KHINTCHINE'S INEQUALITIES

Consider the following experiment. You are given N real numbers $s_1, s_2, \ldots, s_N$, and a fair coin. Flip the coin. If it comes up heads, move a pointer to s_1 on the number line. If it comes up tails, move the pointer to $-s_1$. In general, if you've flipped the coin k times (where $k < N$), and the pointer is at position σ_k, you flip the coin again, and move the pointer to

$$\sigma_{k+1} := \sigma_k + s_{k+1}$$

if the coin comes up heads, and move it to

$$\sigma_{k+1} := \sigma_k - s_{k+1}$$

if the coin comes up tails. You stop after N coin flips.

This is a one-dimensional random walk with *varying* step size.

Question: What is the expected value of

$$|\sigma_N| = \left|\sum_1^N \pm s_k\right|? \tag{1}$$

Khintchine's Inequalities give useful bounds for (1). They assert that there are positive constants c_1 and C_1 (which can be estimated!) such that, for *all* N and all such sequences $\{s_k\}_1^N$,

$$c_1 \left(\sum_1^N s_k^2\right)^{1/2} \leq E\left(|\sigma_N|\right) \leq C_1 \left(\sum_1^N s_k^2\right)^{1/2},$$

where we are using $E(\cdot)$ to mean "expected value." But they go further. It turns out that, if $0 < p < \infty$, there are positive constants c_p and C_p such that, for all N and all such sequences,

$$c_p \left(\sum_1^N s_k^2\right)^{1/2} \leq \left(E\left(|\sigma_N|^p\right)\right)^{1/p} \leq C_p \left(\sum_1^N s_k^2\right)^{1/2}. \tag{2}$$

Something you're probably wondering: Why a sum of squares? Why not have the expected values controlled by

$$\left(\sum_1^N s_k^4\right)^{1/4}$$

or something even more exotic? The *numerical* reason for this turns out to be the simple inequality:

$$\cosh(x) \leq \exp(x^2/2). \tag{3}$$

We will give a self-contained proof of (2) (which will include showing how (3) comes in). Our presentation will assume that the audience has some acquaintance with (but, fortunately, no detailed memory of) measure theory. If time permits we will show what (2) implies about the stability of Fourier and other expansions of functions with respect to small errors in the coefficients.

Lemma. *Suppose that* $(\Omega, \mathcal{M}, \mu)$ *is a probability space and* $f : \Omega \mapsto \mathbf{C}$ *is measurable. If* $0 < p < \infty$ *then*

$$\int_\Omega |f(x)|^p \, d\mu(x) = p \int_0^\infty \lambda^{p-1} \mu(\{x \in \Omega : \ |f(x)| > \lambda\}) \, d\lambda.$$

Quick proof: Since $0 < p < \infty$,

$$|f(x)|^p = p \int_0^{|f(x)|} \lambda^{p-1} \, d\lambda.$$

Therefore

$$\begin{aligned}
\int_\Omega |f(x)|^p \, d\mu(x) &= \int_\Omega \left(p \int_{\{\lambda : 0 < \lambda < |f(x)|\}} \lambda^{p-1} \, d\lambda \right) d\mu(x) \\
&= \int_{\{\lambda : 0 < \lambda < \infty\}} p\lambda^{p-1} \left(\int_{\{x \in \Omega : |f(x)| > \lambda\}} d\mu(x) \right) d\lambda \\
&= p \int_0^\infty \lambda^{p-1} \mu(\{x \in \Omega : \ |f(x)| > \lambda\}) \, d\lambda,
\end{aligned}$$

where the interchange of integrals is justified by Fubini-Tonelli.

REMARK: The formula holds in arbitrary measure spaces, but requires a different proof.

MORAL: An asymptotic bound on

$$\mu(\{x \in \Omega : \ |f(x)| > \lambda\})$$

implies a bound on

$$\int_\Omega |f(x)|^p \, d\mu(x).$$

In particular, if

$$\mu(\{x \in \Omega : \ |f(x)| > \lambda\}) \le 2 \exp(-\lambda^2/2)$$

for all $\lambda > 0$ then, for all $0 < p < \infty$,

$$\int_\Omega |f(x)|^p \, d\mu(x) \le C_p,$$

for some finite constant C_p that only depends on p.

Our Ω will be $[0, 1]$, μ will be m, Lebesgue measure, and $\mathcal{M}$ will be the Lebesgue measurable sets.

