Quantitative probability from a logician's perspective: the law of large numbers

Mathematical Logic: Proof Theory, Constructive Mathematics
Oberwolfach
14 November 2023

Morenikeji Neri
Department of Computer Science
University of Bath, Bath, United Kingdom
kejineri.github.io

Overview of talk

- Preliminaries from logic and probability theory.
- What has been done in the quantitative study of the strong laws of large numbers?
- What am I doing?
- Looking towards the future.

Overview of talk

- Preliminaries from logic and probability theory.
- What has been done in the quantitative study of the strong laws of large numbers?
- What am I doing?
- Looking towards the future.

Overview of talk

- Preliminaries from logic and probability theory.
- What has been done in the quantitative study of the strong laws of large numbers?
- What am I doing?
- Looking towards the future.

Overview of talk

- Preliminaries from logic and probability theory.
- What has been done in the quantitative study of the strong laws of large numbers?
- What am I doing?
- Looking towards the future.

Preliminaries from logic and probability theory

Preliminaries from logic

Let us recall the standard definition of Cauchy convergence on an arbitrary metric space (X, d)

$$P := \forall \varepsilon \in \mathbb{Q}_+ \, \exists N \in \mathbb{N} \, \forall n \in \mathbb{N} \, (n \geq N \implies d(a_n, a_N) \leq \varepsilon)$$

A natural question to ask is, from a proof of P can one find a computable function $f : \mathbb{Q} \to \mathbb{N}$ such that

$$\forall \varepsilon \in \mathbb{Q}_+ \, \forall n \in \mathbb{N} \, (n \geq f(\varepsilon) \implies d(a_n, a_{f(\varepsilon)}) \leq \varepsilon)$$

Specker showed this was not always possible, through his famous construction of a bounded monotone sequence of rational numbers converging to a non-computable number [Spe49].

Preliminaries from logic

Let us recall the standard definition of Cauchy convergence on an arbitrary metric space (X, d)

$$P := \forall \varepsilon \in \mathbb{Q}_+ \, \exists N \in \mathbb{N} \, \forall n \in \mathbb{N} \, (n \geq N \implies d(a_n, a_N) \leq \varepsilon)$$

A natural question to ask is, from a proof of P can one find a computable function $f : \mathbb{Q} \to \mathbb{N}$ such that

$$\forall \varepsilon \in \mathbb{Q}_+ \, \forall n \in \mathbb{N} \, (n \geq f(\varepsilon) \implies d(a_n, a_{f(\varepsilon)}) \leq \varepsilon)$$

Specker showed this was not always possible, through his famous construction of a bounded monotone sequence of rational numbers converging to a non-computable number [Spe49].

Preliminaries from logic

Let us recall the standard definition of Cauchy convergence on an arbitrary metric space (X, d)

$$P := \forall \varepsilon \in \mathbb{Q}_+ \, \exists N \in \mathbb{N} \, \forall n \in \mathbb{N} \, (n \geq N \implies d(a_n, a_N) \leq \varepsilon)$$

A natural question to ask is, from a proof of P can one find a computable function $f : \mathbb{Q} \to \mathbb{N}$ such that

$$\forall \varepsilon \in \mathbb{Q}_+ \, \forall n \in \mathbb{N} \, (n \geq f(\varepsilon) \implies d(a_n, a_{f(\varepsilon)}) \leq \varepsilon)$$

Specker showed this was not always possible, through his famous construction of a bounded monotone sequence of rational numbers converging to a non-computable number [Spe49].

Preliminaries from logic

One can show that Cauchy convergence is equivalent, over classical logic, to

$$\forall \varepsilon \in \mathbb{Q}_+ \, \forall g : \mathbb{N} \to \mathbb{N} \, \exists n \, \forall i, j \in [n, n + g(n)] (d(a_i, a_j) \leq \varepsilon)$$

This can be seen as an instance of Kreisel's no counterexample interpretation [Kre52; Kre51]and results in a new computational challenge, that is, to find a functional (known as a rate of metastability)

$$\Phi : \mathbb{Q}_+ \times (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$$

such that

$$\forall \varepsilon \in \mathbb{Q}_+ \, \forall g : \mathbb{N} \to \mathbb{N} \, \exists n \leq \Phi(\varepsilon, g) \forall i, j \in [n, n + g(n)] (d(a_i, a_j) \leq \varepsilon)$$

Preliminaries from logic

One can show that Cauchy convergence is equivalent, over classical logic, to

$$\forall \varepsilon \in \mathbb{Q}_+ \, \forall g : \mathbb{N} \to \mathbb{N} \, \exists n \, \forall i, j \in [n, n + g(n)](d(a_i, a_j) \leq \varepsilon)$$

This can be seen as an instance of Kreisel's no counterexample interpretation [Kre52; Kre51]and results in a new computational challenge, that is, to find a functional (known as a rate of metastability)

$$\Phi : \mathbb{Q}_+ \times (\mathbb{N} \to \mathbb{N}) \to \mathbb{N}$$

such that

$$\forall \varepsilon \in \mathbb{Q}_+ \, \forall g : \mathbb{N} \to \mathbb{N} \, \exists n \leq \Phi(\varepsilon, g) \forall i, j \in [n, n + g(n)](d(a_i, a_j) \leq \varepsilon)$$

Preliminaries from probability theory

Our main focus will be on quantitative results related to the probabilistic convergence of sequences of random variables.

There are many notions for the convergence of sequences of random variables. In this talk, we shall be concerned mostly be concerned with a notion known as almost sure convergence.

Jointly with Thomas Powell, I am also investigating the computational content of theorems concerned with other modes of probabilistic convergence.

Preliminaries from probability theory

Our main focus will be on quantitative results related to the probabilistic convergence of sequences of random variables.

There are many notions for the convergence of sequences of random variables. In this talk, we shall be concerned mostly be concerned with a notion known as almost sure convergence.

Jointly with Thomas Powell, I am also investigating the computational content of theorems concerned with other modes of probabilistic convergence.

Preliminaries from probability theory

Our main focus will be on quantitative results related to the probabilistic convergence of sequences of random variables.

There are many notions for the convergence of sequences of random variables. In this talk, we shall be concerned mostly be concerned with a notion known as almost sure convergence.

Jointly with Thomas Powell, I am also investigating the computational content of theorems concerned with other modes of probabilistic convergence.

