2 Basic Tools

Before reading and studying results on random graphs included in the text, one should become familiar with the basic rules of asymptotic computation, find leading terms in combinatorial expressions, choose suitable bounds for the binomials, and get acquainted with probabilistic tools needed to study tail bounds, i.e., the probability that a random variable exceeds (or is smaller than) some real value. This chapter offers the reader a short description of these important technical tools used throughout the text. For more information about the topic of this chapter, we refer the reader to an excellent expository book, titled *Asymptotia*, written by Joel Spencer with Laura Florescu (see [108]).

2.1 Asymptotics

The study of random graphs and networks is mainly of an asymptotic nature. This means that we explore the behavior of discrete structures of very large "size," say n. It is quite common to analyze complicated expressions of their numerical characteristics, say $f(n)$, in terms of their rate of growth or decline as $n \to \infty$. The usual way is to "approximate" $f(n)$ with a much simpler function $g(n)$.

We say that $f(n)$ is *asymptotically equal* to $g(n)$ and write $f(n) \sim g(n)$ if $f(n)/g(n) \to 1$ as $n \to \infty$.

Example 2.1 The following functions $f(n)$ and $g(n)$ are asymptotically equal:

(a) Let $f(n) = \binom{n}{2}$, $g(n) = n^2/2$. Then $\binom{n}{2} = n(n-1)/2 \sim n^2/2$.
(b) Let $f(n) = 3\binom{n}{3}p^2$, where $p = m/\binom{n}{2}$. Find m such that $f(n) \sim g(n) = 2\omega^2$. Now,

$$f(n) = 3\frac{n(n-1)(n-2)}{6} \cdot \frac{4m^2}{(n(n-1))^2} \sim \frac{2m^2}{n},$$

so m should be chosen as $\omega\sqrt{n}$.

We write $f(n) = O(g(n))$ when there is a positive constant C such that for all sufficiently large n, $|f(n)| \le C|g(n)|$, or, equivalently, $\limsup_{n\to\infty} |f(n)|/|g(n)| < \infty$.

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

Similarly, we write $f(n) = \Omega(g(n))$ when there is a positive constant c such that for all sufficiently large n, $|f(n)| \geq c|g(n)|$ or, equivalently, $\liminf_{n\to\infty} |f(n)|/|g(n)| > 0$.

Finally, we write $f(n) = \Theta(g(n))$ when there exist positive constants c and C such that for all sufficiently large n, $c|g(n)| \leq |f(n)| \leq C|g(n)|$ or, equivalently, $f(n) = O(g(n))$ and $f(n) = \Omega(g(n))$.

Note that $f(n) = O(g(n))$ simply means that the growth rate of $f(n)$ as $n \to \infty$ does not exceed the growth rate of $g(n)$, $f(n) = \Omega(g(n))$ such that $f(n)$ is growing at least as quickly as $g(n)$, while $f(n) = \Theta(g(n))$ states that their order of growth is identical.

Note also that if $f(n) = f_1(n)f_2(n) + \cdots + f_k(n)$, where k is fixed and for $i = 1, 2, \ldots, k$, $f_i(n) = O(g(n))$, then $f(n) = O(g(n))$ as well. In fact, the above property also holds if we replace O by Ω or Θ.

Example 2.2 Let

(a) $f(n) = 5n^3 - 7\log n + 2n^{-1/2}$; then

$$\begin{aligned} f(n) &= O(n^3), \\ f(n) &= 5n^3 + O(\log n), \\ f(n) &= 5n^3 - 7\log n + O(n^{-1/2}). \end{aligned}$$

(b) $f(x) = e^x$; then $f(x) = 1 + x + x^2/2 + O(x^3)$ for $x \to 0$.

We now introduce the frequently used “little o” notation.

We write $f(n) = o(g(n))$ or $f(n) \ll g(n)$ if $f(n)/g(n) \to 0$ as $n \to \infty$. Similarly, we write $f(n) = \omega(g(n))$ or $f(n) \gg g(n)$ if $f(n)/g(n) \to \infty$ as $n \to \infty$. Obviously, if $f(n) \ll g(n)$, then we can also write $g(n) \gg f(n)$.

Note that $f(n) = o(g(n))$ simply means that $g(n)$ grows faster with n than $f(n)$, and the other way around if $f(n) = \omega(g(n))$.

Let us also make a few important observations. Obviously, $f(n) = o(1)$ means that $f(n)$ itself tends to 0 as $n \to \infty$. Also the notation $f(n) \sim g(n)$ is equivalent to the statement that $f(n) = (1 + o(1))g(n)$. One should also note the difference between the $(1 + o(1))$ factor in the expression $f(n) = (1 + o(1))g(n)$ and when it is placed in the exponent, i.e., when $f(n) = g(n)^{1+o(1)}$. In the latter case, this notation means that for every fixed $\varepsilon > 0$ and sufficiently large n, $g(n)^{1-\varepsilon} < f(n) < g(n)^{1+\varepsilon}$. Hence here the $(1 + o(1))$ factor is more accurate in $f(n) = (1 + o(1))g(n)$ than the much coarser factor $(1 + o(1))$ in $f(n) = g(n)^{(1+o(1))}$.

It is also worth mentioning that, regardless of how small a constant $c > 0$ is and however large a positive constant C is, the following hierarchy of growths holds:

$$\ln^C n \ll n^c, \quad n^C \ll (1 + c)^n, \quad C^n \ll n^{cn}. \tag{2.1}$$

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

Example 2.3 Let

(a) $f(n) = \binom{n}{2}p$, where $p = p(n)$. Then $f(n) = o(1)$ if $p = 1/n^{2+\varepsilon}$, where $\varepsilon > 0$, since $f(n) \sim \frac{n^2}{2}n^{-2-\varepsilon} = n^{-\varepsilon}/2 \to 0$.
(b) $f(n) = 3\binom{n}{3}p^2$, where $p = m/\binom{n}{2}$ and $m = n^{1/2}/\omega$, where $\omega = \omega(n) \to \infty$ as $n \to \infty$. Then $f(n) = o(1)$ since $f(n) \le n^4/\left(2\binom{n}{2}^2\omega^2\right) \to 0$.

