12
Second Moment Method

12.1 Chebyshev's inequality

Markov's inequality is an important tool when bounding probability. It asserts that $\mathbf{Pr}(X \geq t) \leq \mathbb{E}[X]/t$ for all $t > 0$. However, can we do better? We sometimes need a sharper bound to control the concentration of random variables.

Theorem 12.1 (Chebyshev's inequality). $\mathbf{Pr}(|X - \mathbb{E}[X]| \geq t) \leq \frac{\mathbf{Var}[X]}{t^2}$.

Notation 12.2. Variance $\mathbf{Var}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$, which is usually denoted σ^2. $\mathbb{E}[X]$ is usually denoted μ.

Proof.

$$\begin{aligned}\sigma^2 = \mathbb{E}\left[(X-\mu)^2\right] &= \mathbf{Pr}(|X-\mu| \geq t)\ \mathbb{E}\left[(X-\mu)^2 \mid |X-\mu| \geq t\right] \\ &\quad + \mathbf{Pr}(|X-\mu| \leq t)\ \mathbb{E}\left[(X-\mu)^2 \mid |X-\mu| \leq t\right] \\ &\geq \mathbf{Pr}(|X-\mu| \geq t)\ t^2. \end{aligned}$$ □

The use of Chebyshev's inequality is called the second moment method. Now, we will introduce two applications.

Question 12.3. Let S be a positive integer set of size k whose all 2^k subset sums are distinct. What is the minimum possible value of the largest element in S?

A simple argument shows that $\max S \geq 2^k/k$ since all subset sums are at most $k \max S$. However, we can bound $\max S$ in a more clever way, because most subset sums "concentrate" to the mean value by the Chebyshev's inequality.

Theorem 12.4. $\max S \gtrsim \frac{2^k}{\sqrt{k}}$.

Proof. Let $S = \{x_1, \ldots, x_k\}$ and $n = \max S$. For $1 \leq i \leq k$, choose $\varepsilon_i \in \{0,1\}$ independently and uniformly at random. Let $X = \sum \varepsilon_i x_i$. Thus, we have $\mu \triangleq \mathbb{E}[X] = \frac{\sum x_i}{2}$. Also, the variance $\sigma^2 \triangleq \mathbf{Var}[X] = \frac{\sum x_i^2}{4} \leq \frac{n^2 k}{4}$.

By Chebyshev's inequality, $\mathbf{Pr}\left(|X - \mu| < n\sqrt{k}\right) \geq \frac{3}{4}$. Since X takes distinct values for distinct $(\varepsilon_1, \ldots, \varepsilon_k) \in \{0,1\}^k$, we have $\mathbf{Pr}(X = r) \leq 2^{-k}$ for all r. Thus, we have

$$\mathbf{Pr}\left(|X - \mu| < n\sqrt{k}\right) \leq 2^{-k} \cdot 2n\sqrt{k} = \frac{n\sqrt{k}}{2^{k-1}},$$

which implies that $n\sqrt{k} \lesssim 2^k$. □

Remark 12.5. In 2020, Dubroff, Fox and Xu showed that

$$\max S \gtrsim \left(\sqrt{\frac{2}{\pi}} + o(1)\right) \frac{2^k}{\sqrt{k}}.$$

Now, we introduce an application of the second moment method to analysis.

Theorem 12.6 (Weierstrass approximation theorem)**.** *Suppose* $f : [0,1] \to \mathbb{R}$ *is a continuous function. For every* $\varepsilon > 0$*, there exists a polynomial* $p(x)$ *such that*

$$\forall x \in [0,1], \quad |p(x) - f(x)| \leq \varepsilon.$$

Proof. (by Bernstein, 1912) Our key idea is to approximate the value of f at sufficiently dense points (e.g., $0, 1/n, 2/n, \ldots, 1$ for some large n) and then apply uniform continuity.

