Topics in TCS: Randomized Algorithms September 19, 2025

Lecture 2: First and Second Moment Methods

Lecturer: Ola Svensson *Scribes: Ola Svensson*

In today's lecture, we will continue to explore the probabilistic method that we briefly saw in the first lecture (when we proved that any graph as at most $\binom{n}{2}$ cuts using Kargers' beautiful randomized algortihm). We will see more examples of how to use this method to prove the existence of nice objects. Specifically, we will cover first and second moment methods and see how to apply them to random graphs and balls-and-bins.

The specific techniques to learn from today's lecture are as follows:

- The probabilistic method:
  - Want to show a certain object exists.
  - Set up a random experiment such that Pr[object exists] > 0.
  - Thus the object must exist.
- First moment method (Markov's inequality)
- Second moment method (Chebyshev's inequality)
- Erdös-Rényi random graphs
- Balls-and-bins model

1 Graphs with no large cliques or independent sets

1.1 Strangers and Friends

Motivation. Have a dinner party. On the one hand, boring if many people already know each other. On the other hand, tense if many people do not know each other.

Claim 1 *Suppose you invite 6 persons then either* 3 *persons are pairwise friends or* 3 *persons are pairwise strangers.*

Proof Let the invited persons be A, B, C, D, E, F and consider person A.

Case 1: A has 3 friends, say B, C, D. If two of them are friends, say B, C then we have a set A, B, C of 3 friends as claimed. If none of them are friends then we have a set (B, C, D) of 3 strangers as claimed.

Case 2: A has 3 strangers, say B, C, D. If two of them are strangers, say B, C then we have a set A, B, C of 3 strangers as claimed. If none of them are strangers then we have a set (B, C, D) of 3 friends as claimed.

■

Nice simple proof. What about if we invite n people?

Theorem 2 *In a graph G with n vertices, there is either an independent set of size s or a clique of size t as long as $n \geq 2^{s+t} - 1$.*

Proof We give a proof by induction on $s+t$. Note that it is trivially true if $s=1$ or 2. Since if the graph has no independent set of size 2, it is a complete graph. Similar reasoning shows that it is trivial if $t=1$ or $t=2$. Therefore it holds for all s,t such that $s=1,2$ or $t=1,2$.

Now assume true for all s,t such that $s+t=k-1$ for some k. We shall prove that it holds for any $s+t=k$. Fix an arbitrary s,t such that $s+t=k$ and consider a n-vertex graph with $n \geq 2^k-1$. Let A be a vertex of G.

Case 1: A is adjacent to at least $\frac{n-1}{2} \geq \frac{2^k-2}{2} = 2^{k-1}-1$ vertices. In that case the graph induced by these vertices either has (by I.H.) an independent set of size s or a clique of size $t-1$. This implies that the graph G has an independent set of size s or a clique of size t.

Case 2: A is non-adjacent to at least $\frac{n-1}{2} \geq \frac{2^k-2}{2} = 2^{k-1}-1$ vertices. In that case the graph induced by these either has (by I.H) an independent set of size $s-1$ or a clique of size t. This implies that the graph G has an independent set of size s or a clique of size t in this case as well.

■

Remark Note that the above theorem implies that A graph always has an independent set of size $\frac{\log n}{2}$ or a clique of size $\frac{\log n}{2}$.

Can we have a better dependency on $s+t$? NO (up to small constant)!

1.2 No Clique or independent set of size $2\log n$ in a graph from $G(n,1/2)$

To prove this we shall use the probabilistic method and random graphs. More specifically, we shall consider the Erdös-Rényi model.

Definition 3 *For $p \in [0,1]$, a graph sampled from $G(n,p)$ is obtained by including each edge with probability p independent from every other edge.*

We shall now show that there *exists* graphs with no independent set or cliques of size $2\log n$.

Theorem 4 *A graph $G \sim G(n,1/2)$ has no independent set or clique of size $2\log n$ almost surely.*

Proof (To simplify calculations we show it for $t=2\log n+1$.)

