right tend to 0 as $n \to \infty$. Hence, $\mathbb{P}[Z=0] \to 0$ as $n \to \infty$. ■

The above analysis can be extended to any graph H of a fixed size. Let Z_H be the number of copies of H in a random graph G generated according to $G_{n,p}$. Using the previous analysis, we have $\mathbb{E}[Z_H] = \Theta\left(n^{|V(H)|} \cdot p^{|E(H)|}\right)$. Hence, $\mathbb{E}[Z] \to 0$ when $p \ll n^{-|V(H)|/|E(H)|}$ and $\mathbb{E}[Z] \to \infty$ when $p \gg n^{-|V(H)|/|E(H)|}$. Thus, it might be tempting to conclude that $p = n^{-|V(H)|/|E(H)|}$ is the threshold probability for finding a copy of H. However, consider the graph in Figure 2. For this

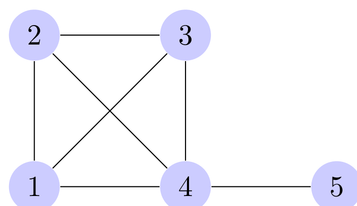

Figure 2: Subgraph H containing K_4

graph, we have $|V(H)|/|E(H)| = 5/7$. But for p such that $p \gg n^{-5/7}$ and $p \ll n^{-2/3}$, a random G is extremely unlikely to contain a copy of K_4 and thus also extremely unlikely to contain a copy of H. For an arbitrary graph H, it was shown by Bollobás [Bol81] that the threshold probability is $n^{-\lambda}$, where

$$\lambda \;=\; \min_{H' \subseteq H} \frac{|V(H')|}{|E(H')|}.$$

References

[ER60] P. Erdős and A. Rényi, "On the Evolution of Random Graphs", *Publ. Math. Inst. Hungar. Acad. Sci.*, **5** (1960), pp. 17–61.

[Bol81] B. Bollobás, "Random Graphs", in *Combinatorics* Proceedings, Swansea 1981, London Mathematical Society Lecture Note Series **52**, Cambridge University Press, pp. 80–102.