Lecture 4: Inequalities and Asymptotic Estimates

We draw materials from [2, 5, 8–10, 17, 18]. Unless specified otherwise, we use μ, σ^2 to denote the mean and variance of the the variable under consideration. This note shall be updated throughout the seminar as I find more useful inequalities.

1 Basic inequalities

Theorem 1.1 (Markov's Inequality). *If X is a random variable taking only non-negative values, then for any $a > 0$*

$$\mathbf{Pr}[X \geq a] \leq \frac{\mathbf{E}[X]}{a}. \tag{1}$$

Proof. We show this for the discrete case only, the continuous case is similar. By definition, we have

$$\mathbf{E}[X] = \sum_x xp(x) = \sum_{x<a} xp(x) + \sum_{x \geq a} xp(x) \geq \sum_{x \geq a} ap(x) = a\mathbf{Pr}[X \geq a].$$

□

Intuitively, when $a \leq \mathbf{E}[X]$ the inequality is trivial. For $a > \mathbf{E}[X]$, it means the larger a is relative to the mean, the harder it is to have $X \geq a$. Thus, the inequality meets common sense. A slightly more intuitive form of (1) is

$$\mathbf{Pr}[X \geq a\mu] \leq \frac{1}{a}. \tag{2}$$

Theorem 1.2 (Chebyshev's Inequality). *If X is a random variable with mean μ and variance σ^2, then for any $a > 0$,*

$$\mathbf{Pr}\big[|X - \mu| \geq a\big] \leq \frac{\sigma^2}{a^2}. \tag{3}$$

Proof. This inequality makes a lot of sense. The probability that X is far from its mean gets smaller when X is further, and smaller when its variance is smaller. The proof is almost an immediate corollary of Markov's. Let $Z = (X - \mu)^2$, then $\mathbf{E}[Z] = \sigma^2$ by definition of variance. Since $|X - \mu| \geq a$ iff $Z \geq a^2$, applying Markov's inequality completes the proof. □

Again, there is a more intuitive way of writing (3):

$$\mathbf{Pr}\big[|X - \mu| \geq a\sigma\big] \leq \frac{1}{a^2}. \tag{4}$$

Theorem 1.3.

$$\mathbf{Pr}[X = 0] \leq \frac{\sigma^2}{\sigma^2 + \mu^2}. \tag{5}$$

Proof. We show this for the discrete case. The continuous case is shown similarly.

$$\mu^2 = \left(\sum_{x\neq 0} x\mathbf{Pr}[X=x]\right)^2 \leq \left(\sum_{x\neq 0} x^2\mathbf{Pr}[X=x]\right)\left(\sum_{x\neq 0}\mathbf{Pr}[X=x]\right) = \sigma^2\left(1-\mathbf{Pr}[X=0]\right).$$

$\square$

Theorem 1.4 (One-sided Chebyshev Inequality). *Let X be a random variable with $\mathbf{E}[X]=\mu$ and Var $[X]=\sigma^2$, then for any $a>0$,*

$$\mathbf{Pr}[X\geq\mu+a] \leq \frac{\sigma^2}{\sigma^2+a^2} \tag{6}$$

$$\mathbf{Pr}[X\leq\mu-a] \leq \frac{\sigma^2}{\sigma^2+a^2}. \tag{7}$$

Proof. Let $t\geq-\mu$ be a variable. Then, $Y=(X+t)^2$ has and

$$\mathbf{E}[Y]=\mathbf{E}[X^2]+2t\mu+t^2=\sigma^2+(t+\mu)^2.$$

Thus, by Markov's inequality we get

$$\mathbf{Pr}[X\geq\mu+a]\leq\mathbf{Pr}[Y\geq(\mu+a+t)^2]\leq\frac{\sigma^2+(t+\mu)^2}{(a+t+\mu)^2}.$$

The right most expression is minimized when $t=\sigma^2/a-\mu$, in which case it becomes $\sigma^2/(\sigma^2+a^2)$ as desired. The other inequality is proven similarly. $\square$

A twice-differentiable function f is *convex* if $f''(x)\geq 0$ for all x, and *concave* when $f''(x)\geq 0$ for all x.