Preliminaries from probability theory: Almost sure convergence

If $\{Y_n\}$ is a sequence of random variables and Y is a random variable, we say Y_n converges to Y almost surely (a.s) if

$$\mathbb{P}(\{\omega \in \Omega : Y_n(\omega) \to Y(\omega)\}) = 1$$

It is not obvious how one would go about giving this definition a meaningful computational interpretation.

However, by Egorov's theorem, one can show that it is equivalent to almost uniform (a.u) convergence

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}(\{\omega \in \Omega : \forall m \geq N \, |Y_m(\omega) - Y(\omega)| \leq \lambda_1\}) > 1 - \lambda_2$$

The definition of a.u convergence can be given a direct computational interpretation and is the usual interpretation of a.s convergence used by probability theorists.

Preliminaries from probability theory: Almost sure convergence

If $\{Y_n\}$ is a sequence of random variables and Y is a random variable, we say Y_n converges to Y almost surely (a.s) if

$$\mathbb{P}(\{\omega \in \Omega : Y_n(\omega) \to Y(\omega)\}) = 1$$

It is not obvious how one would go about giving this definition a meaningful computational interpretation.

However, by Egorov's theorem, one can show that it is equivalent to almost uniform (a.u) convergence

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}(\{\omega \in \Omega : \forall m \geq N \, |Y_m(\omega) - Y(\omega)| \leq \lambda_1\}) > 1 - \lambda_2$$

The definition of a.u convergence can be given a direct computational interpretation and is the usual interpretation of a.s convergence used by probability theorists.

Preliminaries from probability theory: Almost sure convergence

Furthermore, we can also obtain a metastable notion of almost uniform convergence

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \forall K : \mathbb{N} \to \mathbb{N} \, \exists N$$
$$\mathbb{P}(\{\omega \in \Omega : \forall m \in [N, K(N)] |Y_m(\omega) - Y(\omega)| \leq \lambda_1\}) > 1 - \lambda_2$$

This notion of probabilistic convergence (in its Cauchy form) was first studied by Avigad et al. [ADR12; AGT10], with explicit rates being extracted for the pointwise ergodic theorem and a computational version of Egorov's theorem.

What has been done in the quantitative study of the strong laws of large numbers?

The strong law of large numbers

Suppose $X_1, X_2, \ldots$ are pairwise independent identically distributed (iid) real-valued random variables with $\mathbb{E}(X_n) = 0$ and $\mathbb{E}(|X_n|) < \infty$ for all n. Define $S_n = \sum_{i=1}^{n} X_i$. The strong law of large numbers states that,

$$\frac{S_n}{n} \to 0$$

almost surely.

What is the quantitative version of this theorem?

The phrase "Rates for the strong law of large numbers" has appeared in many probability papers, where the notion of a rate means different things to different authors.

There are many notions of a rate used by these authors that provide further information about the speed of the convergence of the sequences of random variables in question. We shall focus on two.

What is the quantitative version of this theorem?

The phrase "Rates for the strong law of large numbers" has appeared in many probability papers, where the notion of a rate means different things to different authors.

There are many notions of a rate used by these authors that provide further information about the speed of the convergence of the sequences of random variables in question. We shall focus on two.

Quantitative almost sure convergence to 0

Let us recall the definition of

$$\frac{S_n}{n} \to 0 \text{ a.u}$$

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}\left(\forall n \geq N \, |\frac{S_n}{n}| \leq \lambda_1\right) > 1 - \lambda_2$$

This is equivalent to

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}\left(\sup_{n \geq N} |\frac{S_n}{n}| > \lambda_1\right) \leq \lambda_2$$

And this is equivalent to

$$\forall \varepsilon \in \mathbb{Q}^+ \, P_{\varepsilon,n} \to 0$$

where

$$P_{n,\varepsilon} = \mathbb{P}\left(\sup_{m \geq n} |\frac{S_n}{n}| > \varepsilon\right)$$

Quantitative almost sure convergence to 0

Let us recall the definition of

$$\frac{S_n}{n} \to 0 \text{ a.u}$$

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}\left(\forall n \geq N \, |\frac{S_n}{n}| \leq \lambda_1\right) > 1 - \lambda_2$$

This is equivalent to

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}\left(\sup_{n \geq N} |\frac{S_n}{n}| > \lambda_1\right) \leq \lambda_2$$

And this is equivalent to

$$\forall \varepsilon \in \mathbb{Q}^+ \, P_{\varepsilon,n} \to 0$$

where

$$P_{n,\varepsilon} = \mathbb{P}\left(\sup_{m \geq n} |\frac{S_n}{n}| > \varepsilon\right)$$

Quantitative almost sure convergence to 0

Let us recall the definition of

$$\frac{S_n}{n} \to 0 \text{ a.u}$$

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}\left(\forall n \geq N \, |\frac{S_n}{n}| \leq \lambda_1\right) > 1 - \lambda_2$$

This is equivalent to

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}\left(\sup_{n \geq N} |\frac{S_n}{n}| > \lambda_1\right) \leq \lambda_2$$

And this is equivalent to

$$\forall \varepsilon \in \mathbb{Q}^+ \, P_{\varepsilon,n} \to 0$$

where

$$P_{n,\varepsilon} = \mathbb{P}\left(\sup_{m \geq n} |\frac{S_n}{n}| > \varepsilon\right)$$

Quantitative almost sure convergence to 0

Let us recall the definition of

$$\frac{S_n}{n} \to 0 \text{ a.u}$$

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}\left(\forall n \geq N \, |\frac{S_n}{n}| \leq \lambda_1\right) > 1 - \lambda_2$$

This is equivalent to

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \exists N \in \mathbb{N} \, \mathbb{P}\left(\sup_{n \geq N} |\frac{S_n}{n}| > \lambda_1\right) \leq \lambda_2$$

And this is equivalent to

$$\forall \varepsilon \in \mathbb{Q}^+ \, P_{\varepsilon, n} \to 0$$

where

$$P_{n,\varepsilon} = \mathbb{P}\left(\sup_{m \geq n} |\frac{S_n}{n}| > \varepsilon\right)$$

A direct approach

One can thus study the speed of convergence of these random variables by studying the speed of convergence of the sequence of real numbers $P_{\varepsilon,n}$, for each fixed ε.

