T-79.7003: Graphs and Networks **Fall 2013**

Lecture 4: Probabilistic tools and Applications I

Lecturer: Charalampos E. Tsourakakis *Oct. 4, 2013*

4.1 Outline

In Lecture 2, we saw that we can use Markov's Inequality to obtain probabilistic inequalities for higher order moments. Specifically, we saw that if ϕ is a strictly monotonically increasing function, then

$$\mathbf{Pr}\left[X \geq t\right] = \mathbf{Pr}\left[\phi(X) \geq \phi(t)\right] \leq \frac{\mathbb{E}\left[\phi(X)\right]}{\phi(t)}.$$

For instance, for $\phi(x) = x^2$ we obtained Chebyshev's inequality.

Theorem 4.1 (Chebyshev's Inequality) *Let X be any random variable. Then,*

$$\mathbf{Pr}\left[|X - \mathbb{E}\left[X\right]| \geq t\right] \leq \frac{\mathbb{V}ar\left[X\right]}{t^2}.$$

Chebyshev's inequality tells us that the random variable X takes value $\mathbb{E}\left[X\right] + O(\lambda \mathbb{V}ar\left[X\right])$ with probability $1 - O(\lambda^{-2})$. This means that the tail of the probability distribution decays as $O(\lambda^{-2})$. In numerous cases, we are able to get control of higher moments of the variable X. So we may ask, whether we can use this to get better tail estimates. Today, we are going to discuss the exponential moment method which results in deriving the famous Chernoff bounds. The typical setting we are going to deal with is when the random variable is the sum of random variables which are either jointly independent or negatively associated or "almost" independent, in the sense that there may be dependencies but they will be weak. We are going to focus typically on integer-valued, non-negative variables in the context of our class, but keep in mind that these results apply to other settings as well. For instance, there exist Chernoff bounds for complex-valued random variables. Furthermore, in this and the next lectures we are going to see applications of probabilistic tools on random graphs. Finally, since almost all of you who are registered in this class are computer scientists, it is worth noting that these tools come up very often in the analysis of randomized algorithms. See for instance, see the classic book of Motwani-Raghavan [Motwani and Raghavan, 2010]. We will also see an example in Section 4.3.4.

4.2 Exponential Moment Method

In the place of ϕ above, we will use the exponential function. Specifically, let $t > 0, \lambda \in \mathbb{R}$. Then, we obtain

$$\mathbf{Pr}\left[X \geq \lambda\right] = \mathbf{Pr}\left[e^{tX} \geq e^{t\lambda}\right] \leq \frac{\mathbb{E}\left[e^{tX}\right]}{e^{t\lambda}},$$

and

$$\mathbf{Pr}\left[X \leq -\lambda\right] = \mathbf{Pr}\left[e^{-tX} \geq e^{t\lambda}\right] \leq \frac{\mathbb{E}\left[e^{-tX}\right]}{e^{t\lambda}}.$$

The core idea of Chernoff bounds is to set t to a value that minimizes the right-hand side probabilities. Sometimes, we may choose sub-optimal values of t in order to get simpler bounds that will still be good enough for our purposes. The function $M_X(t) = \mathbb{E}\left[e^{tX}\right]$ has a special name since it is an important function. So let's define it so that we can frequently refer to it.

Definition 4.2 (Moment Generating Function) *The function $t \mapsto \mathbb{E}\left[e^{tX}\right]$ is known as the moment generating function (***mgf***) of X.*

It is called like this since by the Taylor series for the exponential

$$\mathbb{E}\left[e^{tX}\right] = 1 + t\mathbb{E}\left[X\right] + \frac{t^2}{2!}\mathbb{E}\left[X^2\right] + \ldots + \frac{t^n}{n!}\mathbb{E}\left[X^n\right] + \ldots$$

we see that all moments of X appear. Specifically, if we take the n-th derivative of $M_X(t)$ and set $t = 0$, then we see that $\mathbb{E}\left[X^n\right] = M_X^{(n)}(0)$. There is a technicality here: we assumed that we can exchange the operands of expectation and differentiation. In general, this is valid when the moment generating function exists in a neighborhood of zero. A well-known distribution whose moment generating function takes infinite values for all $t \neq 0$ is the Cauchy distribution with density $f(x) = \frac{1}{\pi}\frac{1}{1+x^2}$. In the cases we are going to see in this class, the following assumption is going to hold and therefore we can safely exchange the operands of expectation and differentiation.