Exercises

2.1.1 Let $f(n) = \binom{n}{3}p^2(1-p)^{2(n-3)}$, where $\log n - \log\log n \le np \le 2\log n$. Then show that $f(n) = O(n^3p^2e^{-2np}) = o(1)$.
2.1.2 Let $f(n) = \left(1 - \frac{c}{n}\right)^{n(\log n)^2}$, where c is a constant. Then show that $f(n) = o(n^{-2})$.
2.1.3 Let $p = \frac{\log n}{n}$ and $f(n) = \sum_{k=2}^{n/2} n^k p^{k-1}(1-p)^{k(n-k)}$. Then show that $f(n) = o(1)$.
2.1.4 Suppose that $k = k(n) = \lceil 2\log_{1/(1-p)} n\rceil$ and $0 < p < 1$ is a constant. Then show that $\binom{n}{k}(1-p)^{k(k-1)/2} \to 0$.

2.2 Binomials

We start with the famous asymptotic estimate for $n!$, known as Stirling's formula.

Lemma 2.4

$$n! = (1 + o(1))n^n e^{-n}\sqrt{2\pi n}.$$

Moreover,

$$n^n e^{-n}\sqrt{2\pi n} \le n! \le n^n e^{-n}\sqrt{2\pi n}\, e^{1/12n}.$$

Example 2.5 Consider the coin-tossing experiment where we toss a fair coin $2n$ times. What is the probability that this experiment results in *exactly* n heads and n tails? Let A denote such an event. Then $\mathbb{P}(A) = \binom{2n}{n}2^{-2n}$.

By Stirling's approximation,

$$\binom{2n}{n} = \frac{(2n)!}{(n!)^2} \sim \frac{(2n)^{2n}e^{-2n}\sqrt{2\pi(2n)}}{(n^ne^{-n})^2(2\pi n)} = \frac{2^{2n}}{\sqrt{\pi n}}.$$

Hence $\mathbb{P}(A) \sim 1/\sqrt{\pi n}$.

Example 2.6 What is the number of digits in 100!? To answer this question we shall use sharp bounds on $n!$ given in Lemma 2.4. Notice that

$$1 < \frac{n!}{n^n e^{-n}\sqrt{2\pi n}} < e^{1/12n}.$$

Hence, taking logarithms,

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

$$0 < \ln n! - \left(n + \frac{1}{2}\right) \ln n - n + \frac{1}{2} \ln 2\pi < \frac{1}{12n}.$$

Now, since the number of digits in a positive integer n is $\lfloor \log_{10} n + 1 \rfloor$, we divide both sides by $\ln 10$, to get

$$0 < \log_{10} n! - \frac{\left(n + \frac{1}{2}\right) \ln n - n + \frac{1}{2} \ln 2\pi}{\ln 10} < \frac{1}{12n \ln 10}.$$

Substituting $n = 100$, we obtain

$$0 < \log_{10} 100! - 157.96 < 0.00036.$$

Thus 100! has exactly 158 digits.

Before we move to the analysis of the asymptotic behavior of the binomial coefficient $\binom{n}{k}$, let us prove some simple, often-used upper and lower bounds, valid for all fixed n and k.

Lemma 2.7 *For every integer n and k, $k \leq n$,*

$$\binom{n}{k} \leq \left(\frac{ne}{k}\right)^k, \tag{2.2}$$

$$\frac{n^k}{k!}\left(1 - \frac{k(k-1)}{2n}\right) \leq \binom{n}{k} \leq \frac{n^k}{k!}\left(1 - \frac{k}{2n}\right)^{k-1}, \tag{2.3}$$

$$\binom{n}{k} \leq \frac{n^k}{k!} e^{-k(k-1)/(2n)}. \tag{2.4}$$

Proof To prove (2.2), note that

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{(n)_k}{k!},$$

where

$$(n)_k = n(n-1)(n-2)\cdots(n-k+1) \leq n^k.$$

By Lemma 2.4, $k! > (k/e)^k$ and the first bound holds.

To see that remaining bounds on $\binom{n}{k}$ are true we have to estimate $(n)_k$ more carefully. Note first that

$$\binom{n}{k} = \frac{n^k}{k!}\frac{(n)_k}{n^k} = \frac{n^k}{k!}\prod_{i=0}^{k-1}\left(1 - \frac{i}{n}\right). \tag{2.5}$$

The upper bound in (2.3) follows from the observation that for $i = 1, 2, \ldots, \lfloor k/2 \rfloor$,

$$\left(1 - \frac{i}{n}\right)\left(1 - \frac{k-i}{n}\right) \leq \left(1 - \frac{k}{2n}\right)^2.$$

The lower bound in (2.3) is implied by the Weierstrass product inequality, which states that

$$\prod_{r=1}^{s}(1-a_r)+\sum_{r=1}^{s}a_r \geq 1 \tag{2.6}$$

for $0 \leq a_1, a_2, \ldots, a_s \leq 1$, and can be easily proved by induction. Hence

$$\prod_{i=0}^{i-1}\left(1-\frac{i}{n}\right) \geq 1-\sum_{i=0}^{k-1}\frac{i}{n} = 1-\frac{k(k-1)}{2n}.$$

The last bound given in (2.4) immediately follows from the upper bound in (2.3) and the simple observation that for every real x,

$$1+x \leq e^x. \tag{2.7}$$

□

Example 2.8 To illustrate an application of (2.2), let us consider the function

$$f(n,k) = \binom{n}{k}(1-2^{-k})^n,$$

where n, k are positive integers, and denote by n_k the smallest n (as a function of k) such that $f(n,k) < 1$. We aim for an upper estimate of n_k as a function of k, when $k \geq 2$. In fact, we claim that

$$n_k \leq \left(1+\frac{3\log_2 k}{k}\right)k^2 2^k \ln 2. \tag{2.8}$$

Now, by (2.2) and (2.7),

$$f(n,k) = \binom{n}{k}(1-2^{-k})^n \leq \left(\frac{ne}{k}\right)^k e^{-n/2^k}.$$

If $m = (1+\varepsilon)k^2 2^k \ln 2$, then

$$\left(\frac{me}{k}\right)^k e^{-m/2^k} = ((1+\varepsilon)2^k k 2^{-(1+\varepsilon)k} e \ln 2)^k. \tag{2.9}$$

If $\varepsilon = 3\log_2 k/k$, then the right-hand side (RHS) of (2.9) equals $((1+\varepsilon)k^{-2}e\ln 2)^k$, which is less than 1. This implies that n_k satisfies (2.8).

In the following chapters, we shall also need the bounds given in the next lemma.

