The best constant in the Khintchine inequality for slightly dependent random variables

Orli Herscovici [1]
Joint work with Susanna Spektor [2]

[1] Department of Mathematics
University of Haifa

[2] Department of Mathematics and Statistics Sciences
Sheridan College Institute of Technology and Advanced Learning

Online Asymptotic Geometric Analysis Seminar
June 20, 2020

Khintchine inequality (1923)

$\forall p \in (0,\infty) \quad \exists\, A_p, B_p \quad$ s.t. for arbitrary $N \in \mathbb{N}$

$$A_p \left(\sum_{i=1}^{N} a_i^2 \right)^{\frac{1}{2}} \leq \mathbb{E} \left(\left| \sum_{i=1}^{N} a_i \varepsilon_i \right|^p \right)^{\frac{1}{p}} \leq B_p \left(\sum_{i=1}^{N} a_i^2 \right)^{\frac{1}{2}},$$

where for $i = 1, \dots, N$

- $a_i \in \mathbb{R}$,
- ε_i – mutually independent Rademacher r.v.s.,

$$\mathbb{P}(\varepsilon_i = 1) = \mathbb{P}(\varepsilon_i = -1) = \frac{1}{2}$$

Some related researches

1930 Littlewood; Paley and Zigmund – more systematic study of the inequality

1961 Stečkin: $B_{2n} = ((2n-1)!!)^{\frac{1}{2n}}$

1964 Kahane – generalization to normed spaces

1970s
- Young, Szarek, Haagerup – best constants
- Maurey and Pisier – study of geometric properties of Banach spaces
- Tomczak-Jaegermann – geometric properties, convexity

1980s Ball, Milman, Garling

1990s – Kahane, Latała, Oleszkiewicz, Tomczak-Jaegermann, Litvak, Milman, König, Peškir, Eskenazis, Nayar, Tkocz, Spektor
- convex bodies
- logconcave random variables
- Kahane-Khintchine inequality
- Steinhaus random variables
- ...

and many many others...

Khintchine inequality

D.J.H. Garling "Inequalities. A journey into linear analysis" (2007)

Theorem 12.3.1

There exist positive constants $\widetilde{A}_p$, $\widetilde{B}_p$, for $0 < p < \infty$, such that if $a_1, \dots, a_N$ are real numbers and $\varepsilon_1, \dots, \varepsilon_N$ are Rademacher random variables, then

$$\widetilde{B}_p \left(\mathbb{E} \left| \sum_{i=1}^{N} \varepsilon_i a_i \right|^p \right)^{\frac{1}{p}} \leq \left(\sum_{i=1}^{N} a_i^2 \right)^{\frac{1}{2}} \leq \widetilde{A}_p \left(\mathbb{E} \left| \sum_{i=1}^{N} \varepsilon_i a_i \right|^p \right)^{\frac{1}{p}}.$$

If $0 < p < 2$, we can take $\widetilde{B}_p = 1$ and $\widetilde{A}_p \leq 3^{\frac{1}{p} - \frac{1}{2}}$. If $2 \leq p \leq \infty$ we can take $\widetilde{B}_p \sim \sqrt{\frac{e}{p}}$ as $p \to \infty$, and $\widetilde{A}_p = 1$.

Khintchine inequality: proof

Consider the case $2 < p < \infty$.
If $2k-2 < p < 2k$, then

$$\left(\mathbb{E}\left|\sum_{i=1}^{N}\varepsilon_i a_i\right|^{2k-2}\right)^{\frac{1}{2k-2}} \leq \left(\mathbb{E}\left|\sum_{i=1}^{N}\varepsilon_i a_i\right|^{p}\right)^{\frac{1}{p}} \leq \left(\mathbb{E}\left|\sum_{i=1}^{N}\varepsilon_i a_i\right|^{2k}\right)^{\frac{1}{2k}}$$

Thus it is sufficient to establish the existence and asymptotic properties of $\widetilde{B}_{2k}$.