Assumption: Throughout this class, we will assume that the **mgf** exists in the sense that there exists a positive number $b > 0$ such that $M_X(t)$ is finite for all $|t| < b$.

Two more facts which are useful to keep in mind about **mgf**s follow.

Fact 1: The moment generating function uniquely defined the distribution. Speficially, let X, Y be two random variables. If $M_X(t) = M_Y(t)$ for all $t \in (-\delta, \delta)$ for some $\delta > 0$ then X, Y have the same distribution.

Fact 2: Let's assume that X, Y are two independent random variables. Then the **mgf** $M_{X+Y}(t)$ of the random variables X+Y is $M_X(t)M_Y(t)$. Since the proof is one line, let's see why this is true. The only things we need to use are definitions and the fact that e^{tX}, e^{tY} are independent.

$$M_{X+Y}(t) = \mathbb{E}\left[e^{t(X+Y)}\right] = \mathbb{E}\left[e^{tX}e^{tY}\right] = \mathbb{E}\left[e^{tX}\right]\mathbb{E}\left[e^{tY}\right] = M_X(t)M_Y(t).$$

Example: Let's compute the **mgf** of the binomial $Bin(n, p)$. The binomial is the sum of n independent Bernoulli random variables with parameter p, i.e., if $X \sim Bin(n, p)$, then $X = X_1 + \ldots + X_n$ where $X_i \sim Bernoulli(p)$ for all i. The **mgf** of such a variable is $\mathbb{E}\left[e^{tX_1}\right] = pe^t + (1 - p)$. By fact 2, the **mgf** of X is $M_X(t) = (pe^t + (1 - p))^n$.

Exercise: Work out the **mgf**s of the following two discrete probability distributions: the Poisson distribution with parameter λ and the geometric distribution with parameter p.

The following theorem shows the way we apply the exponential moment method.

Theorem 4.3 *Let X_i for $1 \leq i \leq n$ be jointly independent random variables with $\mathbf{Pr}\left[X_i = 1\right] = \mathbf{Pr}\left[X_i = -1\right] = \frac{1}{2}$. Let $S_n = \sum_{i=1}^n X_i$ and $\alpha > 0$. Then*

$$\mathbf{Pr}\left[|S_n| > \alpha\right] < 2e^{-\frac{\alpha^2}{2n}}.$$

Proof: By symmetry it suffices to prove that $\mathbf{Pr}\left[S_n > \alpha\right] < e^{-\frac{\alpha^2}{2n}}$. Let $t > 0$ be arbitrary. For $1 \leq i \leq n$

$$\mathbb{E}\left[e^{tX_i}\right] = \frac{e^t + e^{-t}}{2} = \cosh(t).$$

Since $\cosh(t) \leq e^{t^2/2}$, see Lemma 4.4, for all $t > 0$ we obtain the following valid inequality:

$$\mathbb{E}\left[e^{tS_n}\right] = \prod_{i=1}^{n} \mathbb{E}\left[e^{tX_i}\right] = (\cosh(t))^n \leq e^{nt^2/2}.$$

Therefore, by the exponential moment method we obtain

$$\mathbf{Pr}\left[S_n > \alpha\right] \leq e^{nt^2/2 - t\alpha}.$$

Setting $t = \frac{\alpha}{n}$ to minimize the right-hand side we get

$$\mathbf{Pr}\left[S_n > \alpha\right] < e^{-\frac{\alpha^2}{2n}}.$$

■

Lemma 4.4 *For all $t > 0$*

$$\cosh(t) \leq e^{t^2/2}.$$

Proof: We will compare the Taylor expansions of the two hand-sides. By the definition of $\cosh(t)$ and the Taylor expansion for e^t we get

$$\cosh(t) = \frac{e^t + e^{-t}}{2} = \sum_{i=0}^{+\infty} \frac{t^{2i}}{(2i)!},$$

since the odd terms of the Taylor expansions for the exponentials cancel out, and the even terms get multiplied by 2 and then divided by 2. Check it. The Taylor expansion for the right-hand side is

$$e^{t^2/2} = \sum_{i=0}^{+\infty} \frac{(t^2/2)^i}{i!} = \sum_{i=0}^{+\infty} \frac{t^{2i}}{2^i i!}.$$

It is easy now to check that the inequality is true. Notice that for all $i \geq 1$

$$(2i)! = (2i)(2i-1)\ldots(i+1) \times i \ldots 1 \geq 2^i i!.$$

■

One can use the exponential moment method to derive the following useful Chernoff bounds, stated as facts. In the next homework you will have the chance to practice the exponential moment method.