Lemma 2.9 *If $a \geq b$, then*

$$\left(\frac{k-b}{n-b}\right)^b\left(\frac{n-k-a+b}{n-a}\right)^{a-b} \leq \frac{\binom{n-a}{k-b}}{\binom{n}{k}} \leq \left(\frac{k}{n}\right)^b\left(\frac{n-k}{n-b}\right)^{a-b}.$$

Proof To see this note that

$$
\begin{aligned}
\frac{\binom{n-a}{k-b}}{\binom{n}{k}} &= \frac{(n-a)!k!(n-k)!}{(k-b)!(n-k-a+b)!n!} \\
&= \frac{k(k-1)\cdots(k-b+1)}{n(n-1)\cdots(n-b+1)} \times \frac{(n-k)(n-k-1)\cdots(n-k-a+b+1)}{(n-b)(n-b-1)\cdots(n-a+1)} \\
&\leq \left(\frac{k}{n}\right)^b \left(\frac{n-k}{n-b}\right)^{a-b}.
\end{aligned}
$$

The lower bound follows similarly. □

Example 2.10 Let us show that

$$\frac{\binom{\binom{n}{2}-2l+r}{m-2l+r}}{\binom{\binom{n}{2}}{m}} = O\left(\frac{(2m)^{2l-r}}{n^{4l-2r}}\right)$$

assuming that $2l - r \ll m, n$.

Applying Lemma 2.9 with n replaced by $\binom{n}{2}$ and with $k = m, a = b = 2l - r$, we see that

$$\frac{\binom{\binom{n}{2}-2l+r}{m-2l+r}}{\binom{\binom{n}{2}}{m}} \leq \left(\frac{m}{\binom{n}{2}}\right)^{2l-r}.$$

We will also need precise estimates for the binomial coefficient $\binom{n}{k}$ when $k = k(n)$. They are based on the Stirling approximation of factorials and estimates given in Lemma 2.7.

Lemma 2.11 *Let k be fixed or grow with n as $n \to \infty$. Then*

$$\binom{n}{k} \sim \frac{n^k}{k!} \qquad \textit{if } k = o(n^{1/2}), \tag{2.10}$$

$$\binom{n}{k} \sim \frac{n^k}{k!} \exp\left\{-\frac{k^2}{2n}\right\} \qquad \textit{if } k = o(n^{2/3}), \tag{2.11}$$

$$\binom{n}{k} \sim \frac{n^k}{k!} \exp\left\{-\frac{k^2}{2n} - \frac{k^3}{6n^2}\right\} \qquad \textit{if } k = o(n^{3/4}). \tag{2.12}$$

Proof The asymptotic formula (2.10) follows directly from (2.3) and (2.4). We only prove (2.12) since the proof of (2.11) is analogous. In fact, in the proofs of these bounds we use the Taylor expansion of $\ln(1-x)$, $0 < x < 1$. In the case of (2.12), we take

$$\ln(1-x) = -x - \frac{x^2}{2} + O(x^3). \tag{2.13}$$

Now,

$$\begin{aligned}\binom{n}{k} &= \frac{n^k}{k!}\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right)\\ &= \frac{n^k}{k!}\exp\left\{\ln\left[\prod_{i=0}^{k-1}\left(1-\frac{i}{n}\right)\right]\right\}\\ &= \frac{n^k}{k!}\exp\left\{\sum_{i=0}^{k-1}\ln\left(1-\frac{i}{n}\right)\right\}\\ &= \frac{n^k}{k!}\exp\left\{-\sum_{i=0}^{k-1}\left(\frac{i}{n}+\frac{i^2}{2n^2}\right)+O\left(\frac{k^4}{n^3}\right)\right\}.\end{aligned}$$

Hence

$$\binom{n}{k} = \frac{n^k}{k!}\exp\left\{-\frac{k^2}{2n}-\frac{k^3}{6n^2}+O\left(\frac{k^4}{n^3}\right)\right\}, \tag{2.14}$$

and equation (2.12) follows. □

Example 2.12 Let n be a positive integer, $k = o(n^{1/2})$ and $m = o(n)$. Applying (2.10) and the bounds from Lemma 2.9 we show that

$$\frac{1}{2}\binom{n}{k}(k-1)!\frac{\binom{\binom{n}{2}-k}{m-k}}{\binom{\binom{n}{2}}{m}} \sim \frac{1}{2}\frac{n^k}{k!}(k-1)!\left(\frac{2m}{n^2}\right)^k \sim \frac{\left(\frac{2m}{n}\right)^k}{2k}.$$

Example 2.13 As an illustration of the application of (2.11) we show that if $k = k(n) \gg n^{2/5}$, then

$$f(n,k) = \binom{n}{k}k^{k-2}\left(\frac{1}{n}\right)^{k-1}\left(1-\frac{1}{n}\right)^{\binom{k}{2}-k+1+k(n-k)} = o(1).$$

By (2.11) and Stirling's approximation of $k!$ (Lemma 2.4), we get

$$\binom{n}{k} \sim \frac{n^k}{k!}e^{-k^2/2n} \sim e^{-k^2/2n}\left(\frac{ne}{k}\right)^k(2\pi k)^{-1/2}.$$

Moreover, since

$$\binom{k}{2}-k+1+k(n-k) = kn-\frac{k^2}{2}+O(k)$$

and

$$\ln\left(1-\frac{1}{n}\right) = -\frac{1}{n}+O(n^{-2}),$$

we have

$$\left(1-\frac{1}{n}\right)^{\binom{k}{2}-k+1+k(n-k)} = \exp\left\{-k+\frac{k^2}{2n}+o(1)\right\}.$$

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

Hence

$$f(n,k) \sim e^{-k^2/2n}\left(\frac{ne}{k}\right)^k k^{k-2}(2\pi k)^{-1/2} n^{-k+1} e^{-k+k^2/2n}$$
$$\sim nk^{-5/2}(2\pi)^{-1/2} = o(1).$$

Example 2.14 Let

$$f(n,k,l) = \binom{n}{k} C(k,k+l) p^{k+l}(1-p)^{\binom{k}{2}-(k+l)+k(n-k)}, \tag{2.15}$$

where $k \le n,\ l = o(k)$ and $np = 1+\varepsilon,\ 0 < \varepsilon < 1$.