Theorem 1.5 (Jenssen's inequality). *Let $f(x)$ be a convex function, then*

$$\mathbf{E}[f(X)]\geq f(E[X]). \tag{8}$$

The same result holds for multiple random variables.

Proof. Taylor's theorem gives

$$f(x)=f(\mu)+f'(\mu)(x-\mu)+f''(\xi)(x-\mu)^2/2,$$

where ξ is some number between x and μ. When $f(x)$ is convex, $f''(\xi)\geq 0$, which implies

$$f(x)\geq f(\mu)+f'(\mu)(x-\mu).$$

Consequently,

$$\mathbf{E}[f(X)]\geq f(\mu)+f'(\mu)\mathbf{E}[X-\mu]=f(\mu).$$

$\square$

2 Elementary Inequalities and Asymptotic Estimates

Fact 2.1. For $p \in [0,1]$, $(1-p) \leq e^{-p}$. The inequality is good for small p.

Fact 2.2. For any $x \in [-1,1]$, $(1+x) \leq e^{x}$. The inequality is good for small x.

The following theorem was shown by Robbins [16].

Theorem 2.3 (Stirling's approximation). *For each positive integer n, there is an α_n, where $\frac{1}{12n+1} < \alpha_n < \frac{1}{12n}$, such that*

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\alpha_n}. \tag{9}$$

We often find it useful to remember the asymptotic form of Stirling's approximation:

$$n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1+o(1)). \tag{10}$$

The following theorem follows from trivial applications of the Taylor's expansions for $\ln(1+t)$ and $\ln(1-t)$.

Theorem 2.4 (Estimates of $ln(1+t)$). *(a) If $t > -1$, then*

$$\ln(1+t) \leq \min\{t, t - \frac{1}{2}t^2 + \frac{1}{3}t^3\}. \tag{11}$$

(b)

$$\ln(1+t) > t - \frac{1}{2}t^2. \tag{12}$$

(c)

$$\ln(1+t) > t - \frac{1}{2}t^2 + \frac{1}{4}t^3. \tag{13}$$

(d)

$$\ln(1-t) > -t - t^2. \tag{14}$$

(e)

$$\ln(1-t) > -t - \frac{1}{2}t^2 - \frac{1}{2}t^3. \tag{15}$$

Lemma 2.5. *Let $\cosh(x) = (e^x + e^{-x})/2$, and $\sinh(x) = (e^x - e^{-x})/2$. Then for all reals α, x with $|\alpha| \leq 1$,*

$$\cosh(x) + \alpha \sinh(x) \leq e^{x^2/2 + \alpha x}. \tag{16}$$

Proof. This follows from elementary analysis. □

Corollary 2.6. *The following are often more useful than the general result above*

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

(ii) For all $p \in [0,1]$, and all t,

$$pe^{t(1-p)} + (1-p)e^{-tp} \leq e^{t^2/8} \tag{17}$$

Proof. Firstly, (i) follows from Lemma 2.5 by setting $\alpha = 0$, $t = x$. On the other hand, (ii) follows by setting $p = (1+\alpha)/2$ and $t = 2x$. □

3 Chernoff bounds

The following idea from Chernoff (1952, [6]) is infuential on showing many different "tail inequalities".

Theorem 3.1 (Chernoff bound). *Let X be a random variable with moment generating function $M(t) = \mathbf{E}[e^{tX}]$. Then,*

$$\mathbf{Pr}[X \geq a] \leq e^{-ta} M(t) \quad \textit{for all} \quad t > 0$$
$$\mathbf{Pr}[X \leq a] \leq e^{-ta} M(t) \quad \textit{for all} \quad t < 0.$$

Proof. The best bound can be obtained by minimizing the function on the right hand side. We show the first relation, the second is similar. When $t > 0$, by Markov's inequality we get

$$\mathbf{Pr}[X \geq a] = \mathbf{Pr}[e^{tX} \geq e^{ta}] \leq \mathbf{E}[e^{tX}] e^{-ta}.$$