- For every $S \subseteq V$, define the random indicator variable

$$X_S = \begin{cases} 1, & \text{if } S \text{ is a clique,} \\ 0, & \text{otherwise.} \end{cases}$$

- Let X equal the number of cliques of size $t=2\log n+1$. Then

$$\begin{aligned} \mathbb{E}[X] &= \mathbb{E}\left[\sum_{|S|=t} X_S\right] \\ &= \sum_{|S|=t} \mathbb{E}[X_S] \\ &= \binom{n}{t} 2^{-\binom{t}{2}} \end{aligned}$$

We now upper bound $\mathbb{E}[X]$.

$$\begin{aligned}\mathbb{E}[X] &= \binom{n}{t} 2^{-\binom{t}{2}} \\ &\le \frac{1}{t!} \frac{n^t}{2^{t(t-1)/2}} \\ &= \frac{1}{t!} \left(\frac{n}{2^{(t-1)/2}}\right)^t \\ &= \frac{1}{t!} \\ &= o(1).\end{aligned}$$

As X is a non-negative integer variable, by Markov's inequality we have

$$\begin{aligned}\Pr[X \neq 0] &= \Pr[X \geq 1] \\ &\le \mathbb{E}[X] \\ &= o(1).\end{aligned}$$

Hence, almost surely the graph has no clique of size $2\log n + 1$. As the probability of having an independent set of size $2\log n + 1$ is the same for $G(n, 1/2)$, we have by the union bound that almost surely the graph has no independent set or clique of size $2\log n + 1$.

■

IMPORTANT TOOLS.

- **Markov's Inequality:** If X is a non-negative random variable and $a > 0$, then

$$\Pr[X \geq a] \leq \frac{\mathbb{E}[X]}{a}.$$

  (It follows from the definition of expectation: $\mathbb{E}[X] = \sum_x x \Pr[X = x] \geq \sum_{x \geq a} x \Pr[X = x] \geq a \sum_{x \geq a} \Pr[X = x] = a \Pr[X \geq a]$, where the first inequality uses that X is non-negative.)

- **Union Bound:** We have

$$\Pr[A \vee B] \leq \Pr[A] + \Pr[B]\,.$$

  This is because $\Pr[A \vee B] = \Pr[A] + \Pr[B] - \Pr[A \wedge B] \leq \Pr[A] + \Pr[B]$.

  It is often useful to apply the union bound over many events $A_1, A_2, \ldots, A_k$:

$$\Pr[A_1 \vee A_2 \vee \ldots \vee A_k] \leq \Pr[A_1] + \Pr[A_2] + \ldots + \Pr[A_k]\,.$$

The First Moment Method in Action. The proof we just saw regarding random graphs is a classic application of the first moment method. The general strategy when we want to show that a certain "bad event" (like the existence of a large clique) does not happen with high probability is as follows:

1. Define a non-negative integer random variable X that counts the number of occurrences of the bad event.
2. Calculate the expectation $\mathbb{E}[X]$ (the first moment).
3. Show that the expectation tends to zero as the problem size grows, i.e., $\mathbb{E}[X] = o(1)$.
4. Apply Markov's inequality: $\Pr[X \geq 1] \leq \mathbb{E}[X] = o(1)$.
5. Conclude that the probability of zero bad events occurring is high:

$$\Pr[X = 0] = 1 - \Pr[X \geq 1] \geq 1 - \mathbb{E}[X] = 1 - o(1).$$

Thus, with high probability (w.h.p.), "zero bad things happen."

2 Balls and Bins

We now consider another classic application of the first moment method: the balls and bins problem.

The Model. Suppose we have n balls and n bins. We throw each ball into a bin uniformly at random (u.a.r.), independently of all other balls.

Expected Load. We can first calculate the expected load of a specific bin, say bin 1. By linearity of expectation:

$$\mathbb{E}[\text{load of bin 1}] = \sum_{j=1}^{n} \Pr[\text{ball } j \text{ lands in bin 1}] = \sum_{j=1}^{n} \frac{1}{n} = 1.$$

The expected load of any bin is 1.

Maximum Load. We are more interested in the maximum load over all bins. Even though the expected load is 1, the distribution is not perfectly uniform. We want to know how large the maximum load can get.

Theorem 5 *When n balls are thrown into n bins u.a.r., the maximum load is $O\left(\frac{\log n}{\log \log n}\right)$ with high probability.*

Proof We use the first moment method. Let $s = \frac{8 \log n}{\log \log n}$. (We assume log is base 2. The constant 8 is chosen to make the analysis work out.) We want to bound the probability that the maximum load is at least s.

If the maximum load is at least s, it implies that there exists at least one bin containing at least s balls. This, in turn, means there must be some set of s balls that fell into the same bin.