Assuming that $f(n,k,l) \le n/k$ and applying (2.14) and (2.13), we look for an asymptotic upper bound on $C(k,k+l)$ as follows:

$$\begin{aligned} f(n,k,l) &= C(k,k+l)p^{k+l}\frac{n^k}{k!}\exp\left\{-\frac{k^2}{2n} - \frac{k^3}{6n^2} + O\left(\frac{k^4}{n^3}\right)\right\} \\ &\quad \times \exp\left\{\left(-p - \frac{p^2}{2} + O(p^3)\right)\left(\binom{k}{2} - (k+l) + k(n-k)\right)\right\} \\ &= C(k,k+l)\frac{(np)^{k+l}}{n^l k!}\exp\left\{-\frac{k^2}{2n} - \frac{k^3}{6n^2} - pkn + \frac{pk^2}{2}\right\} \\ &\quad \times \exp\left\{O\left(\frac{k^4}{n^3} + pk + p^2kn\right)\right\}. \end{aligned}$$

Recalling that $f(n,k,l) \le n/k,\ p = (1+\varepsilon)/n$ and using the Stirling approximation for $k!$, we get

$$\begin{aligned} C(k,k+l) &\le n^{l+1}(k-1)!\exp\left\{-\varepsilon k + \frac{\varepsilon^2 k}{2} + \frac{k^3}{6n^2} + k + \varepsilon k - \frac{\varepsilon k^2}{2n}\right\} \\ &\quad \times \exp\left\{O\left(\frac{k^4}{n^3} + \varepsilon l\right)\right\} \\ &\le 3n^{l+1}k^{k-\frac{1}{2}}\exp\left\{\frac{\varepsilon^2 k}{2} + \frac{k^3}{6n^2} - \frac{\varepsilon k^2}{2n} + O\left(\frac{k^4}{n^3} + \varepsilon l\right)\right\}. \end{aligned}$$

Exercises

2.2.1 Let $f(n,k) = 2\frac{\binom{k}{2}\binom{n-k}{k-2}}{\binom{n}{k}}$, where $k = k(n) \to \infty$ as $n \to \infty,\ k = o(n^{1/2})$. Show that $f(n,k) \sim k^4/n^2$.

2.2.2 Let $f(n,k) = \left(\frac{\binom{n-2}{k}}{\binom{n-1}{k}}\right)^{n-1}$, where $k = k(n)$. Show that $f(n,k) \sim e^{-k}$.

2.2.3 Let $f(n,k) = \sum_{k=2}^{n}\sum_{j=0}^{n-k} k^2\binom{n}{k}\binom{n}{j}(k-1)!j!\left(\frac{c}{n}\right)^{k+j+1}$, where $c < 1$. Show that $f(n,k) = O(1/n)$.

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

2.2.4 Apply (2.2) to show that

$$f(n,k) = \sum_{k=1}^{n/1000} \binom{n}{k}\binom{n}{2k}\left(\frac{\binom{3k}{2}}{\binom{n-1}{2}}\right)^{2k} = o(1).$$

2.2.5 Prove that n_k in Example 2.8 satisfies $n_k \geq k^2 2^k \ln k$ for sufficiently large k. Use equation (2.3) and $\ln(1-x) \geq -\frac{x}{1-x}$ if $0 < x < 1$ to get a lower bound for $\left(1-2^{-k}\right)^n$. The latter inequality following from $\ln(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}$.

2.3 Tail Bounds

One of the most basic and useful tools in the study of random graphs is *tail bounds*, i.e., upper bounds on the probability that a random variable exceeds a certain real value. We first explore the potential of the simple but indeed very powerful Markov inequality.

Lemma 2.15 (Markov Inequality) *Let X be a non-negative random variable. Then, for all $t > 0$,*

$$\mathbb{P}(X \geq t) \leq \frac{\mathbb{E}X}{t}.$$

Proof Let

$$I_A = \begin{cases} 1 & \text{if event } A \text{ occurs,} \\ 0 & \text{otherwise.} \end{cases}$$

Notice that

$$X = XI_{\{X\geq t\}} + XI_{\{X<t\}} \geq XI_{\{X\geq t\}} \geq tI_{\{X\geq t\}}.$$

Hence,

$$\mathbb{E}X \geq t\mathbb{E}I_{\{X\geq t\}} = t\,\mathbb{P}(X \geq t).$$

□

Example 2.16 Let X be a random variable with the expectation $\mathbb{E}\,X = n((n-2)/n))^m$, where $m = m(n)$. Find m such that

$$\mathbb{P}(X \geq \sqrt{n}) \leq e^{-c},$$

where $c > 0$ is a constant. By the Markov inequality

$$\mathbb{P}(X \geq \sqrt{n}) \leq \frac{n\left(\frac{n-2}{n}\right)^m}{\sqrt{n}} \leq \sqrt{n}e^{-2m/n}.$$

So m should be chosen as $m = \frac{1}{2}n(\log n^{1/2} + c)$.

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

Example 2.17 Let X be a random variable with the expectation

$$\mathbb{E}\,X_k = \binom{n}{k} k^{k-2} p^{k-1},$$

where $k \geq 3$ is fixed. Find $p = p(n)$ such that $\mathbb{P}(X \geq 1) = O(\omega^{1-k})$, where $\omega = \omega(n)$. Note that by the Markov inequality $\mathbb{P}(X \geq 1) \leq \mathbb{E}\,X$; hence

$$\mathbb{P}(X \geq 1) \leq \binom{n}{k} k^{k-2} p^{k-1} \leq \left(\frac{ne}{k}\right)^k k^{k-2} p^{k-1}.$$

Now put $p = 1/(\omega n^{k/(k-1)})$ to get

$$\mathbb{P}(X \geq 1) \leq \left(\frac{ne}{k}\right)^k k^{k-2} \left(\frac{1}{\omega n^{k/(k-1)}}\right)^{k-1} = \frac{e^k}{k^2 \omega^{k-1}} = O(\omega^{1-k}).$$

We are very often concerned with bounds on the *upper and lower tail* of the distribution of S, i.e., on $\mathbb{P}(X \geq \mathbb{E}\,X + t)$ and $\mathbb{P}(X \leq \mathbb{E}\,X - t)$, respectively. The following joint tail bound on the deviation of a random variable from its expectation is a simple consequence of Lemma 2.15.