□

Let us first consider a set of mutually independent Bernulli random variables $X_1, \ldots, X_n$, where $\mathbf{Pr}[X_i = 1] = p_i$, and $\mathbf{Pr}[X_i = 0] = 1 - p_i$, for $0 < p_i < 1$. Let $S_n = X_1 + \cdots + X_n$, then $\mu = \mathbf{E}[S_n] = p_1 + \cdots + p_n$. Note that when $p_i = p$, S_n has the usual Binomial distribution Binomial(n, p).

Theorem 3.2. *Under the above assumptions, for any $a > 0$,*

$$\mathbf{Pr}[S_n \geq a] \leq e^{-ta} \left(1 + p(e^t - 1)\right)^n. \tag{18}$$

Proof. The proof makes use of Chernoff's idea: for any $t > 0$, Markov's inequality gives

$$\mathbf{Pr}[S_n \geq a] = \mathbf{Pr}[e^{tS_n} \geq e^{ta}] \leq e^{-ta}\mathbf{E}[e^{tS_n}] = e^{-ta}\mathbf{E}[e^{tX_1 + \cdots + tX_n}] = e^{-ta}\mathbf{E}[e^{tX_i}] \ldots \mathbf{E}[e^{tX_n}]. \tag{19}$$

Note that the independence assumption is crucial. On the other hand,

$$f(p_i) := \ln(\mathbf{E}[e^{tX_i}]) = \ln(p_i e^t + (1 - p_i)) = \ln(1 + p_i(e^t - 1))$$

is concave in p_i, which - by Jensen's inequality - implies

$$\sum_{i=1}^{n} \ln(\mathbf{E}[e^{tX_i}]) \leq n \ln(1 + p(e^t - 1)).$$

Exponentiating both sides and recall inequality (19), we get

$$\mathbf{Pr}[S_n \geq a] \leq e^{-ta} \left(1 + p(e^t - 1)\right)^n,$$

as desired. □

Theorem 3.3. *Let $X_1, \ldots, X_n$ be mutually independent random variables with $|X_i| \leq c_i$ and $\mathbf{E}[X_i] = 0$, where $c_i > 0$ is a function on i. Let $S = X_1 + \cdots + X_n$, then*

$$\mathbf{Pr}[S \geq a] \leq \exp\left(-\frac{a^2}{2(c_1^2 + \cdots + c_n^2)}\right). \tag{20}$$

Proof. For any $t > 0$, Chernoff's bound gives

$$\mathbf{Pr}[S \geq a] \leq e^{-ta}\mathbf{E}[e^{tS}] = e^{-ta}\mathbf{E}[e^{tX_1+\cdots+tX_n}] = e^{-ta}\mathbf{E}[e^{tX_1}]\dots\mathbf{E}[e^{tX_n}].$$

Note that for $x \in [-c, c]$, we have $e^{tx} \leq f(x)$, where

$$f(x) = \frac{e^{ct}+e^{-ct}}{2} + \frac{e^{ct}-e^{-ct}}{2c}x = \cosh{(ct)} + x\sinh{(ct)}.$$

To see $e^{tx} \leq f(x)$, note that $y = f(x)$ is the chord through the points $x = -c, x = c$ of the convex curve $y = e^{tx}$. Thus,

$$\mathbf{E}[e^{tX_i}] \leq \mathbf{E}[f(X_i)] = f(\mathbf{E}[X_i]) = f(0) = \cosh(c_i t) \leq e^{(c_i t)^2/2}.$$

Consequently,

$$\mathbf{Pr}[S \geq a] \leq e^{-ta}e^{(c_1^2+\cdots+c_n^2)t^2/2}.$$

Pick $t = a/(\sum_i c_i^2)$ to minimize the right hand side, we get the desired result. □

4 Martingale Tail Inequalities

Theorem 4.1 (Kolmogorov-Doob Inequality). *Let $X_0, X_1, \dots$ be a martingale sequence. Then, for any $a > 0$,*

$$\mathbf{Pr}[\max_{0\leq i\leq n} X_i \geq a] \leq \frac{\mathbf{E}[|X_n|]}{a}. \tag{21}$$

Proof. TBD. □

The following result was shown by Hoeffding (1963, [12]) and Azuma (1967, [3]).