This can be approached directly. That is, once can try to find a function $f : \mathbb{Q}^+ \times \mathbb{Q}^+ \to \mathbb{N}$ such that,

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \forall n \geq f(\lambda_1, \lambda_2) \, P_{\lambda_1,n} \leq \lambda_2$$

This has been done in [Luz18; Fil83; Sie75], for example, under further assumptions added to the original statement of the strong law of large numbers.

A direct approach

One can thus study the speed of convergence of these random variables by studying the speed of convergence of the sequence of real numbers $P_{\varepsilon,n}$, for each fixed ε.

This can be approached directly. That is, once can try to find a function $f : \mathbb{Q}^+ \times \mathbb{Q}^+ \to \mathbb{N}$ such that,

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \forall n \geq f(\lambda_1, \lambda_2) \, P_{\lambda_1,n} \leq \lambda_2$$

This has been done in [Luz18; Fil83; Sie75], for example, under further assumptions added to the original statement of the strong law of large numbers.

A direct approach

One can thus study the speed of convergence of these random variables by studying the speed of convergence of the sequence of real numbers $P_{\varepsilon,n}$, for each fixed ε.

This can be approached directly. That is, once can try to find a function $f : \mathbb{Q}^+ \times \mathbb{Q}^+ \to \mathbb{N}$ such that,

$$\forall \lambda_1, \lambda_2 \in \mathbb{Q}^+ \, \forall n \geq f(\lambda_1, \lambda_2) \, P_{\lambda_1,n} \leq \lambda_2$$

This has been done in [Luz18; Fil83; Sie75], for example, under further assumptions added to the original statement of the strong law of large numbers.

An example

Theorem (Luzia 2018)

Let $\{X_n\}$ is a sequence of pairwise independent, identically distributed random variables with $\mathbb{E}(X_1) = 0$, $Var(X_1) = \sigma^2 < \infty$ and $\mathbb{E}(|X_1|) = \tau < \infty$. For all $\beta > 1$, $0 < \varepsilon \leq 1$ and $n \geq n(\varepsilon, \beta)$

$$P_{n,\varepsilon} = \mathbb{P}\left(\sup_{m \geq n} |\frac{S_m}{m}| > \varepsilon\right) = \frac{\sigma^2}{n\varepsilon^2}(C_\beta + D_\beta log(n)^{\beta-1})$$

Where

$$n(\varepsilon, \beta) := \max\left\{\frac{6\tau}{\varepsilon}, exp\left(\left(\frac{9\tau}{\beta\varepsilon}\right)^{\frac{1}{\beta-1}}\right)\right\}$$

$$C_\beta := 72 + 72\beta\lfloor\beta\rfloor!$$

$$D_\beta := 72 + 72(e-1)\beta\lfloor\beta\rfloor!$$

An example

Observe that Luzia's rate is independent of the distribution of the random variables, however, the other mentioned rates ([Fil83; Sie75]) depend heavily on the distribution of the random variables, along with much stronger assumptions.

The "Baum-Katz" approach

The direct approach appears not to be that popular amongst probability theorists, who prefer the approach motivated by the following idea appearing in [BK65].

> *Given a sequence $\{c_n\}$ of bounded, non-negative numbers converging to zero, one method of measuring the rate of convergence is to determine which, if any, of the series $\sum n^r c_n$ converge where $r \geq -1$*

There are many strong law of large numbers type results, (by which we mean results about sequences of random variables concluding that $\frac{S_n}{n} \to 0$ a.s.) [Kor17; CS16; Pet69] for example. In most cases where probability theorists find rates for these results, this is the approach taken [Kor18; Kuc16; Sto10] for example.

The "Baum-Katz" approach

The direct approach appears not to be that popular amongst probability theorists, who prefer the approach motivated by the following idea appearing in [BK65].

> *Given a sequence $\{c_n\}$ of bounded, non-negative numbers converging to zero, one method of measuring the rate of convergence is to determine which, if any, of the series $\sum n^r c_n$ converge where $r \geq -1$*

There are many strong law of large numbers type results, (by which we mean results about sequences of random variables concluding that $\frac{S_n}{n} \to 0$ a.s.) [Kor17; CS16; Pet69] for example. In most cases where probability theorists find rates for these results, this is the approach taken [Kor18; Kuc16; Sto10] for example.

The "Baum-Katz" approach

The direct approach appears not to be that popular amongst probability theorists, who prefer the approach motivated by the following idea appearing in [BK65].

> *Given a sequence $\{c_n\}$ of bounded, non-negative numbers converging to zero, one method of measuring the rate of convergence is to determine which, if any, of the series $\sum n^r c_n$ converge where $r \geq -1$*

There are many strong law of large numbers type results, (by which we mean results about sequences of random variables concluding that $\frac{S_n}{n} \to 0$ a.s.) [Kor17; CS16; Pet69] for example. In most cases where probability theorists find rates for these results, this is the approach taken [Kor18; Kuc16; Sto10] for example.

The Baum-Katz-Chow theorem

Theorem (c.f. [Cho73; BK65])

Let $\{X_n\}$ be a sequence of iid random variables satisfying $\mathbb{E}(X_1) = 0$ and let $r \geq -1$. The following are equivalent:

(i) $\mathbb{E}(|X_1|^{r+2}) < \infty$

(ii) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\left|\frac{1}{n} S_n\right| > \varepsilon\right) < \infty$

(iii) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\sup_{m \geq n} \left|\frac{1}{m} S_m\right| > \varepsilon\right) < \infty$

(iv) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\max_{1 \leq m \leq n} |S_m| > n\varepsilon\right) < \infty$

The above theorem is attributed to Baum, Katz and Chow. When probability theorists find "rates" for strong law of large number type results, it is usually demonstrated that one of the above sums converges for their situation.

Ergodic theory

Theorem (Birkhoff's pointwise ergodic theorem)

Let τ be a measure-preserving map on the probability space. Define the operator T, on all square-integrable functions, by $Tf = f \circ \tau$.

$$A_n := \frac{1}{n}\left(f + Tf + \ldots + T^{n-1}f\right)$$

converges almost surely

In [AGT10], Avigad et al calculate a (Cauchy) rate of metastable almost uniform convergence that depends on $\|f\|_2$.
Birkhoff's Ergodic theorem can be used to prove the strong law of large numbers when we assume the random variables are independent. The rate in [AGT10] carries over to this result if we assume further that the random variables have finite second moment.