$$\left(\mathbb{E}\left|\sum_{i=1}^{N}\varepsilon_i a_i\right|^{2k}\right) = \mathbb{E}\left(\sum_{i=1}^{N}\varepsilon_i a_i\right)^{2k}$$

$$= \sum_{k_1+\cdots+k_N=2k} \frac{(2k)!}{k_1!\cdots k_N!} a_1^{k_1}\cdots a_N^{k_N}\,\mathbb{E}(\varepsilon_1^{k_1}\cdots\varepsilon_N^{k_N})$$

$$= \sum_{k_1+\cdots+k_N=2k} \frac{(2k)!}{k_1!\cdots k_N!} a_1^{k_1}\cdots a_N^{k_N}\,\mathbb{E}(\varepsilon_1^{k_1})\cdots\mathbb{E}(\varepsilon_N^{k_N})$$

Khintchine inequality: proof (cont.)

$$\mathbb{E}(\varepsilon_i^{k_i}) = \begin{cases} 1, & \text{if } k_i \text{ is even} \\ 0, & \text{if } k_i \text{ is odd} \end{cases}$$

Many of therms in the sum are 0, and

$$\mathbb{E}\Big|\sum_{i=1}^{N} \varepsilon_i a_i\Big|^{2k} = \sum_{k_1+\cdots+k_N=k} \frac{(2k)!}{(2k_1)!\cdots(2k_N)!} a_1^{2k_1}\cdots a_N^{2k_N}$$

But $(2k_1)!\cdots(2k_N)! \geq 2^{k_1}k_1!\cdots 2^{k_N}k_N! = 2^k k_1!\cdots k_N!$, and so

$$\mathbb{E}\Big|\sum_{i=1}^{N} \varepsilon_i a_i\Big|^{2k} \leq \frac{(2k)!}{2^k k!} \sum_{k_1+\cdots+k_N=k} \frac{k!}{k_1!\cdots k_N!} a_1^{2k_1}\cdots a_N^{2k_N}$$

$$= \frac{(2k)!}{2^k k!}||a||_2^{2k}$$

Slightly dependent Rademacher r.v.s.

Based on the O. Herscovici, S. Spektor, "The best constant in the Khintchine inequality for slightly dependent random variables", arXiv:1806.03562v3.

Our assumption:

$$\sum_{i=1}^{N} \varepsilon_i = M$$

$M = 0$ $\rightarrow$ we'll consider,
$M > 0$ generalizes the previous case,
$M < 0$ similar to the case $M > 0$.

Note

$$\mathbb{E}_M \left| \sum_{i=1}^{N} \varepsilon_i a_i \right|^{2p} = \sum_{\substack{p_1+\ldots+p_N=2p \\ p_i \in \{0,\ldots,2p\}}} \frac{(2p)!}{p_1! \ldots p_N!} a_1^{p_1} \ldots a_N^{p_N} \, \mathbb{E}_M \left(\prod_{i=1}^{N} \varepsilon_i^{p_i} \right)$$

$$= \mathbb{E}_M \left(\prod_{i=1}^{N} \varepsilon_i^{p_i} \right) \cdot (a_1 + \ldots + a_N)^{2p}.$$

Slightly dependent Rademacher r.v.s.: case $M = 0$

Theorem (HS, 2020)

Let ε_i, $1 \leq i \leq N$, *be Rademacher random variables satisfying condition* $\sum_{i=1}^N \varepsilon_i = 0$. *Let* $a = (a_1, \dots, a_N) \in \mathbb{R}^N$. *Then for any* $p \in \mathbb{N}$,

$$\mathbb{E}_M \left| \sum_{i=1}^N \varepsilon_i a_i \right|^{2p} \leq C_{2p}^{2p} ||a||_2^{2p},$$

where

$$C_{2p}^{2p} = \left(\frac{N}{2}\right)^{p+1} \cdot \frac{\sqrt{\pi}\,\Gamma\left(\frac{N}{2}\right)}{\Gamma(p + \frac{N}{2} + \frac{1}{2})} \cdot \frac{(2p)!}{2^p p!}.$$

Case $M = 0$: proof

Note that $\prod_{i=1}^{N} \varepsilon_i^{p_i} = \pm 1$. In our case

$$M = 0 \quad \Longrightarrow \quad \big|\{i \,|\, \varepsilon_i = 1\}\big| = \big|\{i \,|\, \varepsilon_i = -1\}\big| = \ell$$