Theorem 4.2 (Hoeffding-Azuma Inequality). *Let $X_0, \dots, X_n$ be a martingale sequence such that for each $k = 1, \dots, n$,*

$$|X_k - X_{k-1}| \leq c_k, \tag{22}$$

where c_k is a function on k. Then, for all $m \geq 0$, $a > 0$,

$$\mathbf{Pr}[|X_m - X_0| \geq a] \leq 2\exp\left(\frac{-a^2}{2\sum_{k=1}^m c_k^2}\right) \tag{23}$$

Condition (22) on a martingale sequence is often called the *Lipschitz condition*.

Proof. Let $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_n$ be a *filtration* corresponding to the martingale sequence, i.e.

$$\mathbf{E}[X_k \mid \mathcal{F}_{k-1}] = X_{k-1}, \text{ or } \mathbf{E}[X_k - X_{k-1} \mid \mathcal{F}_{k-1}] = 0.$$

Note also that X_i is $\mathcal{F}_j$-measurable for all $j \geq i$, i.e. X_i is constant on the elementary events of F_j. Hence, for any function f on X_i, we have $\mathbf{E}[f(X_i) \mid F_j] = f(X_i)$ for all $j \geq i$.

For $k = 1, \dots, n$, let $Y_k = X_k - X_{k-1}$. Then, $X_m - X_0 = Y_1 + \cdots + Y_m$ and $|Y_k| \leq c_k$. It is easy to see that, for any $t > 0$,

$$\mathbf{E}[e^{tY_1+\cdots+tY_m}] = \mathbf{E}\left[e^{tY_1+\cdots+tY_{m-1}}\mathbf{E}[e^{tY_m} \mid \mathcal{F}_{m-1}]\right].$$

We first try to bound the upper tail, proceeding in the same way as in the proof of Theorem 3.3. For any $t > 0$, Chernoff bound gives

$$\begin{aligned}
\mathbf{Pr}[Y_1 + \cdots + Y_m \geq a] &\leq e^{-ta}\mathbf{E}[e^{tY_1+\cdots+tY_m}] \\
&= e^{-ta}\mathbf{E}\left[e^{tY_1+\cdots+tY_{m-1}}\mathbf{E}[e^{tY_m} \mid \mathcal{F}_{m-1}]\right] \\
&\leq e^{-ta}e^{c_m^2 t^2/2}\mathbf{E}\left[e^{tY_1+\cdots+tY_{m-1}}\right] \\
&\leq e^{-ta}e^{(c_1^2+\cdots+c_m^2)t^2/2}.
\end{aligned}$$

The rest is the same as in Theorem 3.3. We get half of the right hand side of (23). To show the same upper bound for $\mathbf{Pr}[X_m - X_0 \leq -a]$, we can just let $Y_k = X_{k-1} - X_k$. □

We next develop two more general versions of tail inequalities for martingales, one comes from Maurey (1979, [15]), the other from Alon-Kim-Spencer (1997, [1]).

Let A, B be finite sets, A^B denote the set of all mappings from B into A. (It might be instructive to try to explain the choice of the notation A^B on your own.) For example, if B is a set of edges of a graph G, and $A = \{0, 1\}$, then A^B can be thought of as the set of all spanning subgraphs of G.

Now, let $\Omega = A^B$, and define a measure on Ω by giving values p_{ab} and, for each $g \in A^B$, define

$$\mathbf{Pr}[g(b) = a] = p_{ab},$$

where $g(b)$ are mutually independent.