What am I doing?

How is quantitative probability theory being extended?

A large part of my most recent research has been to obtain rates for strong law of large number type results through a direct approach.

Doing this involves realising the existential quantifier in the definition of almost sure convergence in terms of computational interpretations given to the hypothesis of these results. This is done by an analysis of the proofs of these results.

Furthermore, as is the case with many proof mining results, one can sometimes demonstrate certain theorems cannot be given computable interpretations, through the construction of Specker-like sequences. In these cases, one can usually obtain a metastable result.

How is quantitative probability theory being extended?

A large part of my most recent research has been to obtain rates for strong law of large number type results through a direct approach.

Doing this involves realising the existential quantifier in the definition of almost sure convergence in terms of computational interpretations given to the hypothesis of these results. This is done by an analysis of the proofs of these results.

Furthermore, as is the case with many proof mining results, one can sometimes demonstrate certain theorems cannot be given computable interpretations, through the construction of Specker-like sequences. In these cases, one can usually obtain a metastable result.

How is quantitative probability theory being extended?

A large part of my most recent research has been to obtain rates for strong law of large number type results through a direct approach.

Doing this involves realising the existential quantifier in the definition of almost sure convergence in terms of computational interpretations given to the hypothesis of these results. This is done by an analysis of the proofs of these results.

Furthermore, as is the case with many proof mining results, one can sometimes demonstrate certain theorems cannot be given computable interpretations, through the construction of Specker-like sequences. In these cases, one can usually obtain a metastable result.

Improving known rates

Recall that in [Luz18], it is shown that if $\{X_n\}$ is a sequence of pairwise independent, identically distributed random variables with $\mathbb{E}(X_1) = 0$, $Var(X_1) = \sigma^2 < \infty$ and $\mathbb{E}(|X_1|) = \tau < \infty$ then for all $\beta > 1$

$$P_{n,\varepsilon} = \mathbb{P}\left(\sup_{m\geq n} |\frac{S_m}{m}| > \varepsilon\right) = O\left(\frac{log(n)^{\beta-1}}{n}\right)$$

This result was obtained by essentially analysing the elementary proof of the strong law of large numbers given by Etemadi in [Ete81].

However, in [Luz18] the author's analysis deviated significantly from Etemadi's proof and an analysis that closer follows this proof results in,

$$P_{n,\varepsilon} = O\left(\frac{1}{n}\right)$$

Improving known rates

Recall that in [Luz18], it is shown that if $\{X_n\}$ is a sequence of pairwise independent, identically distributed random variables with $\mathbb{E}(X_1) = 0$, $Var(X_1) = \sigma^2 < \infty$ and $\mathbb{E}(|X_1|) = \tau < \infty$ then for all $\beta > 1$

$$P_{n,\varepsilon} = \mathbb{P}\left(\sup_{m\geq n} |\frac{S_m}{m}| > \varepsilon\right) = O\left(\frac{log(n)^{\beta-1}}{n}\right)$$

This result was obtained by essentially analysing the elementary proof of the strong law of large numbers given by Etemadi in [Ete81].

However, in [Luz18] the author's analysis deviated significantly from Etemadi's proof and an analysis that closer follows this proof results in,

$$P_{n,\varepsilon} = O\left(\frac{1}{n}\right)$$

Improving known rates

Recall that in [Luz18], it is shown that if $\{X_n\}$ is a sequence of pairwise independent, identically distributed random variables with $\mathbb{E}(X_1) = 0$, $Var(X_1) = \sigma^2 < \infty$ and $\mathbb{E}(|X_1|) = \tau < \infty$ then for all $\beta > 1$

$$P_{n,\varepsilon} = \mathbb{P}\left(\sup_{m \geq n} |\frac{S_m}{m}| > \varepsilon\right) = O\left(\frac{log(n)^{\beta-1}}{n}\right)$$

This result was obtained by essentially analysing the elementary proof of the strong law of large numbers given by Etemadi in [Ete81].

However, in [Luz18] the author's analysis deviated significantly from Etemadi's proof and an analysis that closer follows this proof results in,

$$P_{n,\varepsilon} = O\left(\frac{1}{n}\right)$$

Improving known rates

One can demonstrate that this bound is optimal, in a sense.

For every $\delta > 0$, we can obtain a sequence of pairwise iid random variables with mean 0, finite first moment and finite variance, such that,

$$\mathbb{P}\left(\sup_{m \geq n} |\frac{1}{m} S_m| > \varepsilon\right) \geq \frac{\omega}{n^{1+\delta}}$$

for some $\omega > 0$.

Our construction, however, does not rule out the possibility that $\mathbb{P}\left(\sup_{m \geq n} |\frac{1}{m} S_m| > \varepsilon\right) = O\left(\frac{1}{nlog(n)}\right)$, for example.

Improving known rates

One can demonstrate that this bound is optimal, in a sense.

For every $\delta > 0$, we can obtain a sequence of pairwise iid random variables with mean 0, finite first moment and finite variance, such that,

$$\mathbb{P}\left(\sup_{m\geq n}|\frac{1}{m}S_m| > \varepsilon\right) \geq \frac{\omega}{n^{1+\delta}}$$

for some $\omega > 0$.

Our construction, however, does not rule out the possibility that $\mathbb{P}\left(\sup_{m\geq n}|\frac{1}{m}S_m| > \varepsilon\right) = O\left(\frac{1}{nlog(n)}\right)$, for example.

Improving known rates

One can demonstrate that this bound is optimal, in a sense.

For every $\delta > 0$, we can obtain a sequence of pairwise iid random variables with mean 0, finite first moment and finite variance, such that,

$$\mathbb{P}\left(\sup_{m\geq n}|\frac{1}{m}S_m| > \varepsilon\right) \geq \frac{\omega}{n^{1+\delta}}$$

for some $\omega > 0$.

Our construction, however, does not rule out the possibility that $\mathbb{P}\left(\sup_{m\geq n}|\frac{1}{m}S_m| > \varepsilon\right) = O\left(\frac{1}{nlog(n)}\right)$, for example.

A new Baum-Katz type result for pairwise independent random variables

Let us recall the Baum-Katz-Chow theorem.

