# ZERO-ONE LAWS

MLADEN SAVOV

## 1. Introduction

The zero-one laws are one of the several fascinating features exhibited by a sequence of independent random variables. The aim of this file is to discuss them also in terms of random variables which is a bit easier and more illustrative compared to the usual statement of the Kolmogorov's zero one law which is in terms of sigma algebras.

In this file we will assume the availability of a common probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$ for a sequence of independent random variables $(\xi_n)_{n\geq 1}, \xi_n : \Omega \mapsto \mathbb{R}$ on this probability space. We set $S_n = \sum_{k=1}^{n} \xi_k, n \geq 0$, and when in addition $\xi_k$'s are *identically* distributed $S_n$ is the well-known *random walk*. Assume the latter and that $\mathbb{E}[\xi_1] = \mu$ with $\mathbb{E}[|\xi_1|] < \infty$. Then the Strong Law of Large Numbers (SLLN) states that

$$\lim_{n\to\infty} \frac{S_n}{n} = \mu \text{ almost surely} \tag{1.1}$$

or equivalently $\mathbb{P}\left(\lim_{n\to\infty} \frac{S_n}{n} = \mu\right) = 1$. Note that **implicitly** we have that each random variable $\xi_k$ depends on $\omega \in \Omega$ ($\xi_k = \xi_k(\omega)$) and in this sense

$$\{\lim_{n\to\infty} \frac{S_n}{n} = \mu\} = \{\omega \in \Omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \mu\} = A \subseteq \Omega.$$

Then (1.1) yields that $\mathbb{P}(A) = 1$. Let us try to explain why without any prior knowledge other than the common probability space and the independence of the random variables necessarily $\mathbb{P}(A) \in \{0, 1\}$. The name ***zero-one*** stems from the fact that such probabilities can be only zero or one.

## 2. Kolmogorov's zero-one law

Each random variable $\xi_k$ generates information in the form of sigma sub-algebra of $\mathcal{F}$. Denote $\mathcal{F}_k = \sigma(\xi_k) = \{\xi_k^{-1}(C) : C \in \mathcal{B}(\mathbb{R})\} \subseteq \mathcal{F}$. $\mathcal{F}_k$ consists of all events $B \in \mathcal{F}$ whose occurrence or not can be fully established by the knowledge of the realization of $\xi_k$. This is why $\mathcal{F}_k$ is related to the information brought in by $\xi_k$.

We come back to the example related to (1.1). Note that for any $K \geq 1$ we have that

$$\begin{aligned}\lim_{n\to\infty} \frac{S_n}{n} &= \lim_{n\to\infty} \frac{\sum_{k=1}^{K} \xi_k}{n} + \lim_{n\to\infty} \frac{\sum_{k=K+1}^{n} \xi_k}{n} \\ &= \lim_{n\to\infty} \frac{\sum_{k=K+1}^{n} \xi_k}{n}.\end{aligned} \tag{2.1}$$

We conclude that the limit ***does not really depend on the information*** generated by the first $K$ random variables and $K$ is arbitrary. Therefore, the limit does not depend on any finite number of random variables. Let us try to formalize this fact.

The information generated by all random variables is the smallest sigma algebra $\mathcal{G}_1$ generated by all invidual sigma-algebras $\mathcal{F}_k, k \geq 1$. Therefore, $\mathcal{G}_1 = \sigma(\xi_1, \xi_2, \ldots) = \sigma\left(\bigcup_{k=1}^{\infty} \mathcal{F}_k\right)$. Similarly, $\mathcal{G}_{K+1} = \sigma(\xi_{K+1}, \xi_{K+2}, \ldots) = \sigma\left(\bigcup_{k=K+1}^{\infty} \mathcal{F}_k\right)$ is the information generated by all random variables except the first $K$. Note that $\mathcal{G}_1 \supseteq \mathcal{G}_2 \supseteq \mathcal{G}_3 \supseteq \cdots$. ***The events that do not depend on the information generated by any finite number of random variables are contained in the tail sigma-algebra***

$\mathcal{T} = \bigcap_{k=1}^{\infty} \mathcal{G}_k$. Note that

$$\mathcal{G}_1 \supseteq \mathcal{G}_2 \supseteq \mathcal{G}_3 \supseteq \cdots \supseteq \mathcal{T}.$$

Since $\mathcal{T} \subseteq \mathcal{G}_K$ for any $K \geq 1$ and event $A \in \mathcal{T}$ does not depend on $(\xi_1, \xi_2, \ldots, \xi_K)$.

**Theorem 2.1.** *Let $(\xi_n)_{n\geq 1}$ be a sequence of independent random variables on a common probability space $\{\Omega, \mathcal{F}, \mathbb{P}\}$ and let $\mathcal{T}$ be the tail sigma algebra. If $A \in \mathcal{T}$ then either $\mathbb{P}(A) = 1$ or $\mathbb{P}(A) = 0$.*

**Remark 2.2.** *One might be tempted to think that since $\mathcal{T} = \bigcap_{k=1}^{\infty} \mathcal{G}_k$ then only $\emptyset, \Omega \in \mathcal{T}$. It is not the case but the intuition is right in the sense that any $A \in \mathcal{T}$ must be indistinguishable in probability from $\emptyset, \Omega$. In the end its probability is zero or one.*

**Remark 2.3.** *From* (2.1) *we deduct that $\mathbb{P}\left(\omega \in \Omega : \lim_{n\to\infty} \frac{S_n(\omega)}{n} = \mu\right)$ is either zero or one. No further information is needed to have this dichotomy.*