Lemma 2.18 (Chebyshev Inequality) *If X is a random variable with a finite mean and variance, then, for $t > 0$,*

$$\mathbb{P}(|X - \mathbb{E}\,X| \geq t) \leq \frac{\operatorname{Var} X}{t^2}.$$

Proof

$$\mathbb{P}(|X - \mathbb{E}\,X| \geq t) = \mathbb{P}((X - \mathbb{E}\,X)^2 \geq t^2) \leq \frac{\mathbb{E}(X - \mathbb{E}\,X)^2}{t^2} = \frac{\operatorname{Var} X}{t^2}.$$

□

Example 2.19 Consider a standard coin-tossing experiment where we toss a fair coin n times and count, say, the number X of heads. Note that $\mu = \mathbb{E}\,X = n/2$, while $\operatorname{Var} X = n/4$. So, by the Chebyshev inequality,

$$\mathbb{P}\left(\left|X - \frac{n}{2}\right| \geq \varepsilon n\right) \leq \frac{n/4}{(\varepsilon n)^2} = \frac{1}{4n\varepsilon^2}.$$

Hence,

$$\mathbb{P}\left(\left|\frac{X}{n} - \frac{1}{2}\right| \geq \varepsilon\right) \leq \frac{1}{4n\varepsilon^2},$$

so if we choose, for example, $\varepsilon = 1/4$, we get the following bound:

$$\mathbb{P}\left(\left|\frac{X}{n} - \frac{1}{2}\right| \geq \frac{1}{4}\right) \leq \frac{4}{n}.$$

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

Suppose again that X is a random variable and $t > 0$ is a real number. We focus our attention on the observation due to Bernstein [17], which can lead to the derivation of stronger bounds on the lower and upper tails of the distribution of the random variable X.

Let $\lambda \geq 0$ and $\mu = \mathbb{E}\, X$; then

$$\mathbb{P}(X \geq \mu + t) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(\mu+t)}) \leq e^{-\lambda(\mu+t)}\, \mathbb{E}(e^{\lambda X}) \tag{2.16}$$

by the Markov inequality (see Lemma 2.15).

Similarly for $\lambda \leq 0$,

$$\mathbb{P}(X \leq \mu - t) = \mathbb{P}(e^{\lambda X} \geq e^{\lambda(\mu-t)}) \leq e^{-\lambda(\mu-t)}\, \mathbb{E}(e^{\lambda X}). \tag{2.17}$$

Combining (2.16) and (2.17) one can obtain a bound for $\mathbb{P}(|X - \mu| \geq t)$. A bound of such type was considered above, that is, the Chebyshev inequality.

We will next discuss in detail tail bounds for the case where a random variable is the sum of independent random variables. This is a common case in the theory of random graphs. Let

$$S_n = X_1 + X_2 + \cdots + X_n,$$

where $X_i, i = 1, \ldots, n$ are independent random variables.

Assume that $0 \leq X_i \leq 1$ and $\mathbb{E}\, X_i = \mu_i$ for $i = 1, 2, \ldots, n$. Let $\mathbb{E}\, S_n = \mu_1 + \mu_2 + \cdots + \mu_n = \mu$. Then, by (2.16), for $\lambda \geq 0$,

$$\mathbb{P}(S_n \geq \mu + t) \leq e^{-\lambda(\mu+t)} \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i}), \tag{2.18}$$

and, by (2.16), for $\lambda \leq 0$,

$$\mathbb{P}(S_n \leq \mu - t) \leq e^{-\lambda(\mu-t)} \prod_{i=1}^{n} \mathbb{E}(e^{\lambda X_i}). \tag{2.19}$$

In the above bounds we applied the observation that the expected value of the product of independent random variables is equal to the product of their expectations. Note also that $\mathbb{E}(e^{\lambda X_i})$ in (2.18) and (2.19), likewise $\mathbb{E}(e^{\lambda X})$ in (2.16) and (2.17), are the moment-generating functions of the X_i and X, respectively. So finding bounds boils down to the estimation of these functions. Now the convexity of e^x and $0 \leq X_i \leq 1$ implies that

$$e^{\lambda X_i} \leq 1 - X_i + X_i e^{\lambda}.$$

Taking expectations, we get

$$\mathbb{E}(e^{\lambda X_i}) \leq 1 - \mu_i + \mu_i e^{\lambda}.$$

Equation (2.18) becomes, for $\lambda \geq 0$,

$$\begin{aligned} \mathbb{P}(S_n \geq \mu + t) &\leq e^{-\lambda(\mu+t)} \prod_{i=1}^{n} (1 - \mu_i + \mu_i e^{\lambda}) \\ &\leq e^{-\lambda(\mu+t)} \left(\frac{n - \mu + \mu e^{\lambda}}{n} \right)^{n}. \end{aligned} \tag{2.20}$$

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

The second inequality follows from the fact that the geometric mean is at most the arithmetic mean, i.e., $(x_1x_2\cdots x_n)^{1/n} \leq (x_1+x_2+\cdots+x_n)/n$ for non-negative $x_1, x_2, \ldots, x_n$. This in turn follows from Jensen's inequality and the concavity of $\log x$. The RHS of (2.20) attains its minimum, as a function of λ, at

$$e^{\lambda} = \frac{(\mu+t)(n-\mu)}{(n-\mu-t)\mu}. \tag{2.21}$$

Hence, by (2.20) and (2.21), assuming that $\mu + t < n$,

$$\mathbb{P}(S_n \geq \mu+t) \leq \left(\frac{\mu}{\mu+t}\right)^{\mu+t}\left(\frac{n-\mu}{n-\mu-t}\right)^{n-\mu-t}, \tag{2.22}$$

while for $t > n - \mu$ this probability is zero.

Now let

$$\begin{aligned}\varphi(x) &= (1+x)\log(1+x) - x, \quad x \geq -1,\\ &= \sum_{k=2}^{\infty}\frac{(-1)^k x^k}{k(k-1)} \quad \text{for } |x| \leq 1,\end{aligned}$$

and let $\varphi(x) = \infty$ for $x < -1$. Now, for $0 \leq t < n - \mu$, we can rewrite the bound (2.22) as

$$\mathbb{P}(S_n \geq \mu+t) \leq \exp\left\{-\mu\varphi\left(\frac{t}{\mu}\right) - (n-\mu)\varphi\left(\frac{-t}{n-\mu}\right)\right\}.$$

Since $\varphi(x) \geq 0$ for every x, we get

$$\mathbb{P}(S_n \geq \mu+t) \leq e^{-\mu\varphi(t/\mu)}. \tag{2.23}$$

Similarly, putting $n - S_n$ for S_n, or by an analogous argument, using (2.19), we get, for $0 \leq t \leq \mu$,

$$\mathbb{P}(S_n \leq \mu-t) \leq \exp\left\{-\mu\varphi\left(\frac{-t}{\mu}\right) - (n-\mu)\varphi\left(\frac{t}{n-\mu}\right)\right\}.$$