Since $[0,1]$ is compact, f is uniformly continuous and bounded. Without loss of generality, assume $|f(x)| \leq 1$. There exists $\delta > 0$ such that $|f(x) - f(y)| \leq \frac{\varepsilon}{2}$ for all $|x - y| \leq \delta$.

Now, we approximate f by

$$P_n(x) = \sum_{i=0}^{n} E_i(x) f\left(\frac{i}{n}\right),$$

where $E_i(x)$ is a polynomials and the sum of all $E_i(x)$ is $\sum_i E_i(x) = 1$. For each i, we hope $E_i(x)$ peaks at $\frac{i}{n}$ and decays (rapidly) away from $\frac{i}{n}$. Let $\mathrm{Bin}(n, p)$ be the binomial distribution with parameter n and p.

We can choose

$$E_i(x) = \mathbf{Pr}(\mathrm{Bin}(n,x) = i) = \binom{n}{i} x^i (1-x)^i \,.$$

Since $\mathrm{Bin}(n,x)$ has expectation nx and variance $nx(1-x) \le \frac{n}{4}$, by Chebyshev's inequality we have

$$\sum_{i:|i-nx|>n^{2/3}} E_i(x) = \mathbf{Pr}\Big(|\mathrm{Bin}(n,x) - nx| > n^{2/3}\Big) \le \frac{n^{-1/3}}{4} \,.$$

Note that $\sum_{i=0}^{n} E_i(x) = 1$. Taking $n > \max\{\varepsilon^{-3}, \delta^{-3}\}$, we obtain

$$\begin{aligned} |P_n(x) - f(x)| &\le \sum_{i=0}^{n} E_i(x) \left| f\left(\frac{i}{n}\right) - f(x) \right| \\ &\le \sum_{|i-nx| \le n^{2/3}} \frac{\varepsilon}{2} E_i(x) + \sum_{|i-nx|>n^{2/3}} 2E_i(x) \le \varepsilon \,, \end{aligned}$$

which completes the proof. □

12.2 *Threshold function for graph properties*

We now study the properties of random graphs $\mathcal{G}(n,p)$.

Definition 12.7. A graph property $\mathcal{P}$ is a subset of all graphs.

We say a graph property $\mathcal{P}$ is monotone increasing/decreasing if $G \in \mathcal{P}$ implies supergraphs/subgraphs of G are in $\mathcal{P}$. For instance, for a fixed graph H, the graph property

Namely, the graphs obtained from G by adding / removing edges.

$$\mathcal{P}_1 = \{G \mid H \text{ is an induced sub-graph of } G\}$$

is monotone increasing. The graph property

$$\mathcal{P}_2 = \{G \mid G \text{ is a planar graph}\}$$

is monotone decreasing. However,

$$\mathcal{P}_3 = \{G \mid G \text{ contains a vertex of degree } 1\}$$

is not monotone. A graph property $\mathcal{P}$ is non-trivial if for any sufficiently large n, there always exists a graph with n vertices in $\mathcal{P}$ and another graph not in $\mathcal{P}$.

What we want to discuss is the following natural problem.

Question 12.8. Given a graph property $\mathcal{P}$, for which $p = p_n$ is $\mathcal{G}(n,p)$ in $\mathcal{P}$ with high probability?

Notation 12.9. We will use $f \ll g$ to denote $f = o(g)$, and use $f \gg g$ to denote $g = o(f)$.

Let's start from the easiest case. Suppose $\mathcal{P} = \{G : K_3 \subseteq G\}$. Now, consider $G \sim \mathcal{G}(n, p_n)$. Let X be the number of K_3 in graph G, which is a random variable. Clearly, $\mathbb{E}[X] = \binom{n}{3}p^3$.

