Khinchin's inequality and Etemadi's inequality

Jordan Bell
jordan.bell@gmail.com
Department of Mathematics, University of Toronto

September 4, 2015

1 Khinchin's inequality

We will use the following to prove Khinchin's inequality.[1]

Lemma 1. *Let* $X_1, \ldots, X_n$ *be independent random variables each with the Rademacher distribution. For* $a_1, \ldots, a_n \in \mathbb{R}$ *and* $\lambda > 0$,

$$P\left(|S_n| > \lambda \left(\sum_{k=1}^n a_k^2 \right)^{1/2} \right) \leq 2e^{-\lambda^2/2},$$

where

$$S_n = \sum_{k=1}^n a_k X_k.$$

Proof. For $t \in \mathbb{R}$,

$$E(e^{ta_k X_k}) = \int_{\mathbb{R}} e^{ta_k x} d\left(\frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1\right)(x) = \frac{1}{2}(e^{-ta_k} + e^{ta_k}) = \cosh(ta_k).$$

Because the X_k are independent,

$$E(e^{tS_n}) = \prod_{k=1}^n E(e^{ta_k X_k}) = \prod_{k=1}^n \cosh(ta_k),$$

and because $\cosh x \leq e^{x^2/2}$ for all $x \in \mathbb{R}$, we have

$$E(e^{tS_n}) \leq \prod_{k=1}^n e^{\frac{t^2 a_k^2}{2}} = \exp\left(\frac{t^2}{2} \sum_{k=1}^n a_k^2\right).$$

[1] Camil Muscalu and Wilhelm Schlag, *Classical and Multilinear Harmonic Analysis*, volume I, p. 113, Lemma 5.4.

Let $\sigma^2 = \sum_{k=1}^n a_k^2$, with which

$$E(e^{tS_n}) \leq \exp\left(\frac{t^2\sigma^2}{2}\right).$$

Because $t \mapsto e^{\lambda\sigma t}$ is nonnegative and nondecreasing, for $t > 0$ we have

$$1_{S_n > \lambda\sigma} e^{\lambda\sigma t} < e^{tS_n},$$

which yields $P(S_n > \lambda\sigma) \leq e^{-\lambda\sigma t} E(e^{tS_n})$, and hence

$$P(S_n > \lambda\sigma) \leq e^{-\lambda\sigma t} \exp\left(\frac{t^2\sigma^2}{2}\right) = \exp\left(-\lambda\sigma t + \frac{t^2\sigma^2}{2}\right).$$

The minimum of the right-hand side occurs when $\lambda\sigma = t\sigma^2$, i.e. $t = \frac{\lambda}{\sigma}$, at which

$$P(S_n > \lambda\sigma) \leq \exp\left(-\lambda^2 + \frac{\lambda^2}{2}\right) = e^{-\lambda^2/2}.$$

For $t > 0$,

$$1_{S_n < -\lambda\sigma} e^{\lambda\sigma t} < e^{-tS_n},$$

which yields $P(S_n < -\lambda\sigma) \leq e^{-\lambda\sigma t} E(e^{-tS_n})$, and hence

$$P(S_n < -\lambda\sigma) \leq e^{-\lambda\sigma t} \exp\left(\frac{(-t)^2\sigma^2}{2}\right) = \exp\left(-\lambda\sigma t + \frac{t^2\sigma^2}{2}\right),$$

whence

$$P(S_n < -\lambda\sigma) \leq e^{-\lambda^2/2}.$$

Therefore

$$P(|S_n| > \lambda\sigma) = P(S_n > \lambda\sigma) + P(S_n < -\lambda\sigma) \leq 2e^{-\lambda^2/2},$$

proving the claim. □

Corollary 2. *Let $X_1, \ldots, X_n$ be independent random variables each with the Rademacher distribution. For $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ and $\lambda > 0$,*

$$P\left(|S_n| > \lambda \left(\sum_{k=1}^n |\alpha_k|^2\right)^{1/2}\right) \leq 4e^{-\lambda^2/2},$$

where

$$S_n = \sum_{k=1}^n \alpha_k X_k.$$

Proof. Write $\alpha_k = a_k + ib_k$. If

$$|S_n(\omega)| > \lambda \left(\sum_{k=1}^{n} |\alpha_k|^2 \right)^{1/2},$$

then

$$|S_n(\omega)|^2 > \lambda^2 \sum_{k=1}^{n} (a_k^2 + b_k^2).$$

But

$$|S_n(\omega)|^2 = \left(\sum_{k=1}^{n} a_k X_k(\omega) \right)^2 + \left(\sum_{k=1}^{n} b_k X_k(\omega) \right)^2,$$

so at least one of the following is true:

$$\left| \sum_{k=1}^{n} a_k X_k(\omega) \right| > \lambda \left(\sum_{k=1}^{n} a_k^2 \right)^{1/2}, \qquad \left| \sum_{k=1}^{n} b_k X_k(\omega) \right| > \lambda \left(\sum_{k=1}^{n} b_k^2 \right)^{1/2}.$$