Fix a gradation $\emptyset = B_0 \subset B_1 \subset \cdots \subset B_m = B$. (In the simplest case, $|B_i - B_{i-1}| = 1$, $m = |B|$, and thus the gradation defines a total order on B.) The gradation induces a filtration $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_m$ on Ω, where the elementary events of $\mathcal{F}_i$ are sets of functions from B into A whose restrictions on B_i are identical. Thus, there are $|A|^{|}B_i|$ elementary events for $\mathcal{F}_i$, each correspond to a distinct element of A^{B_i}.

To this end, let $L : A^B \to \mathbb{R}$ be a functional (like χ, ω, α in the $G(n, p)$ case), which could be thought of as a random variable on Ω. The sequence $X_i = \mathbf{E}[L \mid \mathcal{F}_i]$ is a martingale. It is easy to see that $X_0 = \mathbf{E}[L]$ and $X_m = L$.

Definition 4.3. The functional L is said to satisfy the *Lipschitz condition* relative to the gradation if, $\forall k \in [m]$,

$$g \text{ and } h \text{ differ only on } B_k - B_{k-1} \text{ implies } |L(g) - L(h)| \leq c_k,$$

where c_k is a function of k.

The following lemma helps generalize Hoeffding-Azuma's inequality.

Lemma 4.4. *Let L satisfy Lipschitz condition, then the corresponding martingale satisfies*

$$|X_k - X_{k-1}| \leq c_k, \quad \forall k \in [m].$$

Proof. TBD. □

Corollary 4.5 (Generalized Hoeffding-Azuma Inequality). *In the setting of Lemma 4.4, let $\mu = \mathbf{E}[L]$. Then, for all $a > 0$,*

$$\mathbf{Pr}[L \geq \mu + a] \leq \exp\left(\frac{-a^2}{2\sum_{k=1}^m c_k^2}\right), \tag{24}$$

and

$$\mathbf{Pr}[L \leq \mu - a] \leq \exp\left(\frac{-a^2}{2\sum_{k=1}^m c_k^2}\right), \tag{25}$$

Proof. This follows directly from Lemmas 4.4 and 4.2. □

5 Lovász Local Lemma

Let $A_1, \dots, A_n$ be events on an arbitrary probability space. A directed graph $G = (V, E)$ with $V = [n]$ is called a *dependency digraph* for $A_1, \dots, A_n$ if each A_i is independent from the set of events $\{A_j \mid (i, j) \notin E\}$. (In other words, A_i is *at most* dependent on its neighbors.) The following lemma, often referred to as Lovász Local Lemma, was originally shown in Erdős and Lovász (1975, [8]). The lemma is very useful when showing a certain event has positive probability, albeit exponentially small. It is most useful when the dependency digraph has small maximum degree.

Lemma 5.1 (Lovász Local Lemma). *Let $G = (V, E)$ be a dependency digraph for the events $A_1, \dots, A_n$. Suppose there are real numbers $\alpha_1, \dots, \alpha_n$, such that $0 \leq \alpha_i < 1$, $\forall i$, and*

$$\mathbf{Pr}[A_i] \leq \alpha_i \prod_{j:(i,j)\in E} (1 - \alpha_j).$$

Then,

(a) For all $S \subset [n]$, $|S| = s < n$, and any $i \notin S$,

$$\mathbf{Pr}\left[A_i \mid \bigwedge_{j\in S} \bar{A}_j\right] \leq \alpha_i. \tag{26}$$

(b) Moreover, the probability that none of the A_i happens is positive. In particular

$$\mathbf{Pr}\left[\bigwedge_{i=1}^{n} \bar{A}_i\right] \geq \prod_{i=1}^{n}(1 - \alpha_i). \tag{27}$$