If $p \ll \frac{1}{n}$, then $\mathbf{Pr}(X \geq 1) = o(1)$ by Markov's inequality. If $p \gg \frac{1}{n}$, Markov's inequality is not useful. We apply the second moment mothod. Let's first prove that $\mathbf{Var}[X] = o(\mathbb{E}[X]^2)$. Denote S as the set of all subsets of vertices in G of size 3, and denote X_T the indicator variable of the set T inducing a triangle in G. Obviously, $X = \sum_{T \in S} X_T$. Notice that

$$\begin{aligned}
\mathbf{Cov}\left[X_{T_1}, X_{T_2}\right] &= \mathbb{E}\left[X_{T_1} X_{T_2}\right] - \mathbb{E}\left[X_{T_1}\right] \mathbb{E}\left[X_{T_2}\right] \\
&= p^{|E(T_1 \cup T_2)|} - p^{|E(T_1)| + |E(T_2)|} \\
&= \begin{cases} 0 & \text{if } |V(T_1 \cap T_2)| \leq 1 \\ p^5 - p^6 & \text{if } |V(T_1 \cap T_2)| = 2 \end{cases} .
\end{aligned}$$

Also, we have

$$\mathbf{Var}[X_T] = \mathbb{E}\left[X_T^2\right] - \mathbb{E}[X_T]^2 = p^3 - p^6.$$

Therefore,

$$\begin{aligned}
\mathbf{Var}[X] &= \sum_{T \in S} \mathbf{Var}[X_T] + \sum_{\substack{T_1 \neq T_2 \in S \\ T_1 \neq T_2}} \mathbf{Cov}\left[X_{T_1}, X_{T_2}\right] \\
&= \binom{n}{3}\left(p^3 - p^6\right) + \sum_{\substack{T_1 \neq T_2 \in S \\ |V(T_1 \cap T_2)| = 2}} \left(p^5 - p^6\right) \\
&= \binom{n}{3}\left(p^3 - p^6\right) + \binom{n}{2}(n-2)(n-3)\left(p^5 - p^6\right) \\
&\lesssim n^3 p^3 + n^4 p^5 \\
&= o\left(n^6 p^6\right).
\end{aligned}$$

The last equality above holds as $p \gg \frac{1}{n}$. This implies that $\mathbf{Var}[X] = o\left(\mathbb{E}[X]^2\right)$. By Chebyshev's inequality, we can find $\mathbf{Pr}(X = 0) = o(1)$.

We now define threshold functions.

Definition 12.10. We say r_n is a threshold function for some graph property $\mathcal{P}$ if

$$\mathbf{Pr}(\mathcal{G}(n, p_n) \in \mathcal{P}) \to \begin{cases} 0 & \text{if } p_n / r_n \to 0 \\ 1 & \text{if } p_n / r_n \to \infty \end{cases} .$$

From above, we are able to show to the following theorem.

Theorem 12.11. *A threshold function for containing a K_3 is $\frac{1}{n}$.*

Exercise 12.12. *Show that $p = n^{-2/3}$ is a threshold for containing a K_4.*

We now consider general cases. Suppose we have a random variable $X = X_1 + \ldots + X_m$, where X_i is the indicator of event E_i. By Markov's inequality, it is easy to show that $X = 0$ with high probability if $\mathbb{E}[X] = o(1)$. However, it is difficult to show $X > 0$ with high probability if $\mathbb{E}[X] \neq o(1)$. To apply Chebyshev's inequality, we need to bound the variance first.

We say $i \sim j$ if $i \neq j$ and E_i, E_j are not independent. If $i \neq j$ and $i \not\sim j$, we clearly have $\mathbf{Cov}[X_i, X_j] = 0$. Otherwise,

$$\mathbf{Cov}[X_i, X_j] = \mathbb{E}[X_i X_j] - \mathbb{E}[X_i]\,\mathbb{E}[X_j] \leq \mathbb{E}[X_i X_j] = \mathbf{Pr}(E_i \wedge E_j) \ .$$