Let

$i_1, \ldots, i_\ell$ be the indexes of $\varepsilon_{i_j} = -1$,

$i_{\ell+1}, \ldots, i_{2\ell}$ be the indexes of $\varepsilon_{i_j} = 1$,

then

$$\prod_{i=1}^{N} \varepsilon_i^{p_i} = \prod_{j=1}^{\ell} \varepsilon_{i_j}^{p_{i_j}} \cdot \prod_{j=\ell+1}^{2\ell} \varepsilon_{i_j}^{p_{i_j}} = \prod_{j=1}^{\ell} (-1)^{p_{i_j}} \cdot \prod_{j=\ell+1}^{2\ell} 1^{p_{i_j}} = (-1)^{p_{i_1} + \ldots + p_{i_\ell}}$$

From here on we re enumerate the indexes $i_j \to j$, s.t. we have

$$\prod_{i=1}^{N} \varepsilon_i^{p_i} = \begin{cases} 1, & \text{if } p_1 + \ldots + p_\ell \text{ is even,} \\ -1, & \text{if } p_1 + \ldots + p_\ell \text{ is odd.} \end{cases}$$

Case $M = 0$: proof

Lemma (HS, 2020)

Let ε_i, $i \le N$, be Rademacher random variables satisfying condition $\sum_{i=1}^{N} \varepsilon_i = 0$ and let $p_1 + \ldots + p_N = 2p$, $p_i \in \{0, \ldots, 2p\}$. Then,

$$\mathbb{P}_{Dif} := \mathbb{P}_+ - \mathbb{P}_- = \frac{\binom{p+\frac{N}{2}-1}{p}}{\binom{2p+N-1}{2p}},$$

where

$$\mathbb{P}_+ = \mathbb{P}\left(\left\{\prod_{i=1}^{N} \varepsilon_i^{p_i} = 1\right\} \bigcap \left\{\sum_{i=1}^{N} \varepsilon_i = 0\right\}\right),$$

$$\mathbb{P}_- = \mathbb{P}\left(\left\{\prod_{i=1}^{N} \varepsilon_i^{p_i} = -1\right\} \bigcap \left\{\sum_{i=1}^{N} \varepsilon_i = 0\right\}\right).$$

We have that

$$\prod_{i=1}^{N} \varepsilon_i^{p_i} = \begin{cases} 1, & \text{if } p_1 + \ldots + p_\ell \text{ is even,} \\ -1, & \text{if } p_1 + \ldots + p_\ell \text{ is odd.} \end{cases}$$

Notation

T_{even}	the number of all solutions of $p_1 + \ldots + p_N = 2p$, where $p_1 + \ldots + p_\ell$ is even
T_{odd}	the number of all solutions of $p_1 + \ldots + p_N = 2p$, where $p_1 + \ldots + p_\ell$ is odd
T	the number of all solutions of $p_1 + \ldots + p_N = 2p$

$$\Longrightarrow \mathbb{P}_{Dif} = \frac{T_{even} - T_{odd}}{T}$$

Proof of the Lemma about $\mathbb{P}_{Dif}$

$T = \binom{2p+N-1}{2p}$ is the number of weak compositions of $2p$ into N parts.

To find $T_{even} - T_{odd}$, we divide the sequences summing to $2p$ into classes and sum over each class separately.

$$2p = p_1 + p_2 + \ldots + p_\ell + p_{\ell+1} + p_{\ell+2} + \ldots + p_{2\ell}$$

Now we consider the difference $T_{even} - T_{odd}$ for each class $c = (c_1, c_2, \ldots, c_\ell) := (p_1 + p_{\ell+1}, p_2 + p_{\ell+2}, \ldots, p_\ell + p_{2\ell})$.

$$\boxed{\begin{array}{l} \text{if } p_1 + \ldots + p_\ell = even \implies (-1)^{p_1+\ldots+p_\ell} = 1 \\ \text{if } p_1 + \ldots + p_\ell = odd \implies (-1)^{p_1+\ldots+p_\ell} = -1 \end{array}}$$

⇓

$$(T_{even} - T_{odd})_c = \sum_c (-1)^{p_1+\ldots+p_\ell}$$

Proof of the Lemma about $\mathbb{P}_{Dif}$

$$(T_{even} - T_{odd})_c = \sum_c (-1)^{p_1+\ldots+p_\ell}$$
$$= ((-1)^{p_{1,1}} + (-1)^{p_{1,2}} + \ldots)\cdots((-1)^{p_{\ell,1}} + (-1)^{p_{\ell,2}} + \ldots)$$

$$(T_{even} - T_{odd})_c = \prod_{j=1}^{\ell} \sum_{p_j+p_{j+\ell}=c_j} (-1)^{p_j}$$