Theorem (c.f. [Cho73; BK65])

Let $\{X_n\}$ be a sequence of iid random variables satisfying $\mathbb{E}(X_1) = 0$ and let $r \geq -1$. The following are equivalent:

(i) $\mathbb{E}(|X_1|^{r+2}) < \infty$

(ii) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(|\frac{1}{n} S_n| > \varepsilon\right) < \infty$

(iii) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\sup_{m \geq n} |\frac{1}{m} S_m| > \varepsilon\right) < \infty$

(iv) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\max_{1 \leq m \leq n} |S_m| > n\varepsilon\right) < \infty$

It is easy to see how (iv) and (iii) imply (ii), in particular, the implications hold without the assumption of independence. However, all other implications for the sums are not so straightforward and require independence in a crucial way.

A new Baum-Katz type result for pairwise independent random variables

Let us recall the Baum-Katz-Chow theorem.

Theorem (c.f. [Cho73; BK65])

Let $\{X_n\}$ be a sequence of iid random variables satisfying $\mathbb{E}(X_1) = 0$ and let $r \geq -1$. The following are equivalent:

(i) $\mathbb{E}(|X_1|^{r+2}) < \infty$

(ii) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(|\frac{1}{n} S_n| > \varepsilon\right) < \infty$

(iii) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\sup_{m \geq n} |\frac{1}{m} S_m| > \varepsilon\right) < \infty$

(iv) $\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\max_{1 \leq m \leq n} |S_m| > n\varepsilon\right) < \infty$

It is easy to see how (iv) and (iii) imply (ii), in particular, the implications hold without the assumption of independence. However, all other implications for the sums are not so straightforward and require independence in a crucial way.

A new Baum-Katz type result for pairwise independent random variables

The most general analogy of the Baum-Katz-Chow theorem for pairwise iid random variables is

Theorem (Bai-Chen-Sung, cf. Theorem 2.1 of [BCS14])

Suppose $\{X_n\}$ *are pairwise iid random variables with,* $\mathbb{E}(X_1) = 0$. *For all* $-1 \leq r < 0$, $\mathbb{E}(|X_1|^{2+r}) < \infty$ *iff for all* $\varepsilon > 0$

$$\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\max_{1\leq m\leq n} |S_m| > n\varepsilon\right) < \infty$$

Observe, that there is no result for random variables with finite variance.

A new Baum-Katz type result for pairwise independent random variables

The most general analogy of the Baum-Katz-Chow theorem for pairwise iid random variables is

Theorem (Bai-Chen-Sung, cf. Theorem 2.1 of [BCS14])

Suppose $\{X_n\}$ *are pairwise iid random variables with,* $\mathbb{E}(X_1) = 0$*. For all* $-1 \leq r < 0$, $\mathbb{E}(|X_1|^{2+r}) < \infty$ *iff for all* $\varepsilon > 0$

$$\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\max_{1 \leq m \leq n} |S_m| > n\varepsilon\right) < \infty$$

Observe, that there is no result for random variables with finite variance.

A new Baum-Katz type result for pairwise independent random variables

The most general analogy of the Baum-Katz-Chow theorem for pairwise iid random variables is

Theorem (Bai-Chen-Sung, cf. Theorem 2.1 of [BCS14])

Suppose $\{X_n\}$ *are pairwise iid random variables with,* $\mathbb{E}(X_1) = 0$*. For all* $-1 \leq r < 0$, $\mathbb{E}(|X_1|^{2+r}) < \infty$ *iff for all* $\varepsilon > 0$

$$\sum_{n=1}^{\infty} n^r \mathbb{P}\left(\max_{1 \leq m \leq n} |S_m| > n\varepsilon\right) < \infty$$

Observe, that there is no result for random variables with finite variance.

A new Baum-Katz type result for pairwise independent random variables

A simple corollary of our improved bound is,

Theorem

Suppose $\{X_n\}$ are pairwise independent, identically distributed random variables with, $\mathbb{E}(X_1) = 0$ and $Var(X_1) < \infty$. Then, for all $\varepsilon > 0$ and $r < 0$

$$\sum_{n=1}^{\infty} n^r P_{n,\varepsilon} < \infty$$

This result appears to be new and extends the work in [BCS14].

A new Baum-Katz type result for pairwise independent random variables

A simple corollary of our improved bound is,

Theorem

Suppose $\{X_n\}$ are pairwise independent, identically distributed random variables with, $\mathbb{E}(X_1) = 0$ and $Var(X_1) < \infty$. Then, for all $\varepsilon > 0$ and $r < 0$

$$\sum_{n=1}^{\infty} n^r P_{n,\varepsilon} < \infty$$

This result appears to be new and extends the work in [BCS14].

A new Baum-Katz type result for pairwise independent random variables

A simple corollary of our improved bound is,

Theorem

Suppose $\{X_n\}$ are pairwise independent, identically distributed random variables with, $\mathbb{E}(X_1) = 0$ and $Var(X_1) < \infty$. Then, for all $\varepsilon > 0$ and $r < 0$

$$\sum_{n=1}^{\infty} n^r P_{n,\varepsilon} < \infty$$

This result appears to be new and extends the work in [BCS14].

What results have been analysed?

Theorem (Chung's law of large numbers c.f [Chu47])

Suppose $\{X_n\}$ is a sequence of independent real-valued random variables with $\mathbb{E}(X_n) = 0$ for all $n \in \mathbb{N}$. Let $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ be a non-decreasing function such that $\frac{\phi(t)}{t}$ and $\frac{t^2}{\phi(t)}$ are non-decreasing on the positive half-line. If

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}(\phi(|X_n|))}{\phi(n)} < \infty$$

then

$$\frac{S_n}{n} \to 0 \; a.s$$

Direct rates (of convergence and metastability) for the generalisation of this result to random variables taking values in, what are known as, type-p Banach spaces, given in [Woy74] have been calculated. Furthermore, these rates are independent of the distribution of the random variables.

What results have been analysed?