Also note that $\mathbf{Var}[X_i] \leq \mathbb{E}[X_i^2] = \mathbb{E}[X_i]$, which implies that

$$\mathbf{Var}[X] \leq \mathbb{E}[X] + \sum_{i \sim j} \mathbf{Pr}(E_i \wedge E_j) \ .$$

Define $\Delta := \sum_{i \sim j} \mathbf{Pr}(E_i \wedge E_j)$. We hope $\mathbf{Var}[X] = o(\mathbb{E}[X]^2)$, so if $\mathbb{E}[X] \to \infty$, $\Delta = o(\mathbb{E}[X]^2)$ suffices. Moreover, note that

$$\sum_{i \sim j} \mathbf{Pr}(E_i \wedge E_j) = \sum_i \mathbf{Pr}(E_i) \sum_{j \sim i} \mathbf{Pr}(E_j \mid E_i) \ .$$

In many symmetric cases, $\sum_{j \sim i} \mathbf{Pr}(E_j \mid E_i)$ does not depend on i. Denote by Δ^* this value. Therefore, $\Delta = \sum_i \mathbf{Pr}(E_i)\,\Delta^* = \mathbb{E}[X]\ \Delta^*$. This gives us the following lemma.

Or we may set

$$\Delta^* = \max_i \sum_{j \sim i} \mathbf{Pr}(E_j \mid E_i)$$

in asymmetric cases.

Lemma 12.13. *If $\mathbb{E}[X] \to \infty$ and $\Delta^* = o(\mathbb{E}[X])$, then $X > 0$ with high probability.*

In fact, by Chebyshev's inequality, we have

$$\mathbf{Pr}((1-\varepsilon)\mathbb{E}[X] \leq X \leq (1+\varepsilon)\mathbb{E}[X]) \geq 1 - \frac{\mathbf{Var}[X]}{\varepsilon^2 \mathbb{E}[X]^2} = 1 - o(1)$$

for any constant $0 < \varepsilon < 1$.

Now consider the property of containing K_4. For any set S consisting of exactly four vertices, let E_S be the event that S forms a K_4 in the random graph. For any S, T of size 4, $S \sim T$ if and only if $|S \cap T| \geq 2$. There are two cases:

- $|S \cap T| = 2$:

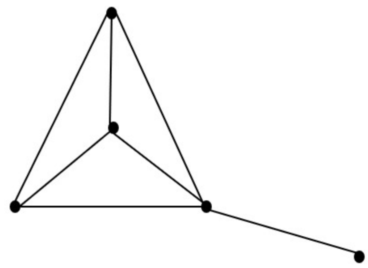
$$\sum_T \mathbf{Pr}(E_T \mid E_S) \leq 6\binom{n}{2}\mathbf{Pr}(E_T \mid E_S) = 6\binom{n}{2}p^5 \approx n^2 p^5;$$

- $|S \cap T| = 3$:

$$\sum_T \mathbf{Pr}(E_T \mid E_S) = 4(n-4)\mathbf{Pr}(E_T \mid E_S) \le 4np^3 \approx np^3.$$

Therefore, $\Delta^* \approx n^2p^5 + np^3 = o(n^4p^6) = o(\mathbb{E}[X])$ if $n^2p \gg 1$ and $np \gg 1$.

One may ask letting X be the number of a general graph H, can we still say that $X > 0$ with high probability if $\mathbb{E}[X] \to \infty$? This is actually not correct. Suppose H is the graph obtained by adding an edge to K_4. Then, $\mathbb{E}[X] \approx n^5p^7 \to \infty$ if $p \gg n^{-5/7}$. However, there is no K_4 in $\mathcal{G}(n,p)$ if $p \ll n^{-2/3}$.



So, can we find a threshold function for containing a general graph? The following theorem tells us the answer.

Definition 12.14. The edge-vertex ratio of $G = (V, E)$ is defined as $\rho(G) = |E|/|V|$. The maximum sub-graph ratio is given by $m(G) = \max_{H \subseteq G} \rho(H)$.