For fixed class c and any c_j in this class, we have

p_j	0	1	2	$\cdots$	c_j
$p_{j+\ell}$	c_j	c_{j-1}	c_{j-2}	$\cdots$	0

$$\Longrightarrow \sum_{p_j+p_{j+\ell}=c_j} (-1)^{p_j} = \begin{cases} 0, & \text{if } c_j \text{ is odd,} \\ 1, & \text{if } c_j \text{ is even.} \end{cases}$$

Proof of the Lemma about $\mathbb{P}_{Dif}$

Thus for any class $c = (c_1, \ldots, c_\ell)$ we have

$$\implies (T_{even} - T_{odd})_c = \begin{cases} 0, & \text{if at least one of } c_j \text{ is odd,} \\ 1, & \text{if } \textit{all } c_j \text{ is even,} \end{cases}$$

which means any $c_j = 2z_j$ and

$$2p = c_1 + \ldots + c_\ell \quad \implies \quad p = z_1 + \ldots + z_\ell$$

$$\Downarrow$$

$$T_{even} - T_{odd} = \#\text{ weak compositions of } p \text{ into } \ell \text{ parts,}$$
$$= \binom{p+\ell-1}{p},$$

and

$$\Downarrow$$

$$\mathbb{P}_{Dif} = \frac{T_{even} - T_{odd}}{T} = \frac{\binom{p+\ell-1}{p}}{\binom{2p+N-1}{2p}}$$

Case $M = 0$: proof (cont.)

Lemma (HS, 2020)

Let ε_i, $i \leq N$, be Rademacher random variables satisfying condition

$$\sum_{i=1}^{N} \varepsilon_i = 0, \tag{1}$$

and let $p_1 + \ldots + p_N = 2p$, $p_i \in \{0, \ldots, 2p\}$. Denote by $\mathbb{E}_M$ an expectation with condition (1). Then,

$$\mathbb{E}_M \left(\prod_{i=1}^{N} \varepsilon_i^{p_i} \right) = \frac{2^N}{\binom{N}{\frac{N}{2}}} \frac{\binom{p+\frac{N}{2}-1}{p}}{\binom{2p+N-1}{2p}}$$

Proof of the Lemma about $\mathbb{E}_M$

$$\mathbb{E}_M\left(\prod_{i=1}^{N}\varepsilon_i^{p_i}\right) = 1\times\mathbb{P}_M\left(\prod_{i=1}^{N}\varepsilon_i^{p_i}=1\right) - 1\times\mathbb{P}_M\left(\prod_{i=1}^{N}\varepsilon_i^{p_i}=-1\right)$$
$$= \frac{\mathbb{P}_{Dif}}{\mathbb{P}\left(\sum_{i=1}^{N}\varepsilon_i=0\right)}.$$

$\mathbb{P}_{Dif}$ – calculated

Let us find $\mathbb{P}\left(\sum_{i=1}^{N}\varepsilon_i=0\right)$. Denote $D=\{i:\varepsilon_i=1\}$ and $D^c=\{i:\varepsilon_i=-1\}$. Note, the cardinalities $\mathrm{card}(D)=\mathrm{card}(D^c)=\frac{N}{2}$.

The event

$$\left\{\sum_{i=1}^{N}\varepsilon_i=0\right\} = \{\varepsilon_i=1\,|\,\forall i\in D\}\biguplus\{\varepsilon_i=-1\,|\,\forall i\in D^c\}.$$

$$\Longrightarrow \quad \mathbb{P}\left(\sum_{i=1}^{N}\varepsilon_i=0\right)=\frac{1}{2^N}\binom{N}{\frac{N}{2}}.$$

Case $M = 0$: proof (cont.)

$$\mathbb{E}_M \left| \sum_{i=1}^{N} \varepsilon_i a_i \right|^{2p} = \mathbb{E}_M \left(\prod_{i=1}^{N} \varepsilon_i^{p_i} \right) \cdot (a_1 + \ldots + a_N)^{2p}$$

For any $a = (a_1, \ldots, a_N)$

- $a_1 + \ldots + a_N \leq |a_1| + \ldots + |a_N| \equiv ||a||_1$
- $||a||_2 \leq ||a||_1 \leq \sqrt{N}||a||_2$

$$(a_1 + \ldots + a_N)^{2p} \leq N^p ||a||_2^{2p}$$

$$\mathbb{E}_M \left| \sum_{i=1}^{N} \varepsilon_i a_i \right|^{2p} \leq \frac{2^N \binom{p + \frac{N}{2} - 1}{p}}{\binom{N}{\frac{N}{2}} \binom{2p+N-1}{2p}} N^p ||a||_2^{2p}$$

Case $M = 0$: proof (cont.)