Let S be a specific set of s balls. Define the indicator variable X_S:

$$X_S = \begin{cases} 1, & \text{if all balls in } S \text{ land in the same bin,} \\ 0, & \text{otherwise.} \end{cases}$$

Let $X = \sum_{|S|=s} X_S$ be the total number of sets of size s that land in the same bin. If the maximum load is $\geq s$, then $X \geq 1$.

We calculate the expected value of X. First, consider $\mathbb{E}[X_S]$. For a fixed set S, the probability that all s balls land in a specific bin j is $(1/n)^s$. Since there are n possible bins they could all land in:

$$\mathbb{E}[X_S] = \Pr[X_S = 1] = n \cdot \left(\frac{1}{n}\right)^s = \frac{1}{n^{s-1}}.$$

By linearity of expectation:

$$\mathbb{E}[X] = \sum_{|S|=s} \mathbb{E}[X_S] = \binom{n}{s} \frac{1}{n^{s-1}}.$$

We use the standard bound $\binom{n}{s} \leq \frac{n^s}{s!}$.

$$\mathbb{E}[X] \leq \frac{n^s}{s!} \frac{1}{n^{s-1}} = \frac{n}{s!}.$$

Now we need a lower bound for $s!$. We can use the bound $s! \geq s^{s/2}$. This can be shown by pairing terms:

$$(s!)^2 = (1 \cdot s) \cdot (2 \cdot (s-1)) \cdots (s \cdot 1).$$

There are s terms in this product. Each term $k(s-k+1) \geq s$ for $1 \leq k \leq s$ (since this is equivalent to $(k-1)(s-k) \geq 0$). Thus $(s!)^2 \geq s^s$, so $s! \geq s^{s/2}$.

Substituting this bound back:

$$\mathbb{E}[X] \leq \frac{n}{s^{s/2}}.$$

We want to show that this expectation is small. We need to analyze the growth of $s^{s/2}$ with our choice of $s = \frac{8 \log n}{\log \log n}$. Let us show that $s^{s/2} \geq n^2$ for sufficiently large n. We analyze this by taking the logarithm:

$$\log(s^{s/2}) = \frac{s}{2} \log s.$$

We want to show $\frac{s}{2} \log s \geq 2 \log n$.

Substituting s:

$$\begin{aligned}
\frac{s}{2} \log s &= \frac{4 \log n}{\log \log n} \cdot \log\left(\frac{8 \log n}{\log \log n}\right) \\
&= \frac{4 \log n}{\log \log n} \cdot (\log 8 + \log \log n - \log \log \log n) \\
&= \frac{4 \log n}{\log \log n} \cdot (3 + \log \log n - \log \log \log n) \\
&= 4 \log n \cdot \left(1 + \frac{3 - \log \log \log n}{\log \log n}\right).
\end{aligned}$$

As $n \to \infty$, the term in the parenthesis approaches 1. Thus, the expression is asymptotically $4 \log n$. Since $4 \log n \geq 2 \log n$, the inequality $s^{s/2} \geq n^2$ holds for sufficiently large n.

Therefore,

$$\mathbb{E}[X] \leq \frac{n}{n^2} = \frac{1}{n}.$$

Finally, by the first moment method (Markov's inequality):

$$\Pr[\text{Max load} \geq s] \leq \Pr[X \geq 1] \leq \mathbb{E}[X] \leq \frac{1}{n}.$$

The maximum load is $O\left(\frac{\log n}{\log \log n}\right)$ with high probability (at least $1 - 1/n$). ■

3 Second Moment Method and Tight Thresholds

Are the above results tight? Yes!! We will show that using the second moment method.

3.1 Second Moment Method

The first moment method is excellent when we want to show that $\Pr[X = 0] \approx 1$ by showing $\mathbb{E}[X] = o(1)$. But what if $\mathbb{E}[X]$ is large (tends to infinity)? We might expect that $X > 0$ with high probability, but Markov's inequality is insufficient to prove this.

Consider a random variable X such that $X = n^2$ with probability $1/n$ and $X = 0$ otherwise. Then $\mathbb{E}[X] = n^2 \cdot (1/n) = n$. The expectation tends to infinity, but $\Pr[X > 0] = 1/n$, which tends to zero.

The second moment method helps in these situations by utilizing the variance of the random variable to show concentration around the mean.