Our first step in proving Khintchine's Inequalities will be to show that if

$$\sigma_N = \sum_1^N \pm s_k$$

and

$$\sum_1^N s_k^2 = 1$$

then, for all $\lambda > 0$, the probability (measure) of the set where $|\sigma_N| > \lambda$ is less than or equal to $2\exp(-\lambda^2/2)$. *This bound does not depend on the particular sequence* $\{s_k\}$.

We make our notion of random ± 1's concrete (and calculable) by means of the Rademacher functions $r_n(x)$. The first Rademacher function, $r_1(x)$, is defined on $[0,1]$ by

$$r_1(x) \equiv \begin{cases} 1 & \text{if } 0 < x < 1/2; \\ -1 & \text{if } 1/2 < x < 1; \\ 0 & \text{otherwise;} \end{cases}$$

and is defined on the rest of $\mathbf{R}$ by making it periodic, with period 1: $r_1(x+1) = r_1(x)$. The n^{th} Rademacher function, $r_n(x)$, is defined by

$$r_n(x) \equiv r_1(2^{n-1}x).$$

We will only be looking at the Rademacher functions on $[0,1]$.

The Rademacher functions are orthonormal in $L^2[0,1]$. More than that, they are *independent random variables*. For those of you whose parents didn't force you to take a rigorous probability course, this means (in our particular situation) that if $\vec{\epsilon} \equiv \{\epsilon_k\}_1^P$ is any sequence of ± 1's, and $n_1 < n_2 < \cdots < n_P$, then the Lebesgue measure of

$$\cap_1^P \{x \in [0,1] : \ r_{n_k}(x) = \epsilon_k\}$$

equals

$$\prod_1^P m\left(\{x \in [0,1] : \ r_{n_k}(x) = \epsilon_k\}\right) = 2^{-P}. \tag{4}$$

(It's not hard to show that if B_1, B_2, $\ldots$, B_P is *any* sequence of Borel sets, then the Lebesgue measure of

$$\cap_1^P \{x \in [0,1] : \ r_{n_k}(x) \in B_k\}$$

equals

$$\prod_1^P m\left(\{x \in [0,1] : \ r_{n_k}(x) \in B_k\}\right),$$

which is the *formal* definition of probabilistic independence.)

The expected value of $|\sigma_N|^p$ equals

$$\int_0^1 \left|\sum_1^N s_k r_k(x)\right|^p \, dx,$$

because σ_N and

$$f(x) \equiv \sum_1^N s_k r_k(x) \tag{5}$$

have exactly the same distribution: For any number R, the probability that $\sigma_N = R$ equals the measure of the set in $[0,1]$ where $f(x) = R$. Our "first step" in proving Khintchine's Inequalities can be rephrased: If

$$\sum_1^N s_k^2 = 1$$

and f is defined as in (5), then, for all $\lambda > 0$,

$$m\left(\{x \in [0,1] : \ |f(x)| > \lambda\}\right) \le 2\exp(-\lambda^2/2).$$

Lemma. *If $f_k : \mathbf{R} \mapsto \mathbf{R}$ $(1 \le k \le P)$ are continuous and $n_1 < n_2 < \cdots < n_P$, then*

$$\int_0^1 \left(\prod_1^P f_k(r_{n_k}(x))\right) dx = \prod_1^P \left(\int_0^1 f_k(r_{n_k}(x))\, dx\right). \tag{6}$$

Remark. This is a consequence of the Rademacher functions' probabilistic independence; a much more general result is true. But this easy, special case will serve our purpose.

Proof. Since

$$\int_0^1 f_k(r_{n_k}(x))\, dx = \frac{f_k(1) + f_k(-1)}{2},$$

the right-hand side of (6) equals

$$2^{-P} \sum_{\vec{\epsilon} \in \{-1,1\}^P} \prod_1^P f_k(\epsilon_k),$$

where $\vec{\epsilon} = (\epsilon_1, \epsilon_2, \ldots, \epsilon_P)$ denotes an arbitrary sequence of ± 1's. But, because of (4), that's also the left-hand side of (6). Done!

We're going to apply the preceding lemma to the functions $f_k(x) \equiv \exp(t_k x)$, where t_k's are real numbers, to be specified presently. The lemma implies that

$$\begin{aligned}\int_0^1 \left(\prod_1^N f_k(r_k(x))\right) dx &= \int_0^1 \exp\left(\sum_1^N t_k r_k(x)\right) dx \\ &= \prod_1^N \left(\int_0^1 \exp\left(t_k r_k(x)\right)\, dx\right).\end{aligned}$$

But

$$\int_0^1 \exp\left(t_k r_k(x)\right)\, dx = \frac{e^{t_k} + e^{-t_k}}{2} = \cosh(t_k),$$

which is less than or equal to $\exp(t_k^2/2)$ (look at their power series). THEREFORE WE HAVE THE FOLLOWING