Hence,

$$\mathbb{P}(S_n \leq \mu-t) \leq e^{-\mu\varphi(-t/\mu)}. \tag{2.24}$$

We can simplify expressions (2.23) and (2.24) by observing that

$$\varphi(x) \geq \frac{x^2}{2(1+x/3)}. \tag{2.25}$$

To see this observe that for $|x| \leq 1$, we have

$$\varphi(x) - \frac{x^2}{2(1+x/3)} = \sum_{k=2}^{\infty}(-1)^k\left(\frac{1}{k(k-1)} - \frac{1}{2\cdot 3^{k-2}}\right)x^k.$$

Equation (2.25) for $|x| \leq 1$ follows from $\frac{1}{k(k-1)} - \frac{1}{2\cdot 3^{k-2}} \geq 0$ for $k \geq 2$. We leave it as Exercise 2.3.3 to check that (2.25) remains true for $x > 1$.

Taking this into account we arrive at the following theorem, see Hoeffding [59].

Theorem 2.20 (Chernoff–Hoeffding inequality) *Suppose that* $S_n = X_1 + X_2 + \cdots + X_n$ *while, for* $i = 1, 2, \ldots, n$,

(i) $0 \le X_i \le 1$,
(ii) $X_1, X_2, \ldots, X_n$ *are independent.*

Let $\mathbb{E}\, X_i = \mu_i$ *and* $\mu = \mu_1 + \mu_2 + \cdots + \mu_n$. *Then for* $t \ge 0$,

$$\mathbb{P}(S_n \ge \mu + t) \le \exp\left\{-\frac{t^2}{2(\mu + t/3)}\right\} \tag{2.26}$$

and for $t \le \mu$,

$$\mathbb{P}(S_n \le \mu - t) \le \exp\left\{-\frac{t^2}{2(\mu - t/3)}\right\}. \tag{2.27}$$

Putting $t = \varepsilon\mu$, for $0 < \varepsilon < 1$, in (2.23), (2.26) and (2.27), one can immediately obtain the following bounds.

Corollary 2.21 *Let* $0 < \varepsilon < 1$; *then*

$$\mathbb{P}(S_n \ge (1+\varepsilon)\mu) \le \left(\frac{e^{\varepsilon}}{(1+\varepsilon)^{1+\varepsilon}}\right)^{\mu} \le \exp\left\{-\frac{\mu\varepsilon^2}{3}\right\}, \tag{2.28}$$

while

$$\mathbb{P}(S_n \le (1-\varepsilon)\mu) \le \exp\left\{-\frac{\mu\varepsilon^2}{2}\right\}. \tag{2.29}$$

□

Note also that the bounds (2.28) and (2.29) imply that, for $0 < \varepsilon < 1$,

$$\mathbb{P}(|S_n - \mu| \ge \varepsilon\mu) \le 2\exp\left\{-\frac{\mu\varepsilon^2}{3}\right\}. \tag{2.30}$$

Example 2.22 Let us return to the coin-tossing experiment from Example 2.19. Notice that the number of heads X is in fact the sum of binary random variables X_i, for $i = 1, 2, \ldots, n$, each representing the result of a single experiment, that is, $X_i = 1$, with probability $1/2$, when head occurs in the ith experiment, and $X_i = 0$, with probability $1/2$, otherwise. Denote this sum by $S_n = X_1 + X_2 + \cdots + X_n$ and notice that random variables X_i are independent. Applying the Chernoff bound (2.30), we get

$$\mathbb{P}\left(\left|S_n - \frac{n}{2}\right| \ge \varepsilon\frac{n}{2}\right) \le 2\exp\left\{-\frac{n\varepsilon^2}{6}\right\}.$$

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

Choosing $\varepsilon = 1/2$, we get

$$\mathbb{P}\left(\left|\frac{S_n}{n} - \frac{1}{2}\right| \geq \frac{1}{4}\right) \leq 2e^{-n/24},$$

a huge improvement over the Chebyshev bound.

Example 2.23 Let S_n now denote the number of heads minus the number of tails after n flips of a fair coin. Find $\mathbb{P}(S_n \geq \omega\sqrt{n})$, where $\omega = \omega(n) \to \infty$ arbitrarily slowly, as $n \to \infty$.

Notice that S_n is again the sum of independent random variables X_i, but now $X_i = 1$, with probability $1/2$, when head occurs in the ith experiment, while $X_i = -1$, with probability $1/2$, when tail occurs. Hence, for each $i = 1, 2, \ldots, n$, expectation $\mathbb{E}\, X_i = 0$ and variance $\operatorname{Var} X_i = 1$. Therefore, $\mu = \mathbb{E}\, S_n = 0$ and $\sigma^2 = \operatorname{Var} S_n = n$. So, by (2.26)

$$\mathbb{P}(S_n \geq \omega\sqrt{n}) \leq e^{-3\omega/2}.$$

To compare, notice that Chebyshev's inequality yields the much weaker bound since it implies that

$$\mathbb{P}(S_n \geq \omega\sqrt{n}) \leq \frac{1}{2\omega^2}.$$

One can "tailor" the Chernoff bounds with respect to specific needs. For example, for small ratios t/μ, the exponent in (2.26) is close to $t^2/2\mu$, and the following bound holds.

Corollary 2.24

$$\mathbb{P}(S_n \geq \mu + t) \leq \exp\left\{-\frac{t^2}{2\mu} + \frac{t^3}{6\mu^2}\right\} \tag{2.31}$$

$$\leq \exp\left\{-\frac{t^2}{3\mu}\right\} \qquad \textit{for } t \leq \mu. \tag{2.32}$$

Proof Use (2.26) and note that

$$(\mu + t/3)^{-1} \geq (\mu - t/3)/\mu^2.$$

□

Example 2.25 Suppose that $p = c/n$ for some constant c and that we create an $n \times n$ matrix A with values 0 or 1, where for all i, j, $\Pr(A(i, j) = 1) = p$ independently of other matrix entries. Let Z denote the number of columns that are all zero. We will show that, for small $\varepsilon > 0$,

$$\Pr(Z \geq (1 + \varepsilon)e^{-c}n) \leq e^{-\varepsilon^2 e^{-c} n/3}.$$

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

Each column of A is zero with probability $q = (1-p)^n = (1+O(1/n))e^{-c}$. Furthermore, Z is the sum of indicator random variables and is distributed as the binomial $Bin(n,q)$. Applying (2.31) with $\mu = nq, t = \varepsilon\mu$, we get

$$\Pr(Z \geq (1+\varepsilon)e^{-c}n) \leq \exp\left\{-\frac{\varepsilon^2\mu}{2} + \frac{\varepsilon^3\mu}{6}\right\} \leq \exp\left\{-\frac{\varepsilon^2 e^{-c}n}{3}\right\}.$$

For large deviations we have the following result.