Theorem 4.5 (Chernoff bound for $Bin(n,p)$**)** *For any* $0 \le t \le np$

$$\boxed{\mathbf{Pr}\left[|Bin(n,p) - np| > t\right] < 2e^{-\frac{t^2}{3np}}.}$$

For $t > np$

$$\boxed{\mathbf{Pr}\left[|Bin(n,p) - np| > t\right] < \mathbf{Pr}\left[|Bin(n,p) - np| > np\right] < 2e^{-\frac{np}{3}}.}$$

For all t

$$\boxed{\mathbf{Pr}\left[|Bin(n,p) - np| > t\right] < 2\exp\Big(- np\Big((1 + \frac{t}{np})\ln(1 + \frac{t}{np}) - \frac{t}{np}\Big)\Big).}$$

Let's prove another Chernoff-type bound for a random variable S_n that is the sum of the n jointly independent random variables $X_1, \ldots, X_n$. We will assume that X_is are bounded. Despite the strong assumptions we make, what we will derive is a very useful bound. We will start by proving the following lemma.

Lemma 4.6 *Let* X *be a random variable with* $|X| \le 1$, $\mathbb{E}\left[X\right] = 0$. *Then for any* $|t| \le 1$ *the following holds:*

$$M_X(t) \le e^{t^2 \mathbb{V}ar[X]}.$$

Proof: Given that $|tX| \le 1$ the inequality $e^{tX} \le 1 + tX + (tX)^2$ holds. By the linearity of expectation and the fact that $\mathbb{E}\left[tX\right] = t\mathbb{E}\left[X\right] = 0$ we obtain

$$\mathbb{E}\left[e^{tX}\right] \le 1 + t^2\mathbb{E}\left[X^2\right] = 1 + t^2\mathbb{V}ar\left[X\right] \le e^{t^2\mathbb{V}ar[X]}.$$

■

Theorem 4.7 (Chernoff for bounded variables) *Assume that* $X_1, \ldots, X_n$ *are jointly independent random variables where* $|X_i - \mathbb{E}\left[X_i\right]| \le 1$ *for all* i. *Let* $S_n = \sum_{i=1}^n X_i$ *and* $\sigma = \sqrt{\mathbb{V}ar\left[S_n\right]}$ *be the standard deviation of* S_n. *Then for any* $\lambda > 0$

$$\boxed{\mathbf{Pr}\left[|S_n - \mathbb{E}\left[S_n\right]| \ge \lambda\sigma\right] \le 2\max\left(e^{-\lambda^2/4}, e^{-\lambda\sigma/2}\right).}$$

Proof: Without loss of generality we may assume that $\mathbb{E}\left[X_i\right] = 0$ since if not we can subtract a constant from each of the X_is and normalize. Observe that by symmetry it suffices to prove $\mathbf{Pr}\left[S_n \ge \lambda\sigma\right] \le e^{-\frac{t\lambda\sigma}{2}}$ where $t = \min\left(\frac{\lambda}{2\sigma}, 1\right)$. Applying the exponential moment method and by taking into account the joint independence of X_is, $\sum_{i=1}^n \mathbb{V}ar\left[X_i\right] = \sigma^2$ and the previous lemma we obtain

$$\mathbf{Pr}\left[X \geq \lambda\sigma\right] \leq e^{-t\lambda\sigma}\prod_{i=1}^{n}\mathbb{E}\left[e^{tX_i}\right] \leq e^{-t\lambda\sigma}\prod_{i=1}^{n}e^{t^2\mathbb{V}ar[X_i]} = e^{-t\lambda\sigma+t^2\sigma^2}.$$

Since $t \leq \lambda/(2\sigma)$, the proof is complete. ■

Finally, a bound due to Hoeffding which is also known as Chernoff bound or Chernoff-Hoeffding bound is the following. Again, we make the same assumptions, namely $S_n = X_1 + \ldots + X_n$ where X_is are jointly independent and bounded.