The constant

$$C_{2p}^{2p} = \frac{2^N \binom{p+\frac{N}{2}-1}{p}}{\binom{N}{\frac{N}{2}} \binom{2p+N-1}{2p}} \cdot N^p$$

$$= \frac{2^{N-1} \left(\frac{N}{2}\right)! \left(p + \frac{N}{2} - 1\right)!}{(2p+N-1)!} \cdot N^p \cdot \frac{(2p)!}{p!}.$$

Since $x! = \Gamma(x+1) = x\Gamma(x)$, we have

- $(p + \frac{N}{2} - 1)! = \Gamma(p + \frac{N}{2})$,
- $(2p + N - 1)! = \Gamma(2p + N)$.

Applying duplication formula $\Gamma(2x) = \pi^{-\frac{1}{2}} 2^{2x-1} \Gamma(x)\Gamma(x + \frac{1}{2})$ to $\Gamma(2p + N)$, we obtain

$$C_{2p}^{2p} = \left(\frac{N}{2}\right)^{p+1} \cdot \frac{\sqrt{\pi}\,\Gamma\left(\frac{N}{2}\right)}{\Gamma(p + \frac{N}{2} + \frac{1}{2})} \cdot \frac{(2p)!}{2^p p!}$$

Case $M = 0$: asymptotic approximation

Proposition (HS, 2020)

The constant C_{2p}^{2p} has the following upper bound.

$$C_{2p}^{2p} \leq \frac{2^N N^p}{(N+1)^p} \cdot \frac{\left(\frac{N}{2}!\right)^2}{N!} \cdot \frac{(2p)!}{2^p p!} \sim e^{-\frac{p}{N}} \sqrt{\frac{\pi N}{2}} \cdot \frac{(2p)!}{2^p p!},$$

as $N \to \infty$.

PROOF.

$$C_{2p}^{2p} = \left(\frac{N}{2}\right)^{p+1} \cdot \frac{\sqrt{\pi}\,\Gamma\left(\frac{N}{2}\right)}{\Gamma(p + \frac{N}{2} + \frac{1}{2})} \cdot \frac{(2p)!}{2^p p!}$$

From $\Gamma(x+1) = x\Gamma(x)$ we obtain

$$\Gamma\left(\frac{N}{2} + p + \frac{1}{2}\right) = \prod_{i=0}^{p-1} \left(\frac{N+1}{2} + i\right) \Gamma\left(\frac{N}{2} + \frac{1}{2}\right)$$

Case $M = 0$: asymptotic approximation

It is easy to see that

$$\prod_{i=0}^{p-1}\left(\frac{N+1}{2}+i\right) \geq \left(\frac{N+1}{2}\right)^{p}.$$

Duplication formula with integer n gives $\Gamma\left(n+\frac{1}{2}\right) = \frac{(2n)!\sqrt{\pi}}{2^{2n}n!}$.

Therefore

$$C_{2p}^{2p} \leq \left(\frac{N}{2}\right)^{p+1} \cdot \frac{2^{N}\left(\frac{N}{2}\right)!\Gamma\left(\frac{N}{2}\right)}{\left(\frac{N+1}{2}\right)^{p} N!} \cdot \frac{(2p)!}{2^{p}p!}$$

$$= \frac{N^{p}}{(N+1)^{p}} \cdot \frac{2^{N}\left(\frac{N}{2}!\right)^{2}}{N!} \cdot \frac{(2p)!}{2^{p}p!}.$$

Let us consider now an asymptotic behaviour of the constant C_{2p}^{2p} as $N \to \infty$.