Corollary 2.26 *If* $c > 1$, *then*

$$\mathbb{P}(S_n \geq c\mu) \leq \left(\frac{e}{ce^{1/c}}\right)^{c\mu}. \tag{2.33}$$

Proof Put $t = (c-1)\mu$ into (2.23). □

Example 2.27 Let $X_1, X_2, \ldots, X_n$ be independent binary random variables, that is, $X_i \in \{0,1\}$ with the Bernoulli distribution: $\mathbb{P}(X_i = 1) = p$, $\mathbb{P}(X_i = 0) = 1-p$, for every $1 \leq i \leq n$, where $0 < p < 1$. Then $S_n = \sum_{i=1}^n X_i$ has the binomial distribution with the expectation $\mathbb{E}\, S_n = \mu = np$. Applying Corollary 2.26 one can easily show that for $t = 2e\mu$,

$$\mathbb{P}(S_n \geq t) \leq 2^{-t}.$$

Indeed, for $c = 2e$,

$$\mathbb{P}(S_n \geq t) = \mathbb{P}(S_n \geq c\mu) \leq \left(\frac{e}{ce^{1/c}}\right)^{c\mu} \leq \left(\frac{1}{2e^{1/(2e)}}\right)^{2e\mu} \leq 2^{-t}.$$

We also have the following:

Corollary 2.28 *Suppose that* $X_1, X_2, \ldots, X_n$ *are independent random variables and that* $a_i \leq X_i \leq b_i$ *for* $i = 1, 2, \ldots, n$. *Let* $S_n = X_1 + X_2 + \cdots + X_n$ *and* $\mu_i = \mathbb{E}(X_i)$, $i = 1, 2, \ldots, n$ *and* $\mu = \mathbb{E}(S_n)$. *Then for* $t > 0$ *and* $c_i = b_i - a_i$, $i = 1, 2, \ldots, n$, *we have*

$$\mathbb{P}(S_n \geq \mu + t) \leq \exp\left\{-\frac{2t^2}{c_1^2 + c_2^2 + \cdots + c_n^2}\right\}, \tag{2.34}$$

$$\mathbb{P}(S_n \leq \mu - t) \leq \exp\left\{-\frac{2t^2}{c_1^2 + c_2^2 + \cdots + c_n^2}\right\}. \tag{2.35}$$

Proof We can assume without loss of generality that $a_i = 0, i = 1, 2, \ldots, n$. We just subtract $A = \sum_{i=1}^n a_i$ from S_n. We proceed as before.

$$\mathbb{P}(S_n \geq \mu + t) = \mathbb{P}\left(e^{\lambda S_n} \geq e^{\lambda(\mu+t)}\right) \leq e^{-\lambda(\mu+t)}\, \mathbb{E}\left(e^{\lambda S_n}\right) = e^{-\lambda t} \prod_{i=1}^n \mathbb{E}\left(e^{\lambda(X_i - \mu_i)}\right).$$

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

Note that $e^{\lambda x}$ is a convex function of x, and since $0 \leq X_i \leq c_i$, we have

$$e^{\lambda(X_i-\mu_i)} \leq e^{-\lambda\mu_i}\left(1 - \frac{X_i}{c_i} + \frac{X_i}{c_i}e^{\lambda c_i}\right)$$

and so

$$\begin{aligned}\mathbb{E}(e^{\lambda(X_i-\mu_i)}) &\leq e^{-\lambda\mu_i}\left(1 - \frac{\mu_i}{c_i} + \frac{\mu_i}{c_i}e^{\lambda c_i}\right)\\ &= e^{-\theta_i p_i}\left(1 - p_i + p_i e^{\theta_i}\right),\end{aligned} \tag{2.36}$$

where $\theta_i = \lambda c_i$ and $p_i = \mu_i/c_i$.

Then, taking the logarithm of the RHS of (2.36), we have

$$\begin{aligned}f(\theta_i) &= -\theta_i p_i + \log\left(1 - p_i + p_i e^{\theta_i}\right),\\ f'(\theta_i) &= -p_i + \frac{p_i e^{\theta_i}}{1 - p_i + p_i e^{\theta_i}},\\ f''(\theta_i) &= \frac{p_i(1-p_i)e^{-\theta_i}}{((1-p_i)e^{-\theta_i} + p_i)^2}.\end{aligned}$$

Now $\frac{\alpha\beta}{(\alpha+\beta)^2} \leq 1/4$ and so $f''(\theta_i) \leq 1/4$, and therefore

$$f(\theta_i) \leq f(0) + f'(0)\theta_i + \frac{1}{8}\theta_i^2 = \frac{\lambda^2 c_i^2}{8}.$$

It follows then that

$$\mathbb{P}(S_n \geq \mu + t) \leq e^{-\lambda t}\exp\left\{\sum_{i=1}^{n}\frac{\lambda^2 c_i^2}{8}\right\}.$$

We obtain (2.34) by putting $\lambda = \frac{4}{\sum_{i=1}^n c_i^2}$, and (2.35) is proved in a similar manner. □

There are many cases when we want to use our inequalities to bound the upper tail of some random variable Y and (i) Y does not satisfy the necessary conditions to apply the relevant inequality, but (ii) Y is dominated by some random variable X that does.

We say that a random variable X *stochastically dominates* a random variable Y and write $X \succ Y$ if

$$\mathbb{P}(X \geq t) \geq \mathbb{P}(Y \geq t) \quad \text{for all real } t. \tag{2.37}$$

Clearly, we can use X as a surrogate for Y if (2.37) holds.

The following case arises quite often. Suppose that $Y = Y_1 + Y_2 + \cdots + Y_n$, where $Y_1, Y_2, \ldots, Y_n$ are not independent, but instead we have that for all t in the range $[A_i, B_i]$ of Y_i,

$$\mathbb{P}(Y_i \geq t \mid Y_1, Y_2, \ldots, Y_{i-1}) \leq \varphi(t),$$

where $\varphi(t)$ decreases monotonically from 1 to 0 in $[A_i, B_i]$.