Theorem 4.8 (Chernoff-Hoeffding bound) *Suppose $a_i \leq X_i \leq b_i$ for $i = 1, \ldots, n$. Then for all $t > 0$*

$$\boxed{\mathbf{Pr}\left[S_n \geq \mathbb{E}\left[S_n\right] + t\right] \leq e^{-\frac{2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}},}$$

and

$$\boxed{\mathbf{Pr}\left[S_n \leq \mathbb{E}\left[S_n\right] - t\right] \leq e^{-\frac{2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}}.}$$

Combining them we get that

$$\boxed{\mathbf{Pr}\left[\left|S_n - \mathbb{E}\left[S_n\right]\right| \geq t\right] \leq 2e^{-\frac{2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}}.}$$

4.3 Applications

Before we see two applications of Chernoff bound on random graphs in Sections 4.3.2, 4.3.3 let's see what we have gained with Chernoff bounds over the first and second moment method for a coin tossing experiment in Section 4.3.1. We also see an application of Chernoff bounds in the analysis of a simple randomized algorithm in Section 4.3.4.

4.3.1 Coin tossing

Let's assume we have a fair coin which we toss n times. We count how many times heads appeared. Clearly the number of heads S_n in n experiments follows the binomial distribution with parameters n and $p = \frac{1}{2}$. In expectation we will see heads $n/2$ times, i.e., $\mathbb{E}\left[S_n\right] = \frac{n}{2}$. The variance of S_n is equal to n times the variance of each toss given that the tosses are made independently. Therefore, $\mathbb{V}ar\left[S_n\right] = n\frac{1}{2}\left(1 - \frac{1}{2}\right)$. Now, let's compute the probability that what we observe in our experiment deviates from the expectation by a multiplicative factor of $0 < \delta < 1$.

By Markov's inequality we obtain that $\mathbf{Pr}\left[S_n \geq (1+\delta)\mathbb{E}\left[S_n\right]\right] \leq \frac{1}{1+\delta}$. Similarly, we can bound the probability $\mathbf{Pr}\left[S_n \leq (1-\delta)\mathbb{E}\left[S_n\right]\right] \leq \frac{1}{1+\delta}$ since $n - S_n$ is also distributed binomially with the same parameters as S_n. Therefore, Markov's inequality results in

$$\mathbf{Pr}\left[\left|S_n - \mathbb{E}\left[S_n\right]\right| \geq \delta\mathbb{E}\left[S_n\right]\right] \leq \frac{2}{1+\delta}.$$

Chebyshev's inequality results in

$$\mathbf{Pr}\left[|S_n - \mathbb{E}\left[S_n\right]| \geq \delta\mathbb{E}\left[S_n\right]\right] \leq \frac{\mathbb{V}ar\left[S_n\right]}{\delta^2\mathbb{E}\left[S_n\right]^2} = \frac{1}{\delta^2 n}.$$

Finally, Chernoff's bound for the binomial results in the following

$$\mathbf{Pr}\left[|S_n - \mathbb{E}\left[S_n\right]| \geq \delta\mathbb{E}\left[S_n\right]\right] \leq 2e^{-\delta^2\mathbb{E}[S_n]/3} = 2e^{-\delta^2 n/6}.$$

Notice that the tail decays exponentially fast in n.

4.3.2 Maximum and minimum degree

We use the standard graph-theoretic notation: $\delta(G)$ = minimum degree in G and $\Delta(G)$ = maximum degree in G.

Theorem 4.9 *(a) If $p = \frac{c}{n}$ for some constant $c > 0$ then $\Delta(G(n,p)) = O(\frac{\log n}{\log\log n})$* **whp** *. (b) If $np = \omega(n)\log n$ for some slowly growing function $\omega(n) \to +\infty$ as $n \to +\infty$ then $\delta(G(n,p)) = \Delta(G(n,p)) \sim np$* **whp** .