Definition 1 (Variance and Covariance) *The variance of a random variable X is defined as*

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2.$$

The covariance of two random variables X and Y is

$$\mathrm{Cov}(X, Y) = \mathbb{E}[(X - \mathbb{E}[X])(Y - \mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\,\mathbb{E}[Y].$$

IMPORTANT TOOL: The key tool is Chebyshev's inequality.

Theorem 6 (Chebyshev's Inequality) *For any random variable X and $a > 0$,*

$$\Pr[|X - \mathbb{E}[X]| \geq a] \leq \frac{\mathrm{Var}(X)}{a^2}.$$

Proof Let $Y = (X - \mathbb{E}[X])^2$. Y is a non-negative random variable. By Markov's inequality:

$$\Pr[|X - \mathbb{E}[X]| \geq a] = \Pr[Y \geq a^2] \leq \frac{\mathbb{E}[Y]}{a^2} = \frac{\mathrm{Var}(X)}{a^2}.$$

■

The Strategy. To show that $X > 0$ w.h.p. when $\mathbb{E}[X] \to \infty$, we bound $\Pr[X = 0]$.

$$\begin{aligned}\Pr[X = 0] &\leq \Pr[|X - \mathbb{E}[X]| \geq \mathbb{E}[X]] \quad (\text{Since if } X = 0, \text{ then } |X - \mathbb{E}[X]| = \mathbb{E}[X]) \\ &\leq \frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2} \quad (\text{By Chebyshev's inequality})\end{aligned}$$

If we can show that $\frac{\mathrm{Var}(X)}{(\mathbb{E}[X])^2} = o(1)$ (i.e., the variance is much smaller than the square of the expectation), then $\Pr[X = 0] = o(1)$, meaning $X > 0$ w.h.p.

Calculating Variance. When $X = \sum_{i=1}^{n} X_i$ is a sum of random variables, the variance is:

$$\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j).$$

Calculating the covariance terms is often the most challenging part and requires analyzing the dependencies between the X_i's.

3.2 Tight Thresholds for Balls and Bins

We now use the second moment method to prove that the bound $O\left(\frac{\log n}{\log \log n}\right)$ on the maximum load is tight.

Theorem 7 *When n balls are thrown into n bins u.a.r., the maximum load is $\Omega\left(\frac{\log n}{\log \log n}\right)$ with high probability.*

Proof Let $s = \frac{\log n}{3 \log \log n}$. We want to show that w.h.p., there exists a bin with at least s balls.

Let X_i be the indicator variable that bin i has *exactly* s balls. Let $X = \sum_{i=1}^{n} X_i$ be the total number of bins with exactly s balls. If $X > 0$, the maximum load is at least s.

Step 1: Calculate the Expectation $\mathbb{E}[X]$. First, we calculate the probability that a specific bin i gets exactly s balls:

$$\mathbb{E}[X_i] = \Pr[\text{Bin } i \text{ has } s \text{ balls}] = \binom{n}{s}\left(\frac{1}{n}\right)^s\left(1-\frac{1}{n}\right)^{n-s}.$$

We want to lower bound this probability. We use the bound $\binom{n}{s} \geq \left(\frac{n}{s}\right)^s$. Also, since $s = o(n)$, $\left(1-\frac{1}{n}\right)^{n-s} \approx \left(1-\frac{1}{n}\right)^n \approx \frac{1}{e}$. For sufficiently large n, we can lower bound this term by a constant, say 1/4.

$$\mathbb{E}[X_i] \geq \left(\frac{n}{s}\right)^s \frac{1}{n^s} \cdot \frac{1}{4} = \frac{1}{4s^s}.$$

Now we analyze s^s for our choice of $s = \frac{\log n}{3\log\log n}$. We analyze the logarithm (assuming base 2):

$$\begin{aligned}
\log(s^s) = s\log s &= \frac{\log n}{3\log\log n} \cdot \log\left(\frac{\log n}{3\log\log n}\right) \\
&= \frac{\log n}{3\log\log n} \cdot (\log\log n - \log 3 - \log\log\log n) \\
&= \frac{\log n}{3}\left(1 - \frac{\log 3 + \log\log\log n}{\log\log n}\right).
\end{aligned}$$

As $n \to \infty$, the term in the parenthesis approaches 1. Since the subtracted term is positive (for large enough n), we have $\log(s^s) < \frac{1}{3}\log n = \log(n^{1/3})$. Thus, $s^s < n^{1/3}$.