Theorem 12.15 (Bollobás, 1981)**.** *Let $H = (V, E)$ be a fixed graph. Then $p = n^{-1/m(H)}$ is a threshold function for containing H as a sub-graph. Furthermore, if $p \gg n^{-1/m(H)}$, then X_H (number of copies of H in $\mathcal{G}(n,p)$) with high probability satisfies*

$$X_H \approx \mathbf{E}[X] = \binom{n}{|V|}\frac{|V|!}{|Aut(H)|}p^{|E|} \approx \frac{n^{|V|}p^{|E|}}{|Aut(H)|}.$$

Proof. Let H' be the sub-graph of H achieving the maximum edge-vertex ratio, i.e., $m(H) = \rho(H')$. If $p \ll n^{-1/m(H)}$, then $\mathbf{E}[X_{H'}] = o(1)$, which implies that $X_{H'} = 0$ with high probability.

Now assume that $p \gg n^{-1/m(H)}$. Count the labelled copies of H in $\mathcal{G}(n,p)$. Let L be a labelled copy of H in K_n. A_L be the event of $L \subseteq \mathcal{G}(n,p)$. For fixed L, we have

$$\Delta^* = \sum_{L' \sim L} \mathbf{Pr}[A_{L'} | A_L] = \sum_{L' \sim L} p^{|E(L') \setminus E(L)|}.$$

Note that the number of L' such that $L' \sim L$ is approximately $n^{|V(L') \setminus V(L)|}$, and

$$p \gg n^{-1/m(H)} \gg n^{-1/\rho(L' \cap L)} = n^{-|V(L') \cap V(L)|/|E(L') \cap E(L)|}.$$

So, we have

$$\Delta^* \approx \sum n^{|V(L') \setminus V(L)|} p^{|E(L') \setminus E(L)|} \ll n^{|V(L)|} p^{|E(L)|},$$

which implies that $\Delta^* \ll \mathbf{E}[X_H]$. Therefore, $\mathbf{Var}[X] = \mathbf{E}[X_H] + o(\mathbf{E}[X_H])^2$, which completes the proof. □

12.3 Existence of thresholds

In this section, we consider for which graph property $\mathcal{P}$ does a threshold function exist?

Let's start from a simpler question. Assume that $\mathcal{P}$ is monotone increasing, is $f(p) = \mathbf{Pr}[\mathcal{G}(n,p) \in \mathcal{P}]$ increasing? We first discuss the question on upward closed sets.

Let $\mathcal{F}$ be a family of subsets of $[n]$. We call $\mathcal{F}$ an upward closed set (or up-set) if for any $S \subseteq T$ and $S \in \mathcal{F}$, we have $T \in \mathcal{F}$. We have the following theorem.

Theorem 12.16. *Suppose $\mathcal{F}$ is a non-trivial ($\mathcal{F} \neq \emptyset$ or $2^{[n]}$) up-set of $[n]$. Let $Bin([n], p)$ be a random set where each number in $[n]$ is chosen independently with probability p. Then $f(P) = \mathbf{Pr}[Bin([n], p) \in \mathcal{F}]$ is a strictly increasing function.*

Proof. We prove it by *coupling*. For any $0 \leq p < q < 1$, construct a coupling as follows. Pick a uniform random vector $(x_1, \ldots, x_n) \in [0,1]^n$. Let $A = \{i : x_i \leq p\}$ and $B = \{j : x_j \leq q\}$. Clearly, A has the same distribution as $\text{Bin}([n], p)$ and B has the same distribution as $\text{Bin}([n], q)$. Notice that $A \subseteq B$. Thus, we have

$$f(p) = \mathbf{Pr}[A \in \mathcal{F}] < \mathbf{Pr}[B \in \mathcal{F}] = f(q),$$

which completes the proof. □

Here, we give another proof, which is based on two-round exposure coupling.