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

Let X_i be a random variable taking values in the same range as Y_i and such that $\mathbb{P}(X_i \geq t) = \varphi(t)$. Let $X = X_1 + \cdots + X_n$, where $X_1, X_2, \ldots, X_n$ are independent of each other and $Y_1, Y_2, \ldots, Y_n$. Then we have

Lemma 2.29 *X stochastically dominates Y.*

Proof Let $X^{(i)} = X_1 + \cdots + X_i$ and $Y^{(i)} = Y_1 + \cdots + Y_i$ for $i = 1, 2, \ldots, n$. We will show by induction that $X^{(i)}$ dominates $Y^{(i)}$ for $i = 1, 2, \ldots, n$. This is trivially true for $i = 1$, and for $i > 1$ we have

$$\begin{aligned}\mathbb{P}(Y^{(i)} \geq t \mid Y_1, \ldots, Y_{i-1}) &= \mathbb{P}(Y_i \geq t - (Y_1 + \cdots + Y_{i-1}) \mid Y_1, \ldots, Y_{i-1}) \\ &\leq \mathbb{P}(X_i \geq t - (Y_1 + \cdots + Y_{i-1}) \mid Y_1, \ldots, Y_{i-1}).\end{aligned}$$

Removing the conditioning, we have

$$\mathbb{P}(Y^{(i)} \geq t) \leq \mathbb{P}(Y^{(i-1)} \geq t - X_i) \leq \mathbb{P}(X^{(i-1)} \geq t - X_i) = \mathbb{P}(X^{(i)} \geq t),$$

where the second inequality follows by induction. □

Exercises

2.3.1. Suppose we roll a fair die n times. Show that w.h.p. the number of odd outcomes is within $O(n^{1/2} \log n)$ of the number of even outcomes.

2.3.2. Consider the outcome of tossing a fair coin n times. Represent this by a (random) string of H's and T's. Show that w.h.p. there are $\sim n/8$ occurrences of HTH as a contiguous substring.

2.3.3. Check that (2.25) remains true for $x > 1$.
(Hint: differentiate both sides, twice.)

Problems for Chapter 2

2.1 Show that if $k = o(n)$, then

$$\binom{n}{k} \sim \left(\frac{ne}{k}\right)^k (2\pi k)^{-1/2} \exp\left\{-\frac{k^2}{2n}(1 + o(1))\right\}.$$

2.2 Let c be a constant, $0 < c < 1$, and let $k \sim cn$. Show that for such k,

$$\binom{n}{k} = 2^{n(H(c)+o(1))},$$

where H is an *entropy function*: $H(c) = -c \ln c - (1 - c) \ln(1 - c)$.

2.3 Prove the following strengthening of (2.2),

$$\sum_{\ell=0}^{k} \binom{n}{\ell} \leq \left(\frac{ne}{k}\right)^k.$$

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press

2.4 Let $f(n) = \sum_{k=1}^{n} \frac{1}{k} \prod_{j=0}^{k-1} \left(1 - \frac{j}{n}\right)$. Prove that $f(n) \sim \frac{1}{2} \log n$.

2.5 Suppose that $m = cn$ distinguishable balls are thrown randomly into n boxes. (i) Write down an expression for the expected number of boxes that contain k or more balls. (ii) Show that your expression tends to zero if $k = \lceil \log n \rceil$.

2.6 Suppose that $m = cn$ distinguishable balls are thrown randomly into n boxes. Suppose that box i contains b_i balls. (i) Write down an expression for the expected number of k-sequences such that $b_i = b_{i+1} = \cdots = b_{i+k-1} = 0$. (ii) Show that your expression tends to zero if $k = \lceil \log n \rceil$.

2.7 Suppose that we toss a fair coin. Estimate the probability that we have to make $(2 + \varepsilon)n$ tosses before we see n heads.

2.8 Let $X_1, X_2, \ldots, X_n$ be independent binary random variables, $X_i \in \{0, 1\}$, and let $\mathbb{P}(X_i = 1) = p$ for every $1 \leq i \leq n$, where $0 < p < 1$. Let $\overline{S}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$. Apply the Chernoff–Hoeffding bounds to show that if $n \geq (3/t^2) \ln(2/\delta)$, then $\mathbb{P}(|\overline{S}_n - p)| \leq t) \geq 1 - \delta$.

2.9 Let $Y_1, Y_2, \ldots, Y_m$ be independent non-negative integer random variables. Suppose that for $r \geq 1$ we have $\Pr(Y_r \geq k) \leq C\rho^k$, where $\rho < 1$. Let $\mu = C/(1 - \rho)$. Show that if $Y = Y_1 + Y_2 + \cdots + Y_m$, then

$$\Pr(Y \geq (1 + \varepsilon)\mu m) \leq e^{-B\varepsilon^2 m}$$

for $0 \leq \varepsilon \leq 1$ and some $B = B(C, \rho)$.

2.10 We say that a sequence of random variables $A_0, A_1, \ldots$ is (η, N)-bounded if $A_i - \eta \leq A_{i+1} \leq A_i + N$ for all $i \geq 0$.

(i) Suppose that $\eta \leq N/2$ and $a < \eta m$. Prove that if $0 = A_0, A_1, \ldots$ is an (η, N)-bounded sequence, then $\Pr(A_m \leq -a) \leq \exp\left\{-\frac{a^2}{3\eta m N}\right\}$.

(ii) Suppose that $\eta \leq N/10$ and $a < \eta m$. Prove that if $0 = A_0, A_1, \ldots$ is an (η, N)-bounded sequence, then $\Pr(A_m \geq a) \leq \exp\left\{-\frac{a^2}{3\eta m N}\right\}$.

2.11 Let A be an $n \times m$ matrix, with each $a_{ij} \in \{0, 1\}$, and let $\vec{b}$ be an m-dimensional vector, with each $b_k \in \{-1, 1\}$, where each possibility is chosen with probability $1/2$. Let $\vec{c}$ be the n-dimensional vector that denotes the product of A and $\vec{b}$. Applying the Chernoff–Hoeffding bound show that the following inequality holds for $i \in \{1, 2, \ldots, n\}$:

$$\mathbb{P}(\max\{|c_i|\} \geq \sqrt{4m \ln n}) \leq O(n^{-1}).$$

https://doi.org/10.1017/9781009260268.004 Published online by Cambridge University Press