Theorem (Chung's law of large numbers c.f [Chu47])

Suppose $\{X_n\}$ *is a sequence of independent real-valued random variables with* $\mathbb{E}(X_n) = 0$ *for all* $n \in \mathbb{N}$. *Let* $\phi : \mathbb{R}^+ \to \mathbb{R}^+$ *be a non-decreasing function such that* $\frac{\phi(t)}{t}$ *and* $\frac{t^2}{\phi(t)}$ *are non-decreasing on the positive half-line. If*

$$\sum_{n=1}^{\infty} \frac{\mathbb{E}(\phi(|X_n|))}{\phi(n)} < \infty$$

then

$$\frac{S_n}{n} \to 0 \; a.s$$

Direct rates (of convergence and metastability) for the generalisation of this result to random variables taking values in, what are known as, type-p Banach spaces, given in [Woy74] have been calculated. Furthermore, these rates are independent of the distribution of the random variables.

What results have been analysed?

Theorem (cf. Theorem 2.1 of [CS16])

Suppose $\{X_n\}$ is a sequence of non-negative random variables. Let $z_n = \mathbb{E}(S_n)$. Suppose $\frac{1}{n}z_n = O(1)$ and that there exists a sequence of non-negative real numbers $\{\gamma_n\}$ satisfying,

- $\mathbb{E}(|S_n - z_n|^p) \leq \sum_{k=1}^{n} \gamma_k$
- $\sum_{n=1}^{\infty} \frac{\gamma_n}{n^p} < \infty$

then

$$\frac{S_n}{n} \to 0 \ a.s$$

Again, one can calculate distribution-independent rates for this theorem. Furthermore, this theorem generalises many other strong law of large number type results ([Jab13; NAB04; KB99; Bir88] for example), so the calculated rates carry over to these.

What results have been analysed?

Theorem (cf. Theorem 2.1 of [CS16])

Suppose $\{X_n\}$ is a sequence of non-negative random variables. Let $z_n = \mathbb{E}(S_n)$. Suppose $\frac{1}{n}z_n = O(1)$ and that there exists a sequence of non-negative real numbers $\{\gamma_n\}$ satisfying,

- $\mathbb{E}(|S_n - z_n|^p) \leq \sum_{k=1}^{n} \gamma_k$
- $\sum_{n=1}^{\infty} \frac{\gamma_n}{n^p} < \infty$

then

$$\frac{S_n}{n} \to 0 \; a.s$$

Again, one can calculate distribution-independent rates for this theorem. Furthermore, this theorem generalises many other strong law of large number type results ([Jab13; NAB04; KB99; Bir88] for example), so the calculated rates carry over to these.

Looking towards the future

Looking towards the future

Along with mining proofs in probability theory, efforts are currently being made to formalise (aspects of) the theory in arithmetic in all finite types in order to explain the extraction of rates in some of these results, through metatheorems in the style of [Pis23; GK08; Koh05].

Furthermore, having such a system will not only be interesting from a logical perspective, but I believe it (along with other proof mining ideas) will allow us to further the understanding of certain phenomena in quantitative probability theory, which I shall now introduce in the form of questions.

Looking towards the future

Along with mining proofs in probability theory, efforts are currently being made to formalise (aspects of) the theory in arithmetic in all finite types in order to explain the extraction of rates in some of these results, through metatheorems in the style of [Pis23; GK08; Koh05].

Furthermore, having such a system will not only be interesting from a logical perspective, but I believe it (along with other proof mining ideas) will allow us to further the understanding of certain phenomena in quantitative probability theory, which I shall now introduce in the form of questions.

Do distribution-independent rates exist for the original strong law of large numbers?

For a sequence of pairwise iid random variables $\{X_n\}$ with expected values 0, the strong law of large numbers states that

$$\frac{S_n}{n} \to 0 \text{ a.s}$$

If we assume that $\mathbb{E}(|X_1|) < \infty$.

By [Luz18] and my improved result we observe that we obtain distribution-independent rates for this result if we assume $\mathbb{E}(|X_1|^2) < \infty$. Furthermore, my rates for [Chu47] demonstrate that we get distribution-independent rates for $\mathbb{E}(|X_1|^{1+\delta}) < \infty$ for all $0 < \delta \leq 1$.

However, it does not appear that one can obtain distribution-independent rates, if we assume that $\mathbb{E}(|X_1|) < \infty$. So one could ask if it is possible or if there is a Specker-like sequence demonstrating its impossibility.

Do distribution-independent rates exist for the original strong law of large numbers?

For a sequence of pairwise iid random variables $\{X_n\}$ with expected values 0, the strong law of large numbers states that

$$\frac{S_n}{n} \to 0 \text{ a.s}$$

If we assume that $\mathbb{E}(|X_1|) < \infty$.
By [Luz18] and my improved result we observe that we obtain distribution-independent rates for this result if we assume $\mathbb{E}(|X_1|^2) < \infty$. Furthermore, my rates for [Chu47] demonstrate that we get distribution-independent rates for $\mathbb{E}(|X_1|^{1+\delta}) < \infty$ for all $0 < \delta \leq 1$.

However, it does not appear that one can obtain distribution-independent rates, if we assume that $\mathbb{E}(|X_1|) < \infty$. So one could ask if it is possible or if there is a Specker-like sequence demonstrating its impossibility.

Do distribution-independent rates exist for the original strong law of large numbers?

For a sequence of pairwise iid random variables $\{X_n\}$ with expected values 0, the strong law of large numbers states that

$$\frac{S_n}{n} \to 0 \text{ a.s}$$

If we assume that $\mathbb{E}(|X_1|) < \infty$.
By [Luz18] and my improved result we observe that we obtain distribution-independent rates for this result if we assume $\mathbb{E}(|X_1|^2) < \infty$. Furthermore, my rates for [Chu47] demonstrate that we get distribution-independent rates for $\mathbb{E}(|X_1|^{1+\delta}) < \infty$ for all $0 < \delta \leq 1$.

However, it does not appear that one can obtain distribution-independent rates, if we assume that $\mathbb{E}(|X_1|) < \infty$. So one could ask if it is possible or if there is a Specker-like sequence demonstrating its impossibility.

Do rates exist in the central limit theorem?

The central limit theorem states,

Theorem (The central limit theorem)

Let $\{X_n\}$ be iid random variables satisfying $\mathbb{E}(X_1) = 0$, $Var(X_1) < \infty$. Let

$$S_n = \frac{\sum_{i=1}^{n} X_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$$

, F_n be the cumulative distribution function of S_n and Φ the cumulative distribution function of the standard normal distribution. Then for all $x \in \mathbb{R}$

$$F_n(x) \to \Phi(x)$$

.

Do rates exist in the central limit theorem?

