A Brief Review of Zero-One Laws and Some Applications

Jiaxin Lu, 518030910412

July 1, 2020

Abstract

This document tangled Zero-One Laws in Probability including Borel-Cantelli Lemmas, Kolmogorov's Zero-One Law, Hewitt-Savage Zero-One Law, and Lévy's Zero-One Law. I provided the statement and the proof or the idea of proof to all these lemmas and theorems. I also explored the extensions of Borel-Cantelli Lemma and showed the development of these extensions. Finally, I finished the exercises given in the section E4 of Probability with Martingales, and discussed some applications of the Zero-One Law.

Contents

1 Borel-Cantelli Lemmas

1.1 Preliminaries

Definition 1.1. A probability triple or probability space is a triple $(\Omega, \mathcal{F}, P)$ consisting of:

- The sample space Ω;
- $\mathcal{F}$, a σ-algebra of Ω; and
- The probability measure P, a function $P : \mathcal{F} \to [0,1]$, such that $P(\emptyset) = 0$, $P(\Omega) = 1$ and P is countably additive on disjoint sets.

Now, given a probability triple $(\Omega, \mathcal{F}, P)$, we call Ω the sample space, and any element $\omega \in \omega$ a sample point. We call elements of $\mathcal{F}$ an event and so $\mathcal{F}$ is known as the family of events.

Definition 1.2. Given a probability triple $(\Omega, \mathcal{F}, P)$ and a sequence of events $(E_n)_{n\in\mathbb{N}}$ is a sequence of events.

(a) We define

$$\begin{aligned}(E_n, \text{i.o.}) &:= (E_n \text{ infinitely often}) \\ &:= \limsup E_n := \bigcap_m \bigcup_{n\geq m} E_n \\ &= \{\omega : \text{for every } m,\ \exists n(\omega) \geq m \text{ such that } \omega \in E_{n(\omega)}\} \\ &= \{\omega : \omega \in E_n \text{ for infinitely many } n\}.\end{aligned}$$

(b) We define

$$\begin{aligned}(E_n, \text{ev}) &:= (E_n \text{ eventually} \\ &:= \liminf E_n := \bigcup_m \bigcap_{n\geq m} E_n \\ &= \{\omega : \text{for some } m(\omega), \omega \in E_n, \forall n \geq m(\omega)\} \\ &= \{\omega : \omega \in E_n \text{ for all large } n\}.\end{aligned}$$

(c) Note that $(E_n, \text{ev})^c = (E_n^c, \text{i.o.})$.

1.2 The Borel-Cantelli Lemma

Theorem 1.3 (The Borel-Cantelli Lemma)**.** *Let $E_1, E_2, \ldots \in \mathcal{F}$,*

(i) If $\sum_n P(E_n) < \infty$, then $P(E_n, i.o.) = 0$.

(ii) If $\sum_n P(E_n) = \infty$, E_n are pairwise independent, then $P(E_n, i.o.) = 1$.

Proof. (i) Note that $\forall m \in \mathbb{N}$, we have:

$$P(E_n, \text{i.o.}) = P(\bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} E_k) \leq P(\bigcup_{k=m}^{\infty} E_k)$$

monotonically. Then follows from countable subadditivity,

$$P(\bigcup_{k=m}^{\infty} E_k) \leq \sum_{k=m}^{\infty} P(E_k)$$

Hence, $\forall \epsilon > 0$, there exists $m \in \mathbb{N}$ such that $\sum_{k=m}^{\infty} P(E_k) < \epsilon$. It follows that $P(E_n, \text{i.o.}) < \epsilon$ for any ϵ, hence,

$$P(E_n, \text{i.o.}) = 0$$

(ii) First we prove following properties.

From $\log(1-x) \leq -x \; \forall x \in [0,1]$, we have,

$$\begin{aligned} \log(\prod_{n\in\mathbb{N}} (1-y_n)) &= \sum_{n\in\mathbb{N}} \log(1-y_n) \\ &\leq -\sum_{n\in\mathbb{N}} y_n \end{aligned}$$

Thus,

$$\prod_{n\in\mathbb{N}} (1-y_n) = \exp(-\sum_{n\in\mathbb{N}} y_n) = 0 \tag{1.1}$$

Now using this property, we finish our proof.