Theorem. *If $\{t_k\}_1^N$ is any finite sequence of real numbers, then*

$$\int_0^1 \exp\left(\sum_1^N t_k r_k(x)\right)\, dx \le \exp\left((1/2)\sum_1^N t_k^2\right).$$

Now suppose that f is as defined in (5), with $\sum_1^N s_k^2 = 1$. For $\lambda > 0$, set $t_k = \lambda s_k$. The preceding theorem implies that

$$\begin{aligned}\int_0^1 \exp(\lambda f(x))\, dx &= \int_0^1 \exp\left(\sum_1^N t_k r_k(x)\right)\, dx \\ &\le \exp\left((1/2)\sum_1^N t_k^2\right) \\ &= \exp(\lambda^2/2).\end{aligned}$$

But we also have that

$$\exp(\lambda^2) m\left(\{x \in [0,1] :\ f(x) > \lambda\}\right) \le \int_0^1 \exp(\lambda f(x))\, dx.$$

Therefore

$$m\left(\{x \in [0,1] :\ f(x) > \lambda\}\right) \le \exp(\lambda^2/2)\exp(-\lambda^2) = \exp(-\lambda^2/2).$$

When we apply the same argument to $-f$, we get

$$m\left(\{x \in [0,1] :\ f(x) < -\lambda\}\right) \le \exp(-\lambda^2/2).$$

Combining the two yields

$$m\left(\{x \in [0,1] :\ |f(x)| > \lambda\}\right) \le 2\exp(-\lambda^2/2),$$

as desired.

WE HAVE COMPLETED STEP ONE.

Now we know that if $\sum_1^N s_k^2 = 1$ and $0 < p < \infty$ then

$$\left(\int_0^1 \left|\sum_1^N s_k r_k(x)\right|^p dx\right)^{1/p} \leq C_p$$

for some constant C_p that only depends on p. But this is the same as

$$\left(\int_0^1 \left|\sum_1^N s_k r_k(x)\right|^p dx\right)^{1/p} \leq C_p \left(\sum_1^N s_k^2\right)^{1/2}, \tag{7}$$

because of how we normalized the s_k's. Inequality (7) remains true if we multiply the s_k's by any constant R. Therefore it will be true for *every* finite sequence $\{s_k\}_1^N$. ("True by homogeneity.") We have proved **one-half** of Khintchine's Inequalities.

THE OTHER DIRECTION?

If $p \geq 2$ then

$$\left(\sum_1^N s_k^2\right)^{1/2} = \left(\int_0^1 \left|\sum_1^N s_k r_k(x)\right|^2 dx\right)^{1/2} \leq \left(\int_0^1 \left|\sum_1^N s_k r_k(x)\right|^p dx\right)^{1/p}$$

by Hölder's Inequality. Therefore we only need to worry about the other direction when $0 < p < 2$.

Without loss of generality, we can assume that not ALL of the s_k's are 0.

We write:

$$\begin{aligned} \sum_1^N s_k^2 &= \int_0^1 \left|\sum_1^N s_k r_k(x)\right|^2 dx \\ &= \int_0^1 \left|\sum_1^N s_k r_k(x)\right|^{p/2} \left|\sum_1^N s_k r_k(x)\right|^{2-p/2} dx \\ &\leq \left(\int_0^1 \left|\sum_1^N s_k r_k(x)\right|^p dx\right)^{1/2} \left(\int_0^1 \left|\sum_1^N s_k r_k(x)\right|^{4-p} dx\right)^{1/2}, \end{aligned}$$

where the last line follows from the Cauchy-Schwarz inequality.

Because of (7), we know that

$$\left(\int_0^1 \left|\sum_1^N s_k r_k(x)\right|^{4-p} dx\right)^{1/2} \leq C_{4-p}^{2-p/2} \left(\sum_1^N s_k^2\right)^{1-p/4}.$$

Therefore

$$\sum_1^N s_k^2 \le \left(\int_0^1 \left| \sum_1^N s_k r_k(x) \right|^p dx \right)^{1/2} C_{4-p}^{2-p/2} \left(\sum_1^N s_k^2 \right)^{1-p/4} .$$

Dividing both sides by

$$\left(\sum_1^N s_k^2 \right)^{1-p/4}$$

yields

$$\left(\sum_1^N s_k^2 \right)^{p/4} \le C_{4-p}^{2-p/2} \left(\int_0^1 \left| \sum_1^N s_k r_k(x) \right|^p dx \right)^{1/2} .$$

Raising both sides to the power $2/p$ finishes the proof. QED.