Proof. Firstly, we show that (a) implies (b). This follows as

$$\begin{aligned}
\mathbf{Pr}\left[\bigwedge_{i=1}^{n} \bar{A}_i\right] &= \mathbf{Pr}[\bar{A}_1] \cdot \mathbf{Pr}[\bar{A}_2 \mid \bar{A}_1] \dots \mathbf{Pr}[\bar{A}_n \mid \wedge_{j=1}^{n-1} \bar{A}_j] \\
&= (1 - \mathbf{Pr}[A_1])(1 - \mathbf{Pr}[A_2 \mid \bar{A}_1]) \dots (1 - \mathbf{Pr}[A_n \mid \wedge_{j=1}^{n-1} \bar{A}_j]) \\
&= (1 - \alpha_1)(1 - \alpha_2) \dots (1 - \alpha_n).
\end{aligned}$$

To show (a), we induct on $s = |S|$. There is nothing to do for $s = 0$. For $s \geq 1$, assume that (26) holds for all $|S| \leq s - 1$. Consider some S with $|S| = s \geq 1$. Let $D_i = \{j \in S \mid (i, j) \in E\}$, and $\bar{D}_i = S - D_i$. We have

$$\begin{aligned}
\mathbf{Pr}\left[A_i \mid \bigwedge_{j\in S} \bar{A}_j\right] &= \mathbf{Pr}\left[A_i \mid \left(\bigwedge_{j\in D_i} \bar{A}_j\right) \wedge \left(\bigwedge_{j\in \bar{D}_i} \bar{A}_j\right)\right] \\
&= \frac{\mathbf{Pr}\left[A_i \wedge \left(\bigwedge_{j\in D_i} \bar{A}_j\right) \mid \bigwedge_{j\in \bar{D}_i} \bar{A}_j\right]}{\mathbf{Pr}\left[\bigwedge_{j\in D_i} \bar{A}_j \mid \bigwedge_{j\in \bar{D}_i} \bar{A}_j\right]}.
\end{aligned}$$

We first bound the numerator:

$$\mathbf{Pr}\left[A_i \wedge \left(\bigwedge_{j\in D_i} \bar{A}_j\right) \mid \bigwedge_{j\in \bar{D}_i} \bar{A}_j\right] \leq \mathbf{Pr}\left[A_i \mid \bigwedge_{j\in \bar{D}_i} \bar{A}_j\right] = \mathbf{Pr}[A_i] \leq \alpha_i \prod_{j:(i,j)\in E} (1 - \alpha_j).$$

Next, the denominator (which would be 1 if $\wedge_{j\in\bar{D}_i}\bar{A}_j = \emptyset$) can be bounded with induction hypothesis. Suppose $D_i = \{j_1, \ldots j_k\}$, then

$$\begin{aligned}
&\mathbf{Pr}\left[\bigwedge_{j\in D_i}\bar{A}_j \mid \bigwedge_{j\in\bar{D}_i}\bar{A}_j\right] \\
= \ & \left(1-\mathbf{Pr}\left[A_{j_1} \mid \bigwedge_{j\in\bar{D}_i}\bar{A}_j\right]\right)\left(1-\mathbf{Pr}\left[A_{j_2} \mid \bigwedge_{j\in\bar{D}_i\cup\{j_1\}}\bar{A}_j\right]\right)\ldots \\
& \ldots\left(1-\mathbf{Pr}\left[A_{j_k} \mid \bigwedge_{j\in\bar{D}_i\cup D_i-\{j_k\}}\bar{A}_j\right]\right) \\
\geq \ & \prod_{j\in D_i}(1-\alpha_j) \\
\geq \ & \prod_{j:(i,j)\in E}(1-\alpha_j).
\end{aligned}$$

□

As we have mentioned earlier, the Local Lemma is most useful when the maximum degree of a dependency graph is small. We now give a particular version of the Lemma which helps us make use of this observation:

Corollary 5.2 (Local Lemma; Symmetric Case). *Suppose each event A_i is independent of all others except for at most Δ (i.e. the dependency graph has maximum degree at most Δ), and that $\mathbf{Pr}(A_i) \leq p$ for all $i = 1\ldots, n$.*

If

$$ep(\Delta+1) \leq 1, \tag{28}$$

then $\mathbf{Pr}(\wedge_{i=1}^n \bar{A}_i) > 0$.