Proof. Let $0 \leq p < q \leq 1$. Construct A, B as follows:

- For any $i \in [n]$, add i into A with probability p.
- If $i \in A$, add i into B. Otherwise, add it into B with probability $1 - \frac{1-q}{1-p}$.

Notice that $\mathbf{Pr}[i \in B] = p + (1-p)(1 - \frac{1-q}{1-p}) = q$. Therefore, A has the same distribution as $\text{Bin}([n], p)$ and B has the same distribution as $\text{Bin}([n], q)$. The rest of the proof is the same. □

Now, let's prove that every non-trivial monotone increasing graph property has a threshold function.

Theorem 12.17 (Bollobás & Thomason, 1987)**.** *Every non-trivial monotone increasing graph property has a threshold function.*

Proof. Consider k independent copies $G_1, G_2, \ldots, G_k$ of $\mathcal{G}(n,p)$. Their union $G_1 \cup \ldots \cup G_k$ has the same distribution of $\mathcal{G}(n, 1-(1-p)^k)$.

According to the monotonicity of $\mathcal{P}$, if $G_1 \cup \ldots \cup G_k \notin \mathcal{P}$, then $G_i \notin \mathcal{P}$ for all $1 \leq i \leq k$. Note that these k copies are independent, we have

$$\mathbf{Pr}[\mathcal{G}(n, 1-(1-p)^k) \notin \mathcal{P}] \leq \mathbf{Pr}[\mathcal{G}(n,p) \notin \mathcal{P}]^k.$$

Let $f(p) = f_n(p) = \mathbf{Pr}[\mathcal{G}(n,p) \in \mathcal{P}]$. Note that $(1-p)^k \geq 1-kp$. For any monotone increasing property $\mathcal{P}$ and any positive integer $k \leq \frac{1}{p}$, we have

$$1 - f(kp) \leq 1 - f(1-(1-p)^k) \leq (1-f(p))^k.$$

For any sufficiently large n, define a function as follows. Since $f(p)$ is a continuous strictly increasing function from 0 to 1 as p goes from 0 to 1, there is some critical $p_c = p_c(n)$ such that $f(p_c) = \frac{1}{2}$. We claim that p_c is a threshold function.

If $p = p(n) \gg p_c$, then letting $k = \lceil p/p_c \rceil \to \infty$, we have $1 - f(p) \leq (1-f(p_c))^k = 2^{-k} \to 0$. Therefore, $f(p) \to 1$.

Analogously, if $p \ll p_c$, then letting $\ell = \lceil p/p_c \rceil \to \infty$, we have $\frac{1}{2} = 1 - f(p_c) \leq (1-f(p))^\ell$. Thus, $f(p) \to 0$ as $n \to \infty$. This completes the proof. □

12.4 *Sharp thresholds*

In fact, using the method of moments, the number of triangles in a random graph converges to a Poisson distribution. We have

$$\mathbf{Pr}(\mathcal{G}(n, c_n/n) \text{ has triangles}) \to \begin{cases} 0 & \text{if } c_n \to -\infty \\ 1 - \mathrm{e}^{-c^3/6} & \text{if } c_n \to c \\ 1 & \text{if } c_n \to \infty \end{cases}.$$

However, consider some other properties, such as "no isolated vertex". We have

$$\mathbf{Pr}(\mathcal{G}(n,p) \text{ has no isolated vertex}) = \mathrm{e}^{-\mathrm{e}^{-c}}$$

if $c_n \to c$, where $p = \frac{\log n + c_n}{n}$ and $c \in R \cup \{-\infty, \infty\}$. (We leave it as an exercise.) Note that if $c_n \to -\infty$, even though $c_n = -o(\log n)$, we have the probability goes to $\mathrm{e}^{-\mathrm{e}^{-c}} = 0$. Analogously, $\mathrm{e}^{-\mathrm{e}^{-c}} = 1$ if $c_n \to \infty$, even though $c_n = o(\log n)$. So this property shows a stronger notion of threshold: *sharp threshold*.