*Sketch of proof.* Let $A \in \mathcal{T}$. Since $\mathcal{T} = \bigcap_{k=1}^{\infty} \mathcal{G}_k$ and $A \in \mathcal{G}_k$ for any $k \geq 1$ then $A$ is independent of $(\xi_1, \xi_2, \ldots, \xi_{k-1})$ since $\mathcal{G}_k = \sigma(\xi_k, \cdots)$ is independent of $(\xi_1, \xi_2, \ldots, \xi_{k-1})$. Hence $A$ is independent of $\sigma(\xi_1, \cdots, \xi_{k-1})$. Since this independence is valid for any $k$ it means that $A$ does not depend $\mathcal{G}_1 = \sigma(\xi_1, \xi_2, \cdots) \supseteq \mathcal{T}$ and hence does not depend on $\mathcal{T}$ and ***thereby does not depend on itself***. Therefore,

$$\mathbb{P}(A) = \mathbb{P}(A \cap A) = \mathbb{P}(A)^2$$

or $\mathbb{P}(A) \in \{0, 1\}$. □

2.1. **Borel-Cantelli lemma.** If we have a sequence of mutually independent events $(A_n)_{n\geq 1}$ part of $\{\Omega, \mathcal{F}, \mathbb{P}\}$ we can associate them with a sequence of independent Bernoulli random variables on $\{\Omega, \mathcal{F}, \mathbb{P}\}$ by $\xi_n(\omega) = 1, \omega \in A_n$ and $\xi_n(\omega) = 0, \omega \in A_n^c$. The event ***infinitely often*** can be defined as

$$A^* := \{A_n \text{ i.o.}\} = \bigcap_{n=1}^{\infty} \bigcup_{k\geq n} A_k = \{\omega \in \Omega : \sum_{n=1}^{\infty} \xi_n(\omega) = \infty\} \tag{2.2}$$

and describes the event whereby infinitely many of the events $A_n$ occur simultaneously. The last is a representation in terms of random variables and the sum being infinite means we have infinitely many one's.

The Borel-Cantelli lemma says that

$$\begin{aligned} \sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \mathbb{P}(\xi_n = 1) = \infty &\iff \mathbb{P}(A^*) = 1 \\ \sum_{n=1}^{\infty} \mathbb{P}(A_n) = \sum_{n=1}^{\infty} \mathbb{P}(\xi_n = 1) < \infty &\iff \mathbb{P}(A^*) = 0. \end{aligned} \tag{2.3}$$

Relation (2.3) suggests $A^*$ is in the tail sigma-algebra. Without going into technicalities let us check that $A^*$ does not depend on any finite number of random variables $\xi_n$. This is clear from (2.2) since for any $K \geq 1$

$$\sum_{n=1}^{\infty} \xi_n(\omega) = \infty \iff \sum_{n=K}^{\infty} \xi_n(\omega) = \infty.$$

So the zero-one law applies it remains to establish when $\mathbb{P}(A^*)$ is zero or one, which is specified in (2.3).

2.2. **Exponential sum.** Consider $(\xi_n)_{n\geq 1}$ of independent, identically distributed random variables on the same $\{\Omega, \mathcal{F}, \mathbb{P}\}$ and the exponential sum

$$E := \sum_{n=1} e^{-S_n} \tag{2.4}$$

and the general sum

$$F := \sum_{n=1} f(S_n). \tag{2.5}$$

Whereas $\{E < \infty\}$ can be niftily reduced to the Kolmogorov's zero-one law by writing

$$E = e^{-\xi_1}(1 + \sum_{n=2} e^{-\sum_{k=2}^{n} \xi_k})$$

for $F$ it is not clear. So we could deduce that $\mathbb{P}(E < \infty) \in \{0, 1\}$. However, our intuition says that may be so at least when the underlying random variables are identically distributed.

## 3. Hewitt-Savage zero-one law

Consider $(\xi_n)_{n\geq 1}$ of independent, identically distributed random variables on the same $\{\Omega, \mathcal{F}, \mathbb{P}\}$. We say that a set $A$ is permutationally invariant if a permutation of any finite number of random variables $(\xi_1, \xi_2, \ldots, \xi_k)$ does not effect the occurrence of $A$. For example, take

$$A = \{\omega \in \Omega : \sum_{n=1} f(S_n(\omega)) < \infty\}. \tag{3.1}$$

Pick any $K \geq 1$ and permute the way you like the first $K$ random variables getting the sequence $\{\xi_{\sigma(1)}, \xi_{\sigma(2)}, \cdots, \xi_{\sigma(K)}, \xi_{K+1}, \xi_{K+2}, \cdots\}$ and denote by $S_n^{\sigma}$ the sum of the first $n$ elements of this sequence. Clearly, if $n \geq K$ we have that $S_n^{\sigma} = S_n$ and the occurance of $A$ does not depend on this permutation. Therefore, $A$ is permutationally invariant. The zero-one law of Hewitt-Savage states that

**Theorem 3.1.** *Consider $(\xi_n)_{n\geq 1}$ of independent, identically distributed random variables on the same $\{\Omega, \mathcal{F}, \mathbb{P}\}$. For any permutationally invariant $A$ we have that either $\mathbb{P}(A) = 1$ or $\mathbb{P}(A) = 0$.*

As a consequence we deduce that

$$\mathbb{P}\left(\omega \in \Omega : \sum_{n=1} f(S_n(\omega)) < \infty\right) \in \{0, 1\}.$$

Remarkably simple once the foundation is laid up!