Substituting this back:

$$\mathbb{E}[X_i] \geq \frac{1}{4n^{1/3}}.$$

By linearity of expectation:

$$\mathbb{E}[X] = \sum_{i=1}^{n} \mathbb{E}[X_i] = n \cdot \mathbb{E}[X_i] \geq \frac{n}{4n^{1/3}} = \frac{1}{4}n^{2/3}.$$

The expectation $\mathbb{E}[X]$ tends to infinity as $n \to \infty$.

Step 2: Analyze the Variance $\mathrm{Var}(X)$. We need to calculate the variance:

$$\mathrm{Var}(X) = \sum_{i=1}^{n} \mathrm{Var}(X_i) + \sum_{i \neq j} \mathrm{Cov}(X_i, X_j).$$

We rely on two important facts.

Fact 1: Negative Correlation. In the balls and bins experiment, the loads of the bins are negatively correlated. If bin i has many balls, it leaves fewer balls for the other bins, making it less likely that bin j also has many balls.

Claim 8 *The indicators X_i, X_j are negatively correlated, i.e.,* $\mathrm{Cov}(X_i, X_j) \leq 0$.

(This is a known property of the multinomial distribution. We omit the formal proof here.)

This implies that the sum of covariances is non-positive, so:

$$\mathrm{Var}(X) \leq \sum_{i=1}^{n} \mathrm{Var}(X_i).$$

Fact 2: Variance of Indicators. Since X_i is an indicator variable ($X_i \in \{0,1\}$), we have $X_i^2 = X_i$.

$$\mathrm{Var}(X_i) = \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2 = \mathbb{E}[X_i] - (\mathbb{E}[X_i])^2 \le \mathbb{E}[X_i].$$

Combining these facts:

$$\mathrm{Var}(X) \le \sum_{i=1}^{n} \mathrm{Var}(X_i) \le \sum_{i=1}^{n} \mathbb{E}[X_i] = \mathbb{E}[X].$$

Step 3: Apply Chebyshev's Inequality. We apply the second moment method strategy:

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

Using the bound $\mathrm{Var}(X) \le \mathbb{E}[X]$:

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

Since we established that $\mathbb{E}[X] \ge \frac{1}{4}n^{2/3}$, we have:

$$\Pr[X = 0] \le \frac{1}{\frac{1}{4}n^{2/3}} = \frac{4}{n^{2/3}}.$$

Therefore, the probability that at least one bin has s balls is

$$\Pr[X > 0] = 1 - \Pr[X = 0] \ge 1 - \frac{4}{n^{2/3}} = 1 - o(1).$$

The maximum load is $\Omega\left(\frac{\log n}{\log\log n}\right)$ with high probability. ■

3.3 The Clique Threshold

We return to the Erdős-Rényi random graph $G \sim G(n, 1/2)$. We previously showed (Theorem 4) that the graph almost surely has no clique of size $2\log n + 1$. We now use the second moment method to show that this threshold is essentially tight. This analysis is significantly more involved than the balls and bins case because the indicator variables for cliques are positively correlated.

Theorem 9 *For any constant $\epsilon > 0$, $G \sim G(n, 1/2)$ has a clique of size $t = 2(1-\epsilon)\log n$ almost surely.*

Proof Let $t = 2(1-\epsilon)\log n$. Let X be the number of cliques of size t. As before, $X = \sum_{|S|=t} X_S$.

Step 1: Expectation. We first show that the expected number of cliques of this size is large.

$$\begin{aligned}
\mathbb{E}[X] &= \binom{n}{t} 2^{-\binom{t}{2}} \\
&\ge \left(\frac{n}{t}\right)^t 2^{-t(t-1)/2} \qquad \left(\text{using } \binom{n}{t} \ge \left(\frac{n}{t}\right)^t\right) \\
&= \left(\frac{n}{t \cdot 2^{(t-1)/2}}\right)^t.
\end{aligned}$$

We analyze the term inside the parenthesis.

$$2^{(t-1)/2} = 2^{(2(1-\epsilon)\log n - 1)/2} = 2^{(1-\epsilon)\log n - 1/2} = \frac{n^{1-\epsilon}}{\sqrt{2}}.$$

Substituting this back:

$$\mathbb{E}[X] \geq \left(\frac{n}{t \cdot n^{1-\epsilon}/\sqrt{2}} \right)^t = \left(\frac{\sqrt{2} \cdot n^{\epsilon}}{t} \right)^t.$$

Since $t = O(\log n)$, n^{ϵ} (polynomial growth) dominates t (logarithmic growth). Thus, $\mathbb{E}[X] \to \infty$ as $n \to \infty$.