By Lemma 4,

$$P\left(|S_n| > \lambda \left(\sum_{k=1}^{n} a_k^2 \right)^{1/2} \right) \le 2e^{-\lambda^2/2}$$

and

$$P\left(|S_n| > \lambda \left(\sum_{k=1}^{n} b_k^2 \right)^{1/2} \right) \le 2e^{-\lambda^2/2},$$

thus

$$\begin{aligned} P\left(|S_n| > \lambda \left(\sum_{k=1}^{n} |\alpha_k|^2 \right)^{1/2} \right) &\le P\left(|S_n| > \lambda \left(\sum_{k=1}^{n} a_k^2 \right)^{1/2} \right) \\ &\quad + P\left(|S_n| > \lambda \left(\sum_{k=1}^{n} b_k^2 \right)^{1/2} \right) \\ &\le 4e^{-\lambda^2/2}, \end{aligned}$$

proving the claim. □

We now prove **Khinchin's inequality**.[2]

[2]Camil Muscalu and Wilhelm Schlag, *Classical and Multilinear Harmonic Analysis*, volume I, p. 114, Lemma 5.5; Thomas H. Wolff, *Lectures on Harmonic Analysis*, p. 28, Proposition 4.5.

Theorem 3 (Khinchin's inequality)**.** *For $1 \le p < \infty$, let*

$$C(p) = \left(2^{1+\frac{p}{2}} \cdot p \cdot \Gamma\left(\frac{p}{2}\right)\right)^{1/p},$$

and let $\frac{1}{p} + \frac{1}{q} = 1$. If $X_1, \dots, X_n$ are independent random variables each with the Rademacher distribution and $a_1, \dots, a_n \in \mathbb{C}$, then

$$C(q)^{-1}\left(\sum_{k=1}^n |a_k|^2\right)^{1/2} \le E\left(\left|\sum_{k=1}^n a_k X_k\right|^p\right)^{1/p} \le C(p)\left(\sum_{k=1}^n |a_k|^2\right)^{1/2}.$$

Proof. First we remark that it can be computed that

$$\left(\int_0^\infty pt^{p-1} \cdot 4e^{-t^2/2}dt\right)^{1/p} = \left(2^{1+\frac{p}{2}} \cdot p \cdot \Gamma\left(\frac{p}{2}\right)\right)^{1/p} = C(p).$$

Let $\sigma^2 = \sum_{k=1}^n |a_k|^2$ and let $\alpha_k = \frac{a_k}{\sigma}$; if $\sigma = 0$ then the claim is immediate. To prove the claim it is equivalent to prove that

$$C(q)^{-1} \le E\left(\left|\sum_{k=1}^n \alpha_k X_k\right|^p\right)^{1/p} \le C(p).$$

Write $S_n = \sum_{k=1}^n \alpha_k X_k$. Using the fact that for a random variable X with $P(X \ge 0) = 1$,

$$E(X^p) = \int_0^\infty pt^{p-1} P(X \ge t)dt,$$

we obtain, applying Lemma 2,

$$E(|S_n|^p) = \int_0^\infty pt^{p-1}P(|S_n| \ge t)dt \le \int_0^\infty pt^{p-1} \cdot 4e^{-t^2/2}dt,$$

and thus

$$E(|S_n|^p)^{1/p} \le C(p). \tag{1}$$

Using Hölder's inequality, because the X_k are independent and $E(X_k) = 0$ and $E(|X_k|^2) = 1$,

$$\sum_{k=1}^n |\alpha_k|^2 = E\left(\left|\sum_{k=1}^n \alpha_k X_k\right|^2\right) \le E\left(\left|\sum_{k=1}^n \alpha_k X_k\right|^p\right)^{1/p} E\left(\left|\sum_{k=1}^n \alpha_k X_k\right|^q\right)^{1/q}.$$

Applying (1),

$$E\left(\left|\sum_{k=1}^n \alpha_k X_k\right|^q\right)^{1/q} \le C(q),$$

and as $\sum_{k=1}^n |\alpha_k|^2 = 1$ we obtain

$$1 \le C(q)E\left(\left|\sum_{k=1}^n \alpha_k X_k\right|^p\right)^{1/p}.$$

Thus we have

$$C(q)^{-1} \le E(|S_n|^p)^{1/p} \le C(p),$$

which proves the claim. ∎

2 Etemadi's inequality

The following is **Etemadi's inequality**.[3]

Theorem 4 (Etemadi's inequality). *If $X_1, \dots, X_n$ are independent random variables, then for any $x > 0$,*

$$P\left(\max_{1\le k\le n} |S_k| \ge 3x\right) \le 2P(|S_n| \ge x) + \max_{1\le k\le n} P(|S_k| \ge x) \le 3 \max_{1\le k\le n} P(|S_k| \ge x),$$

where $S_k = \sum_{j=1}^k X_j$.