Proof: (a) Let $k = \frac{\log n}{\log\log n - 2\log\log\log n}$. By substituting the value of k in the expression $k\log k$ and the fact that when x is small then $\frac{1}{1-x} = 1 + x + O(x^2)$ is a good approximation, it is easy to check that

$$k\log k \geq \frac{\log n}{\log\log n}(\log\log n + \log\log\log n + o(1)).$$

Let's use the first moment method and the union bound to prove that **whp** there exists no vertex in $G(n,p)$ with degree greater than k.

$$\mathbf{Pr}\left[\exists v : d(v) \geq k\right] \leq n\binom{n-1}{k}p^k \leq \exp\Big(\log n - k\log k + O(k)\Big) = o(1).$$

(b) We apply the union bound and the Chernoff bound with $\epsilon = \omega(n)^{-1/3}$.

$$\mathbf{Pr}\left[\exists v : |d(v) - (n-1)p| \geq \epsilon p(n-1)\right] \leq n2e^{-\epsilon^2 np/3} = 2n^{1-\frac{\omega^{1/3}(n)}{3}} = o(1).$$

■

Exercise: In Theorem 4.9(a) we proved that **whp** there is no degree greater than a constant fraction of $\frac{\log n}{\log\log n}$. Prove by using the second moment method that there exist vertices with degree $\frac{\log n}{\log\log n + 2\log\log\log n}$. By combining the two results you see that the maximum degree for this range of p is $\Theta(\frac{\log n}{\log\log n})$.

4.3.3 Diameter

Theorem 4.10 *Let $d \geq 2$ be a fixed integer. Suppose $c > 0$ and $p^d n^{d-1} = \ln \frac{n^2}{c}$. Then* **whp** *$diam(G(n,p)) \geq d$.*

Proof: Let's define for $v \in V(G)$ the set $N_k(v)$ to be the set of vertices whose distance is exactly k from v. Namely, let

$$N_k(v) = \{w : d(v,w) = k\}.$$

We will show that **whp** for $0 \leq k < d$, the k-th neighborhood of v is $o(n)$. Notice that since d is a constant, we have at least one witness (actually a huge number of witnesses, namely $(1-o(1))n$) that the diameter is at least d **whp** . Specifically, we will prove that $|N_k(v)| \leq (2np)^k$ **whp** . Notice that $|N_k(v)| \sim Bin]\Big(n - \sum_{i=0}^{k-1} |N_i(v)|, 1-(1-p)^{|N_{k-1}(v)|}\Big)$. Let's define event $\mathcal{A}_i = \{|N_i(v)| \leq (2np)^i\}$ for each i and condition on $\mathcal{A}_1, \ldots, \mathcal{A}_{k-1}$. Then the number of neighbors of v k steps away is distributed binomially as $|N_k(v)| \sim Bin(\nu, q)$ where $\nu < n$ and $q = 1-(1-p)^{|N_{k-1}(v)|} \leq p|N_{k-1}(v)|$. Given what we have conditioned on, $q \leq p(2pn)^{k-1}$. Notice that $q < 1$ given that p satisfies $p^d n^{d-1} = \ln \frac{n^2}{c}$ and $k < d$. The conditional expectation is

$$\mathbb{E}\left[|N_k(v)||\mathcal{A}_1, \ldots, \mathcal{A}_{k-1}\right] = \nu q \leq np|N_{k-1}(v)|.$$

Now we use the Chernoff bound for the binomial to upper bound the probability of the bad event $\mathcal{A}_k|\mathcal{A}_1, \ldots, \mathcal{A}_{k-1}$. Namely,

$$\begin{aligned}
\mathbf{Pr}\left[|N_k(v)| \geq (2np)^k|\mathcal{A}_1, \ldots, \mathcal{A}_{k-1}\right] &\leq \mathbf{Pr}\left[Bin(n, p|N_{k-1}(v)|) \geq (2np)^k|\mathcal{A}_{k-1}\right] \\
&\leq \mathbf{Pr}\left[Bin(n, p(2np)^{k-1}) \geq (2np)^k|\mathcal{A}_{k-1}\right] \\
&\leq e^{-\frac{n^k p^k 2^{k-1}}{3}} \\
&= o(n^{-2}).
\end{aligned}$$

Therefore, the claim follows by observing that

$$\mathbf{Pr}\left[\cup_{k=0}^{d-1} N_k(v) = [n]\right] \leq \sum_{k=1}^{d-1} \mathbf{Pr}\left[\bar{\mathcal{A}}_k|\mathcal{A}_1, \ldots, \mathcal{A}_{k-1}\right] = o(n^{-2}).$$

■

For the sake of completeness and in order to see how parameter c affects the diameter, it is worth mentioning the following theorem.