Case $M = 0$: asymptotic approximation

We have to consider

- $\frac{N^p}{(N+1)^p}$
- $\frac{\left(\frac{N}{2}!\right)^2}{N!}$

For the first term we have

$$\frac{N^p}{(N+1)^p} = \left(1 + \frac{1}{N}\right)^{-p} = \left(1 + \frac{1}{N}\right)^{N \cdot \left(-\frac{p}{N}\right)} \sim e^{-\frac{p}{N}},$$

while the second term can be approximated by applying $\binom{2n}{n} \sim \frac{2^{2n}}{\sqrt{\pi n}}$ (see Elezović, "Asymptotic expansions of central binomial coefficients and Catalan numbers" (2014)).
Finally,

$$C_{2p}^{2p} \sim e^{-\frac{p}{N}} \sqrt{\frac{\pi N}{2}} \cdot \frac{(2p)!}{2^p p!}$$

Case $M > 0$

Theorem (HS, 2020)

Let ε_i, $i \leq N$, be Rademacher random variables satisfying condition $\sum_{i=1}^{n} \varepsilon_i = M \geq 0$. Let $a = (a_1, \ldots, a_N) \in \mathbb{R}^N$. Then for any $p \in \mathbb{N}$,

$$\mathbb{E}_M \left(\left| \sum_{i=1}^{N} \varepsilon_i a_i \right|^{2p} \right) \leq C_{2p}^{2p} ||a||_2^{2p},$$

where

$$C_{2p}^{2p} = \frac{2^N N^p}{\binom{N}{\frac{N+M}{2}} \binom{2p+N-1}{2p}} \sum_{m=0}^{p} \binom{p - m + \frac{N-M}{2} - 1}{p-m} \binom{2m + M - 1}{2m}.$$

Case $M > 0$: sketch of the proof

- $\sum_{i=1}^{N} \varepsilon_i = M \implies |\{i \,|\, \varepsilon_i = -1\}| = \ell$, and $|\{i \,|\, \varepsilon_i = 1\}| = M + \ell$, where $\ell = \frac{N-M}{2}$
- a renumeration of variables:

$$\prod_{i=1}^{N} \varepsilon_i^{p_i} = \prod_{i=1}^{\ell} (-1)^{p_i} \prod_{j=\ell+1}^{2\ell+M} 1^{p_j} = \begin{cases} 1, & p_1 + \ldots + p_\ell \text{ is even}, \\ -1, & p_1 + \ldots + p_\ell \text{ is odd}. \end{cases}$$

- a construction of classes c:

$$2p = p_1 + \ldots + p_\ell + \ldots + p_{2\ell} + \ldots + p_{2\ell+M}$$

$$\Downarrow$$

$$c = (c_1, \ldots, c_\ell, p_{2\ell+1}, \ldots, p_{2\ell+M}),$$

where $c_j = p_j + p_{\ell+j}$, and $c_j, p_i \in \{0, \ldots, 2p\}$.

Case $M > 0$: sketch of the proof

- for each such class c consider the difference $(T_{even} - T_{odd})_c$

$$2p = 2z_1 + \ldots + 2z_\ell + p_{2\ell+1} + \ldots + p_{2\ell+M}$$

$$\Downarrow$$

$$\begin{cases} p_{2\ell+1} + \ldots + p_{2\ell+M} = 2m, \\ 2z_1 + \ldots + 2z_\ell = 2p - 2m \end{cases}$$

for some m.

$$\Downarrow$$

$$(T_{even} - T_{odd})_c = \binom{p-m+\ell-1}{p-m}\binom{2m+M-1}{2m}.$$

Asymptotics of growing sample with given M

Theorem (HS, 2020)

Let $\varepsilon_1, \dots, \varepsilon_N$ *be Rademacher random variables, i.e. such that* $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = 1/2$*, with condition that* $\sum_{i=1}^{N} \varepsilon_i = M$*, where* $0 < M < N$ *is fixed. Then, for any integer* $p \geq 2$

$$C_{2p}^{2p} \sim \sqrt{\frac{\pi(N^2 - M^2)}{2N}} \frac{e^{-\frac{Mp}{N}}}{(M-1)!} \frac{(2p)!}{2^p} \cdot \sum_{m=0}^{p} \frac{(2m+M-1)!}{(p-m)!(2m)!} \left(\frac{2}{N-M}\right)^m,$$

when $N \to \infty$.