Definition 12.18. We say r_n is a *sharp threshold* for some graph property $\mathcal{P}$ if for any $\delta > 0$, we have

$$\mathbf{Pr}[\mathcal{G}(n, p_n) \in \mathcal{P}] \to \begin{cases} 0 & \text{if } p_n \le (1-\delta) r_n \\ 1 & \text{if } p_n \ge (1+\delta) r_n \end{cases}.$$

Roughly speaking, any monotone graph property with a coarse threshold may be approximated by a local property (having some H as a sub-graph). This is the famous Friedgut's sharp threshold theorem, which was proved in 1999.

A well-known conjecture is if the property of not being k-colorable has a sharp threshold for some constant (only depending on k) threshold d_k. Namely, we are interested in whether a constant d_k exists, such that

$$\mathbf{Pr}[\mathcal{G}(n, p_n) \text{ is } k\text{-colorable}] \to \begin{cases} 1 & \text{if } d(n) < d_k \\ 0 & \text{if } d(n) > d_k \end{cases}.$$

The following theorem shows that the property of being k-colorable indeed has a sharp threshold.

Theorem 12.19 (Achlioptas & Friedgut, 2000)**.** *For any $k \ge 3$, there exists a function $d_k(n)$ such that for any $\varepsilon > 0$, we have*

$$\mathbf{Pr}[\mathcal{G}(n, p_n) \text{ is } k\text{-colorable}] \to \begin{cases} 1 & d(n) < d_k(n) - \varepsilon \\ 0 & d(n) > d_k(n) + \varepsilon \end{cases}.$$

However, it still remains an open question whether $d_k(n)$ has a limit d_k.

Example 12.20. We now concern the clique numbers of $\mathcal{G}(n, 1/2)$. Let X be the number of k-cliques in $\mathcal{G}(n, 1/2)$. Then we have

$$\mathbf{E}[X] = \binom{n}{k} 2^{-\binom{k}{2}}.$$

Denote it $f(k)$. Clearly $\omega < k$ if $f(k) \to 0$. Now assume $f(k) \to \infty$. Let A_S be the event that S forms a clique in $\mathcal{G}(n, 1/2)$. Fix S, T of size k. Then $S \sim T$ if $|S \cap T| \ge 2$. So we have

$$\Delta^* = \sum_{T \sim S} \mathbf{Pr}(A_T \mid A_S) = \sum_{\ell=2}^{k-1} \binom{k}{\ell} \binom{n-k}{k-\ell} 2^{\binom{\ell}{2} - \binom{k}{2}}.$$

We claim that $\Delta^* = o(f(k))$ if $f(k) \to \infty$. Thus we have $X > 0$ (i.e., $\omega \ge k$) with high probability. Overall, we showed that

$$\omega(\mathcal{G}(n, 1/2)) \approx 2 \log_2 n$$

with high probability.

In fact, we can show that it is a sharp threshold. For $k = (1 \pm o(1)) 2\log_2(n)$, we have

$$\frac{f(k+1)}{f(k)} = \frac{n-k}{k+1} \cdot 2^{-k} = n^{-1+o(1)}.$$

So $f(k)$ decreases rapidly when $k \approx 2\log_2 n$.

Let $k_0 = k_0(n)$ be the value such that $f(k_0) \geq 1 > f(k_0+1)$. For n such that $f(k_0) \to \infty$ and $f(k_0+1) \to 0$, it is known that

$$\omega(\mathcal{G}(n, 1/2)) = k_0$$

with high probability.

If $f(k_0) = O(1)$ (or $f(k_0+1) = O(1)$, then we increacse k_0 by 1), we have $f(k_0-1) \to \infty$ and $f(k_0+1) \to 0$. Thus,

$$\omega(\mathcal{G}(n, 1/2)) \in \{k_0 - 1, k_0\}$$

with high probability. This completes the proof.