Theorem 4.11 *Let $d \geq 2$ be a fixed integer. Suppose $c > 0$ and $p^d n^{d-1} = \ln \frac{n^2}{c}$. Then with probability $e^{-c/2}$ the diameter is d and with the remaining probability $1 - e^{-c/2}$ the diameter is $d+1$.*

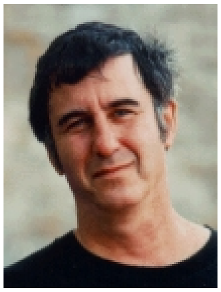

Figure 4.1: Joel Spencer proved his favorite result [Spencer, 1985] known as “six standard deviations suffice” while being in the audience of a talk.

4.3.4 Discrepancy

Consider a set system, a.k.a. hypergraph, $(V, \mathcal{F})$ where $V = [n]$ is the ground set and $\mathcal{F} = \{A_1, \ldots, A_m\}$ where $A_i \subseteq V$. We wish to color the ground set V with two colors, say red and blue, in such way that all sets in the family are colored in a “balanced” way, i.e., each set has nearly the same number of red and blue points. As it can be seen from the family $\mathcal{F} = 2^{[n]}$ this is not possible, since by the pidgeonhole principle at least one color will appear at least $n/2$ times and all the possible subsets of those points will be monochromatic. We formalize the above ideas immediately. It shall be convenient to use in the place of red/blue colorings, the coloring

$$\chi : V \to \{-1, +1\}.$$

For any $A \subseteq V$ define

$$\chi(A) = \sum_{i \in A} \chi(i).$$

Define the discrepancy of $\mathcal{F}$ with respect to χ by

$$\mathrm{disc}_\chi(\mathcal{F}) = \max_{A_i \in \mathcal{F}} |\chi(A_i)|.$$

The discrepancy of $\mathcal{F}$ is

$$\mathrm{disc}(\mathcal{F}) = \min_{\chi} \mathrm{disc}_\chi(\mathcal{F}).$$

It is worth outlining that the discrepancy can be defined in a linear algebraic way. Specifically, let A be the $m \times n$ incidence matrix of $\mathcal{F}$. Then,

$$\mathrm{disc}(\mathcal{F}) = \min_{x \in \{-1,+1\}} ||Ax||_{+\infty}.$$

Let’s prove the next theorem by applying union and Chernoff bounds.

Theorem 4.12

$$disc(\mathcal{F}) \leq \sqrt{2n\log{(2m)}}.$$

Proof: Select a coloring χ uniformly at random from the set of all possible random colorings. Let us call A_i bad if its discrepancy exceeds $t = \sqrt{2n\log 2m}$. Applying the Chernoff-Hoeffding bound for set A_i we obtain:

$$\mathbf{Pr}\left[A_i \text{ is bad}\right] = \mathbf{Pr}\left[|\chi(A_i)| > t\right] < 2\exp\Big(-\frac{t^2}{2|A_i|}\Big) \leq 2\exp\Big(-\frac{t^2}{2n}\Big) = \frac{1}{m}.$$

Using a simple union bound we see that

$$\mathbf{Pr}\left[\text{disc}(\mathcal{F}) > t\right] = \mathbf{Pr}\left[\exists \text{ bad } A_i\right] < m \times \frac{1}{m}.$$

■

Theorem 4.12 serves as our basis for a randomized algorithm that succeeds with as high probability as we want. Let $t = \sqrt{2n\log 2m}$. Since the probability of obtaining a coloring that gives discrepancy larger than t is less than $\frac{1}{\sqrt{m}}$, we can boost the success probability by repeating the random coloring k times. The failure probability is at most $\frac{1}{m^{k/2}}$. Assume $m = n$. We have proved that the discrepancy is $O(\sqrt{n\log n})$. Again, for the sake of completeness, a famous result of Joel Spencer states that $\text{disc}(F) = O(\sqrt{n})$.

References

[Motwani and Raghavan, 2010] Motwani, R. and Raghavan, P. (2010). Randomized algorithms. In *Algorithms and theory of computation handbook*, pages 12–12. Chapman & Hall/CRC.

[Spencer, 1985] Spencer, J. (1985). Six standard deviations suffice. *Transactions of the American Mathematical Society*, 289(2):679–706.