Step 2: Variance. Since the expectation is large, we can hope to use the second moment method. Let $\mu = \mathbb{E}[X]$. We want to show $\text{Var}[X]/\mu^2 = o(1)$.

We analyze the variance using covariances:

$$\text{Var}[X] = \sum_{|A|=t, |B|=t} \left(\mathbb{E}[X_A X_B] - \mathbb{E}[X_A]\, \mathbb{E}[X_B] \right).$$

Crucially, X_A and X_B are independent if they do not share any edges. This happens if $|A \cap B| \leq 1$. In this case, the corresponding term is 0.

We only need to sum over pairs (A, B) that overlap in at least 2 vertices. Let $i = |A \cap B|$, where $i \in \{2, \ldots, t\}$.

If $|A \cap B| = i$, the total number of distinct edges in $A \cup B$ is $2\binom{t}{2} - \binom{i}{2}$. Thus,

$$\mathbb{E}[X_A X_B] = 2^{-(2\binom{t}{2} - \binom{i}{2})}.$$

We can upper bound the variance by bounding the covariance by $\mathbb{E}[X_A X_B]$ (since covariance is positive here):

$$\text{Var}[X] \leq \sum_{i=2}^{t} \sum_{|A|=t} \sum_{|B|=t, |A \cap B|=i} \mathbb{E}[X_A X_B].$$

We count the number of such pairs. There are $\binom{n}{t}$ ways to choose A. Given A, there are $\binom{t}{i}$ ways to choose the intersection, and $\binom{n-t}{t-i}$ ways to choose the remaining vertices of B.

$$\begin{aligned} \text{Var}[X] &\leq \sum_{i=2}^{t} \binom{n}{t} \binom{t}{i} \binom{n-t}{t-i} 2^{-(2\binom{t}{2} - \binom{i}{2})} \\ &= \binom{n}{t} 2^{-2\binom{t}{2}} \sum_{i=2}^{t} \binom{t}{i} \binom{n-t}{t-i} 2^{\binom{i}{2}}. \end{aligned}$$

Step 3: The Ratio $\text{Var}[X]/\mu^2$. Recall that $\mu = \binom{n}{t} 2^{-\binom{t}{2}}$. So $\mu^2 = \binom{n}{t}^2 2^{-2\binom{t}{2}}$.

$$\begin{aligned} \frac{\text{Var}[X]}{\mu^2} &\leq \frac{\binom{n}{t} 2^{-2\binom{t}{2}} \sum_{i=2}^{t} \binom{t}{i} \binom{n-t}{t-i} 2^{\binom{i}{2}}}{\binom{n}{t}^2 2^{-2\binom{t}{2}}} \\ &= \sum_{i=2}^{t} \frac{\binom{t}{i} \binom{n-t}{t-i} 2^{\binom{i}{2}}}{\binom{n}{t}}. \end{aligned}$$

To show this expression is $o(1)$, we analyze the contribution of the terms in the sum. Let $R(i)$ be the i-th term. We examine the extreme cases.

Case $i = 2$ (Small overlap):

$$R(2) = \frac{\binom{t}{2}\binom{n-t}{t-2}2^{\binom{2}{2}}}{\binom{n}{t}} = 2\frac{\binom{t}{2}\binom{n-t}{t-2}}{\binom{n}{t}}.$$

We analyze the ratio of the binomial coefficients. For $n \gg t$:

$$\frac{\binom{n-t}{t-2}}{\binom{n}{t}} = \frac{(n-t)!}{(t-2)!(n-2t+2)!}\frac{t!(n-t)!}{n!} = \frac{t(t-1)\cdot((n-t)!)^2}{n!(n-2t+2)!}.$$

Using the approximation $(N-k)!/N! \approx 1/N^k$, this ratio is approximately $\frac{t^2\cdot(1/n^t)\cdot n^{t-2}}{1} = \frac{t^2}{n^2}$.

$$R(2) \approx 2\binom{t}{2}\frac{t^2}{n^2} \approx 2\frac{t^2}{2}\frac{t^2}{n^2} = \frac{t^4}{n^2}.$$

Since $t = O(\log n)$, this term is $O(\log^4 n/n^2) = o(1)$.