The central limit theorem states,

Theorem (The central limit theorem)

Let $\{X_n\}$ be iid random variables satisfying $\mathbb{E}(X_1) = 0$, $Var(X_1) < \infty$. Let

$$S_n = \frac{\sum_{i=1}^{n} X_i}{\sqrt{\sum_{i=1}^{n} \sigma_i^2}}$$

, F_n be the cumulative distribution function of S_n and Φ the cumulative distribution function of the standard normal distribution. Then for all $x \in \mathbb{R}$

$$F_n(x) \to \Phi(x)$$

.

Do rates exist in the central limit theorem?

Although the theorem is true for random variables with with finite second moment, distribution-independent rates have been calculated if we assume finite third moment. The first of these theorems is attributed to Berry and Esseen,

Theorem (Berry [Ber41] and Esseen [Ess42])

Let $\{X_n\}$ be iid random variables satisfying $\mathbb{E}(X_1) = 0, Var(X_1) = \sigma^2 > 0, \mathbb{E}(|X_1|^3) = \rho < \infty$. Let

$$S_n = \frac{\sum_{i=1}^n X_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$$

F_n be the cumulative distribution function of S_n and Φ the cumulative distribution function of the standard normal distribution. There exists $C > 0$ such that, for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$,

$$|F_n(x) - \Phi(x)| \leq \frac{C\rho}{\sqrt{n}\sigma^3}$$

Do rates exist in the central limit theorem?

Although the theorem is true for random variables with with finite second moment, distribution-independent rates have been calculated if we assume finite third moment. The first of these theorems is attributed to Berry and Esseen,

Theorem (Berry [Ber41] and Esseen [Ess42])

Let $\{X_n\}$ be iid random variables satisfying $\mathbb{E}(X_1) = 0, Var(X_1) = \sigma^2 > 0, \mathbb{E}(|X_1|^3) = \rho < \infty$. Let

$$S_n = \frac{\sum_{i=1}^n X_i}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$$

F_n be the cumulative distribution function of S_n and Φ the cumulative distribution function of the standard normal distribution. There exists $C > 0$ such that, for all $n \in \mathbb{N}$ and $x \in \mathbb{R}$,

$$|F_n(x) - \Phi(x)| \le \frac{C\rho}{\sqrt{n}\sigma^3}$$

Do rates exist in the central limit theorem?

As with the strong law of large numbers, it was also shown that distribution-independent rates exist if we assume $\mathbb{E}(|X_1|^{2+\delta})$ for all $0 < \delta \leq 1$, [Kat63].

So a similar question can be asked. Do distribution-independent rates exist if we only assume $\mathbb{E}(|X_1|^2) < \infty$ or can we show that no such rates exist?

Probability theorists have thought about this.

Do rates exist in the central limit theorem?

As with the strong law of large numbers, it was also shown that distribution-independent rates exist if we assume $\mathbb{E}(|X_1|^{2+\delta})$ for all $0 < \delta \leq 1$, [Kat63].

So a similar question can be asked. Do distribution-independent rates exist if we only assume $\mathbb{E}(|X_1|^2) < \infty$ or can we show that no such rates exist?

Probability theorists have thought about this.

Do rates exist in the central limit theorem?

As with the strong law of large numbers, it was also shown that distribution-independent rates exist if we assume $\mathbb{E}(|X_1|^{2+\delta})$ for all $0 < \delta \leq 1$, [Kat63].

So a similar question can be asked. Do distribution-independent rates exist if we only assume $\mathbb{E}(|X_1|^2) < \infty$ or can we show that no such rates exist?

Probability theorists have thought about this.

Do rates exist in the central limit theorem?

In [Ber41], Berry states,

> *The author originally developed the early sections of this paper for the case of bounded variates, and is indebted to W. Feller who urged the study, in these sections, of the case of finite third order absolute moment*

Furthermore, probability theorists believe that no such rates exist. In the popular graduate probability textbook [Gut13] the author states,

> *The ultimate goal would be to move down to nothing more than second moments, but this is not feasible, since finite variances only yield existence of the limit. More has to be assumed in order to obtain a rate result.*

Do rates exist in the central limit theorem?

In [Ber41], Berry states,

> *The author originally developed the early sections of this paper for the case of bounded variates, and is indebted to W. Feller who urged the study, in these sections, of the case of finite third order absolute moment*

Furthermore, probability theorists believe that no such rates exist. In the popular graduate probability textbook [Gut13] the author states,

> *The ultimate goal would be to move down to nothing more than second moments, but this is not feasible, since finite variances only yield existence of the limit. More has to be assumed in order to obtain a rate result.*

Do rates exist in the central limit theorem?

In [Ber41], Berry states,

> *The author originally developed the early sections of this paper for the case of bounded variates, and is indebted to W. Feller who urged the study, in these sections, of the case of finite third order absolute moment*

Furthermore, probability theorists believe that no such rates exist. In the popular graduate probability textbook [Gut13] the author states,

> *The ultimate goal would be to move down to nothing more than second moments, but this is not feasible, since finite variances only yield existence of the limit. More has to be assumed in order to obtain a rate result.*

Can we explain the previous two phenomena?

We have seen two theorems about sequences of random variables that are true if we assume the random variables have finite pth-moment (say). From the proofs of these theorems, it does not appear to be possible to obtain distribution-independent rates, however, assuming finite $(p + \delta)$th moments does allow us to do so.

Could explain this proof theoretically? Can we show that formalising the proofs of these results in a certain logical system guarantees the existence of distribution-independent rates after we upgrade the moment condition?

Can we explain the previous two phenomena?

We have seen two theorems about sequences of random variables that are true if we assume the random variables have finite pth-moment (say). From the proofs of these theorems, it does not appear to be possible to obtain distribution-independent rates, however, assuming finite $(p+\delta)$th moments does allow us to do so.

Could explain this proof theoretically? Can we show that formalising the proofs of these results in a certain logical system guarantees the existence of distribution-independent rates after we upgrade the moment condition?

Thank you!

References I

References

[Ber41] Andrew C Berry. "The accuracy of the Gaussian approximation to the sum of independent variates". In: *Transactions of the american mathematical society* 49.1 (1941), pp. 122–136.

[Ess42] Carl-Gustav Esseen. "On the Liapunov limit error in the theory of probability". In: *Ark. Mat. Astr. Fys.* 28 (1942), pp. 1–19.