Proof. For $k = 1, \dots, n$, let

$$A_k = \left\{\max_{1\le j\le k-1} |S_j| < 3x\right\} \cap \{|S_k| \ge 3x\},$$

with $A_1 = \{|S_1| \ge 3x\}$. $A_1, \dots, A_n$ are disjoint, and

$$A = \bigcup_{k=1}^n A_k = \left\{\max_{1\le k\le n} |S_k| \ge 3x\right\}.$$

For each $1 \le k \le n$,

$$A_k \cap \{|S_n| < x\} \subset A_k \cap \{|S_n - S_k| > 2x\},$$

and also, the events A_k and $\{|S_n - S_k| > 2x\}$ are independent, and thus

$$\begin{aligned} P(A) &= P(A \cap \{|S_n| \ge x\}) + P(A \cap \{|S_n| < x\}) \\ &\le P(|S_n| \ge x) + P(A \cap \{|S_n| < x\}) \\ &\le P(|S_n| \ge x) + \sum_{k=1}^n P(A_k \cap \{|S_n - S_k| > 2x\}) \\ &= P(|S_n| \ge x) + \sum_{k=1}^n P(A_k) P(|S_n - S_k| > 2x) \\ &\le P(|S_n| \ge x) + \max_{1\le k\le n} P(|S_n - S_k| > 2x) \cdot P(A). \end{aligned}$$

Then, because $|a - b| > 2x$ implies that $|a| > x$ or $|b| > x$,

$$\begin{aligned} P(A) &\le P(|S_n| \ge x) + \max_{1\le k\le n} P(|S_n - S_k| > 2x) \\ &\le P(|S_n| \ge x) + \max_{1\le k\le n} \left(P(|S_n| > x) + P(|S_k| > x)\right). \end{aligned}$$

∎

[3] Allan Gut, *Probability: A Graduate Course*, p. 143, Theorem 7.6.

The following inequality is similar enough to Etemadi's inequality to be placed in this note.[4]

Lemma 5. *Let $\xi_1, \dots, \xi_n$ be independent random variables with sample space $(\Omega, \mathscr{F}, P)$. Let $\zeta_0 = 0$ and for $1 \leq k \leq n$ let $\xi_k = \sum_{i=1}^k \xi_i$. If $P(|\zeta_n - \zeta_k| \leq t) \geq \alpha$ for $0 \leq k \leq n$ then*

$$P\left(\max_{1\leq k\leq n} |\zeta_k| > 2t\right) \leq \alpha^{-1} P(|\zeta_n| > t).$$

Proof. For $0 \leq k \leq n$ let

$$A_k = \{|\zeta_1| \leq 2t, \dots, |\zeta_{k-1}| \leq 2t, |\zeta_k| > 2t\}, \qquad B_k = \{|\zeta_n - \zeta_k| \leq t\},$$

where $A_0 = \Omega$. Because $|\zeta_n| \geq |\zeta_k| - |\zeta_n - \zeta_k|$,

$$A_k \cap B_k \subset \{|\zeta_n| > t\},$$

and so

$$\bigcup_{k=1}^n (A_k \cap B_k) \subset \{|\zeta_n| > t\}.$$

It is apparent that for $j \neq k$ the events A_j and A_k are disjoint, so the sets $A_1 \cap B_1, \dots, A_k \cap B_k$ are pairwise disjoint, hence

$$P(|\zeta_n| > t) \geq P\left(\bigcup_{k=1}^n (A_k \cap B_k)\right) = \sum_{k=1}^n P(A_k \cap B_k).$$

For each k, using that $\xi_1, \dots, \xi_n$ are independent one checks that the events A_k and B_k are independent, and using this,

$$P(|\zeta_n| > t) \geq \sum_{k=1}^n P(A_k)P(B_k) \geq \alpha \sum_{k=1}^n P(A_k) = \alpha P\left(\bigcup_{k=1}^n A_k\right),$$

that is,

$$P(|\zeta_n| > t) \geq \alpha P\left(\max_{1\leq k\leq n} |\zeta_k| > 2t\right),$$

proving the claim. □

[4] K. R. Parthasarathy, *Probability Measures on Metric Spaces*, p. 219, Chapter VII, Lemma 4.1.