Case $i = t$ (Complete overlap, $A = B$):

$$R(t) = \frac{\binom{t}{t}\binom{n-t}{0}2^{\binom{t}{2}}}{\binom{n}{t}} = \frac{2^{\binom{t}{2}}}{\binom{n}{t}} = \frac{1}{\binom{n}{t}2^{-\binom{t}{2}}} = \frac{1}{\mu}.$$

Since we showed $\mu \to \infty$, this term is $o(1)$.

Case $i = t-1$ (Large overlap):

$$\begin{aligned} R(t-1) &= \frac{\binom{t}{t-1}\binom{n-t}{1}2^{\binom{t-1}{2}}}{\binom{n}{t}} \\ &= \frac{t(n-t)2^{\binom{t}{2}-(t-1)}}{\binom{n}{t}} \qquad \left(\text{Since } \binom{t}{2} = \binom{t-1}{2} + t - 1\right) \\ &= \frac{t(n-t)}{2^{t-1}}\cdot\frac{1}{\mu}. \end{aligned}$$

The first fraction grows polynomially in n (it is $O(n\log n)$ divided by $n^{2(1-\epsilon)}$). However, μ grows superpolynomially (since $\log\mu = \Theta(\log^2 n)$). Therefore, $R(t-1) = o(1)$.

Detailed calculations confirm that the sum over all i is indeed negligible, i.e., $\sum_{i=2}^{t} R(i) = o(1)$.

Therefore, $\Pr[X = 0] \le \mathrm{Var}[X]/\mu^2 = o(1)$. The graph almost surely has a clique of size t. ■

4 Summary

In this lecture, we explored the first and second moment methods and applied them to analyze properties of random graphs and the balls and bins model.

Key Results.

- **Balls and Bins:** We showed that when throwing n balls into n bins, the maximum load is $\Theta\left(\frac{\log n}{\log\log n}\right)$ with high probability. A tighter analysis (often using the natural logarithm, ln, see e.g. the textbook by Mitzenmacher and Upfal) reveals the constant factor: the maximum load is $(1+o(1))\frac{\ln n}{\ln\ln n}$ w.h.p.
- **Random Graphs:** We established that the largest clique (and independent set) in $G(n, 1/2)$ is almost surely of size $\approx 2\log_2 n$.

Core Techniques. We formalized the application of two fundamental probability inequalities:

1. **Markov's Inequality:** For $X \geq 0$ and $\lambda > 0$, $\Pr(X \geq \lambda) \leq \frac{\mathbb{E}(X)}{\lambda}$.
2. **Chebyshev's Inequality:** For any random variable X with mean μ, $\Pr(|X-\mu| \geq \lambda) \leq \frac{\mathrm{Var}(X)}{\lambda^2}$.

These lead to the powerful first and second moment methods:

3. **First Moment Method:** Used to show that "bad things" rarely happen. Let X be the count of bad events. If $\mathbb{E}(X) = o(1)$, then $\Pr(X=0) \geq 1 - \mathbb{E}(X) = 1 - o(1)$. W.h.p., zero bad things happen.
4. **Second Moment Method:** Used to show that "good things" likely happen when the expectation is large ($\mathbb{E}[X] \to \infty$). Let X be the count of good events. If $\frac{\mathrm{Var}(X)}{(\mathbb{E}(X))^2} = o(1)$, then $\Pr(X=0) \leq \frac{\mathrm{Var}(X)}{(\mathbb{E}(X))^2} = o(1)$. W.h.p., at least one good thing happens.

Open Problems. The probabilistic method often proves the existence of objects without providing an efficient way to find or construct them.

- **Finding Cliques in Random Graphs:** We know that $G(n, 1/2)$ almost surely has a clique of size $\approx 2\log n$. However, the best known polynomial-time algorithms can only find a clique of size $\approx \log n$ (e.g., by a simple greedy approach). It is a major open problem whether a clique of size significantly larger than $\log n$ (e.g., $(1+\epsilon)\log n$ for any $\epsilon > 0$) can be found in polynomial time.
- **Deterministic Ramsey Constructions:** We showed that there exist graphs with no independent set or clique of size $2\log n$. However, the proof relied on random construction. It is another amazing open problem to deterministically construct such a graph (known as a Ramsey graph). The best known deterministic constructions are still far from the $O(\log n)$ bound achieved by random graphs.