[Chu47] Kai-lai Chung. "Note on Some Strong Laws of Large Numbers". In: *American Journal of Mathematics* 69.1 (1947), pp. 189–192. ISSN: 00029327, 10806377. URL: http://www.jstor.org/stable/2371664 (visited on 06/16/2023).

References II

[Spe49] Ernst Specker. "Nicht Konstruktiv Beweisbare Sätze der Analysis". In: *Journal of Symbolic Logic* 14 (1949), pp. 145–158.

[Kre51] G. Kreisel. "On the Interpretation of Non-Finitist Proofs, Part I". In: *Journal of Symbolic Logic* 16 (1951), pp. 241–267. DOI: 10.1017/S0022481200100581.

[Kre52] G. Kreisel. "On the Interpretation of Non-Finitist Proofs, Part II: Interpretation of Number Theory". In: *Journal of Symbolic Logic* 17 (1952), pp. 43–58. DOI: 10.2307/2267457.

[Kat63] Melvin L Katz. "Note on the Berry-Esseen theorem". In: *The Annals of Mathematical Statistics* 34.3 (1963), pp. 1107–1108.

[BK65] Leonard E Baum and Melvin Katz. "Convergence rates in the law of large numbers". In: *Transactions of the American Mathematical Society* 120.1 (1965), pp. 108–123.

References III

[Pet69] Valentin Vladimirovich Petrov. "On the strong law of large numbers". In: *Theory of Probability & Its Applications* 14.2 (1969), pp. 183–192.

[Cho73] YS Chow. "Delayed sums and Borel summability of independent, identically distributed random variables". In: *Bull. Inst. Math. Acad. Sinica* 1.2 (1973), pp. 207–220.

[Woy74] WA Woyczynski. "Random series and laws of large numbers in some Banach spaces". In: *Theory of Probability & Its Applications* 18.2 (1974), pp. 350–355.

[Sie75] D Siegmund. "Large deviation probabilities in the strong law of large numbers". In: *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete* 31.2 (1975), pp. 107–113.

References IV

[Ete81] Nasrollah Etemadi. "An elementary proof of the strong law of large numbers". In: *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete* 55.1 (1981), pp. 119–122.

[Fil83] James Allen Fill. "Convergence Rates Related to the Strong Law of Large Numbers". In: *The Annals of Probability* 11.1 (1983), pp. 123–142. DOI: 10.1214/aop/1176993663. URL: https://doi.org/10.1214/aop/1176993663.

[Bir88] Thomas Birkel. "A note on the strong law of large numbers for positively dependent random variables". In: *Statistics & probability letters* 7.1 (1988), pp. 17–20.

[KB99] Tae-Sung Kim and Jong Il Baek. "The strong laws of large numbers for weighted sums of pairwise quadrant dependent random variables". In: *Journal of the Korean Mathematical Society* 36.1 (1999), pp. 37–49.

References V

[NAB04] HR Nili Sani, HA Azarnoosh, and A Bozorgnia. “The strong law of large numbers for pairwise negatively dependent random variables”. In: *Iranian Journal of Science* 28.2 (2004), pp. 211–217.

[Koh05] Ulrich Kohlenbach. “Some logical metatheorems with applications in functional analysis”. In: *Transactions of the American Mathematical Society* 357.1 (2005), pp. 89–128.

[GK08] Philipp Gerhardy and Ulrich Kohlenbach. “General logical metatheorems for functional analysis”. In: *Transactions of the American Mathematical Society* 360.5 (2008), pp. 2615–2660.

[AGt10] J. Avigad, P. Gerhardy, and H. towsner. “Local stability of ergodic averages”. In: *Transactions of the American Mathematical Society* 362.1 (2010), pp. 261–288. DOI: 10.1090/S0002-9947-09-04814-4.

References VI

[Sto10] George Stoica. “Rate of convergence in Petrov’s strong law of large numbers”. In: *Adv. Appl. Statist. Sci* 4 (2010), pp. 57–60.

[ADR12] J. Avigad, E. Dean, and J. Rute. “A metastable dominated convergence theorem”. In: *Journal of Logic and Analysis* (Feb. 2012). DOI: 10.1184/R1/6490742.v1. URL: https://kilthub.cmu.edu/articles/journal_contribution/A_metastable_dominated_convergence_theorem/6490742.

[Gut13] Allan Gut. *Probability: a graduate course*. Springer, 2013.

[Jab13] H Jabbari. “On almost sure convergence for weighted sums of pairwise negatively quadrant dependent random variables”. In: *Statistical Papers* 54 (2013), pp. 765–772.

References VII

[BCS14] P Bai, P-Y Chen, and SH Sung. “On complete convergence and the strong law of large numbers for pairwise independent random variables”. In: *Acta Mathematica Hungarica* 142 (2014), pp. 502–518.

[CS16] Pingyan Chen and Soo Hak Sung. “A strong law of large numbers for nonnegative random variables and applications”. In: *Statistics & Probability Letters* 118 (2016), pp. 80–86. ISSN: 0167-7152. DOI: https://doi.org/10.1016/j.spl.2016.06.017. URL: https://www.sciencedirect.com/science/article/pii/S0167715216300980.

[Kuc16] A Kuczmaszewska. “Convergence rate in the Petrov SLLN for dependent random variables”. In: *Acta Mathematica Hungarica* 148 (2016), pp. 56–72.

References VIII

[Kor17] VM Korchevsky. “On the Strong Law of Large Numbers for a Sequence of Independent Random Variables”. In: *Journal of Mathematical Sciences* 225.5 (2017), pp. 788–791.

[Kor18] VM Korchevsky. “On the Rate of Convergence in the Strong Law of Large Numbers for Nonnegative Random Variables”. In: *Journal of Mathematical Sciences* 229.6 (2018), pp. 719–727.

[Luz18] Nuno Luzia. “A SIMPLE PROOF OF THE STRONG LAW OF LARGE NUMBERS WITH RATES”. In: *Bulletin of the Australian Mathematical Society* 97.3 (2018), pp. 513–517. DOI: 10.1017/S0004972718000059.

[Pis23] Nicholas Pischke. “Logical metatheorems for accretive and (generalized) monotone set-valued operators”. In: *Journal of Mathematical Logic* (2023), p. 2350008.