Asymptotics of proportionally growing samples

Theorem (HS, 2020)

Let $\varepsilon_1, \dots, \varepsilon_N$ be Rademacher random variables, i.e. such that $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = 1/2$, with condition that $\sum_{i=1}^N \varepsilon_i = M$. Let a number of negative ε_i is n and a number of positive ε_i is αn, for some fixed real $\alpha > 1$. Then, for any integer $p \geq 2$

$$C_{2p}^{2p} \sim \sqrt{\frac{2\pi n\alpha}{\alpha+1}} \frac{\alpha^{\alpha n} 2^{(\alpha+1)n}}{(\alpha+1)^{(\alpha+1)n}} \frac{(2p)!}{(\alpha+1)^p} \cdot \sum_{m=0}^{p} \frac{(\alpha-1)^{2m} n^m}{(p-m)!(2m)!},$$

when $n \to \infty$.

If $\alpha \to 1$, then

$$C_{2p}^{2p} \sim \sqrt{\pi n} \frac{(2p)!}{2^p p!}$$

$$C_{2p}^{2p} \sim e^{-\frac{p}{N}} \sqrt{\frac{\pi N}{2}} \cdot \frac{(2p)!}{2^p p!}$$

Asymptotics of proportionally growing samples

Proposition (HS, 2020)

Let $\varepsilon_1, \dots, \varepsilon_N$ be Rademacher random variables, i.e. such that $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = 1/2$, with condition that $\sum_{i=1}^{N} \varepsilon_i = M$. Let a number of negative ε_i is n and a number of positive ε_i is αn, for some fixed real $\alpha > 1$. Then for any $p \geq 2$, the coefficient C_{2p}^{2p} has the following upper bound.

$$C_{2p}^{2p} \leq \sqrt{\frac{2\pi n\alpha}{\alpha+1}}\,\frac{\alpha^{\alpha n}\,2^{(\alpha+1)n}}{(\alpha+1)^{(\alpha+1)n}}\,\frac{(\alpha-1)^{2p}n^{p}}{(p+1)^{p}}\,\frac{(2p)!}{(\alpha+1)^{p}p!},$$

where $n \to \infty$.

Asymptotics of proportionally growing samples

Proposition (HS, 2020)

Let $\varepsilon_1, \ldots, \varepsilon_N$ be Rademacher random variables, i.e. such that $P(\varepsilon_i = 1) = P(\varepsilon_i = -1) = 1/2$, with condition that $\sum_{i=1}^{N} \varepsilon_i = M$, with $M = \beta N, \quad 0 < \beta < 1$. Then, for any integer $p \geq 2$,

$$C_{2p}^{2p} \sim \sqrt{\frac{\pi N}{2}} (1-\beta^2)^{\frac{N+1}{2}} \left(\frac{1+\beta}{1-\beta}\right)^{\frac{\beta N}{2}} \frac{(1-\beta)^p (2p)!}{2^p} \sum_{m=0}^{p} \frac{2^m \beta^{2m} N^m}{(1-\beta)^m (p-m)!(2m)!}$$

when $N \to \infty$.

If $\beta \to 0$, then

$$C_{2p}^{2p} \sim \sqrt{\frac{\pi N}{2}} \frac{(2p)!}{2^p p!}$$

$$C_{2p}^{2p} \sim e^{-\frac{p}{N}} \sqrt{\frac{\pi N}{2}} \cdot \frac{(2p)!}{2^p p!}$$

Applications

Moments $\implies$ information about the tail of random variable

Possible applications:

- Risk and portfolio management
- Epidemiological studies

The next example is based on the article by A.B. Kashlak, S. Myroshnychenko, S. Spektor, *“Analytic permutation testing via Kahane-Khintchine inequalities”*, arXiv:2001.01130

Automatic Speech Recognition

One of the methods is a permutation test. It maps a sample of size n onto the symmetric group with $n!$ elements.
Disadvantages of the method:

- long computation time
- a conservative test – more likely to get a false negative and miss a significant result

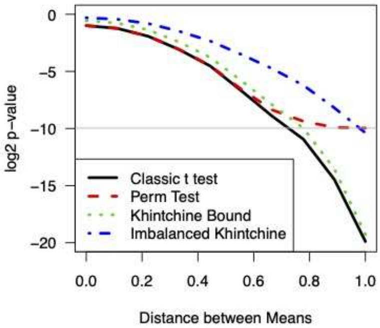
Applications

Proposed solution:

Applying Kahane-Khintchine's type inequality, which effectively consider the entire discrete distribution at once ***without the need to simulate***

Main idea:

the moment bounds from the Kahane-Khintchine's type inequality directly imply bounds on the tail probability of the test statistic, which means bounding the p-value for testing for significance



Thank you for your attention