Proof. The case $\Delta = 0$ is trivial. Otherwise, take $\alpha_i = 1/(\Delta+1)$ (which is < 1) in the Local Lemma, we have

$$P[A_i] \leq p \leq \frac{1}{\Delta+1}\frac{1}{e} \leq \alpha_i\left(1-\frac{1}{\Delta+1}\right)^{\Delta} \leq \alpha_i \prod_{j:(i,j)\in E}(1-\alpha_j).$$

Here, we have used the fact that for $\Delta \geq 1$, $\left(1-\frac{1}{\Delta+1}\right)^{\Delta} > 1/e$, which follows from (14) with $t = 1/(\Delta+1) \leq 0.5$. □

For applications of Lovász Local Lemma and its algorithmic aspects, see Beck [4] and others [7, 11, 13, 14]

References

[1] N. Alon, J.-H. Kim, and J. Spencer, *Nearly perfect matchings in regular simple hypergraphs*, Israel J. Math., 100 (1997), pp. 171–187.

[2] N. Alon and J. H. Spencer, *The probabilistic method*, Wiley-Interscience Series in Discrete Mathematics and Optimization, Wiley-Interscience [John Wiley & Sons], New York, second ed., 2000. With an appendix on the life and work of Paul Erdős.

[3] K. Azuma, *Weighted sums of certain dependent random variables*, Tôhoku Math. J. (2), 19 (1967), pp. 357–367.

[4] J. BECK, *An algorithmic approach to the Lovász local lemma. I*, Random Structures Algorithms, 2 (1991), pp. 343–365.

[5] B. BOLLOBÁS, *Random graphs*, vol. 73 of Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, second ed., 2001.

[6] H. CHERNOFF, *A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations*, Ann. Math. Statistics, 23 (1952), pp. 493–507.

[7] A. CZUMAJ AND C. SCHEIDELER, *Coloring nonuniform hypergraphs: a new algorithmic approach to the general Lovász local lemma*, in Proceedings of the Ninth International Conference "Random Structures and Algorithms" (Poznan, 1999), vol. 17, 2000, pp. 213–237.

[8] P. ERDŐS AND L. LOVÁSZ, *Problems and results on* 3*-chromatic hypergraphs and some related questions*, in Infinite and finite sets (Colloq., Keszthely, 1973; dedicated to P. Erdős on his 60th birthday), Vol. II, North-Holland, Amsterdam, 1975, pp. 609–627. Colloq. Math. Soc. János Bolyai, Vol. 10.

[9] W. FELLER, *An introduction to probability theory and its applications. Vol. I*, John Wiley & Sons Inc., New York, 1968.

[10] ———, *An introduction to probability theory and its applications. Vol. II.*, John Wiley & Sons Inc., New York, 1971.

[11] V. GURUSWAMI, J. HÅSTAD, AND M. SUDAN, *Hardness of approximate hypergraph coloring*, SIAM J. Comput., 31 (2002), pp. 1663–1686 (electronic).

[12] W. HOEFFDING, *Probability inequalities for sums of bounded random variables*, J. Amer. Statist. Assoc., 58 (1963), pp. 13–30.

[13] M. KRIVELEVICH AND V. H. VU, *Choosability in random hypergraphs*, J. Combin. Theory Ser. B, 83 (2001), pp. 241–257.

[14] T. LEIGHTON, C.-J. LU, S. RAO, AND A. SRINIVASAN, *New algorithmic aspects of the local lemma with applications to routing and partitioning*, SIAM J. Comput., 31 (2001), pp. 626–641 (electronic).

[15] B. MAUREY, *Construction de suites symétriques*, C. R. Acad. Sci. Paris Sér. A-B, 288 (1979), pp. A679–A681.

[16] H. ROBBINS, *A remark onStirling's formula*, Amer. Math. Monthly, 62 (1955), pp. 26–29.

[17] S. ROSS, *A first course in probability*, Macmillan Co., New York, second ed., 1984.

[18] S. M. ROSS, *Introduction to probability models*, Harcourt/Academic Press, San Diego, CA, seventh ed., 2000.