Let $A = \limsup_{n\to\infty} E_n$. We shall prove that $P(A^c) = 0$. Let $B_i = \bigcap_{n=i}^{\infty} E_n^c$. Then $A^c = \bigcup_{i=1}^{\infty} B_i$. So, we shall prove that $P(B_i) = 0$ for all i. Now, for each i and $k > i$,

$$\begin{aligned} P(B_i) &= P(\bigcap_{n=i}^{\infty} E_n^c) \\ &\leq P(\bigcap_{n=i}^{k} E_n^c) = \prod_{n=i}^{k} [1 - P(E_n)] \end{aligned}$$

Use Eq. (1.1), we can derive,

$$P(B_i) = \prod_{n=i}^{k} [1 - P(E_n)] = 0$$

Thus, $P(B_i) = 0$ for all $i \in \mathbb{N}$.

$\square$

1.3 Extensions of Borel-Cantelli Lemma

1.3.1 Extension of the First Borel-Cantelli Lemma

Theorem 1.4. *If* $\liminf_{n\to\infty} P(E_n) = 0$ *and* $\sum_{n=1}^{\infty} P(E_n \cap E_{n+1}^c) < \infty$*, then one has* $P(E_n, i.o.) = 0$*.*

Proof. Let $B_n = E_n \cap E_{n+1}^c$, $F = \limsup A_n^c$, $G = \limsup E_n$. Thus we have

$$P(\limsup B_n) \leq P(F^c) = P(\liminf E_n) \leq \liminf_{n\to\infty} P(E_n) = 0$$

Observe that $F \cap G \subset \limsup B_n$. To see this, fix $n \geq 1$ and $\omega \in F \cap G$. As $\omega \in G$, $\exists m \geq n$ such that $\omega \in E_m$.

Put $l = \inf\{k > m : \omega \in E_k^c\}$, by the definition of F, $l < \infty$ and $\omega \in A_l \cap A_{l-1}^c$. Since n is arbitrary, $\omega \in \limsup B_n$.

Hence,

$$P(G) \leq P(F^c) + P(F \cap G) = 0.$$

Thus, $P(E_n, \text{i.o.}) = 0$. □

1.3.2 Extension of the Second Borel-Cantelli Lemma

Many investigations were devoted in the second Borel-Cantelli lemma in attempt to weaken the independence condition. Bellow is one form of extension which raised in the class and I'll discuss more versions later.

Theorem 1.5. *If* $\sum_{n=1}^{\infty} P(A_n) = \infty$ *and*

$$\liminf_{n\to\infty} \frac{\sum_{j=1}^{n}\sum_{k=1}^{n} P(A_i A_j)}{(\sum_{j=1}^{n} P(A_j))^2} = 1 \tag{1.2}$$

then $P(A_n, i.o.) = 1$*.*

This is first proved by Erdös and Rényi [1959], I rewrote the whole proof below.

Proof. Define $a_n = 1_{A_n}$. Then we have $E(a_k) = P(A_k)$, and $E(a_k a_l) = P(A_k A_l)$ and thus putting $\eta_n = \sum_{k=1}^{n} a_k$, we have,

$$\frac{\sum_{k=1}^{n}\sum_{l=1}^{n} P(A_k A_l)}{(\sum_{k=1}^{n} P(A_k))^2} = \frac{E(\eta_n^2)}{E^2(\eta_n)}$$

Thus condition Eq. (1.2) can be written in the equivalent form

$$\liminf_{n\to\infty} \frac{E(\eta_n^2)}{E^2(\eta_n)} = 1 \tag{1.3}$$

or as $E(\eta_n^2) = D^2(\eta_n) + E^2(\eta_n)$, also in form

$$\liminf_{n\to\infty} \frac{D^2(\eta_n)}{E^2(\eta_n} = 0 \tag{1.4}$$

Now by the inequality of Chebyshev according to which for any random variable η we have

$$P(|\eta - E(\eta)| \geq \lambda D(\eta)) \leq \frac{1}{\lambda^2} \quad \text{ if } \lambda > 1, \tag{1.5}$$

we have for any ϵ with $0 < \epsilon < 1$

$$P(\eta_n \leq (1-\epsilon)E(\eta_n)) \leq \frac{D^2(\eta_n)}{\epsilon^2 E^2(\eta_n)}. \tag{1.6}$$

If Eq. (1.4) holds, we can find a sequence $n_k(n_1 < n_2 < \cdots)$ such that

$$\sum_{k=1}^{\infty} \frac{D^2(\eta_n)}{E^2(\eta_n)} < \infty. \tag{1.7}$$

It follows from Eq. (1.6) and Eq. (1.7) that

$$\sum_{k=1}^{\infty} P(\eta_{n_k} \leq (1-\epsilon)E(\eta_{n_K})) < \infty \tag{1.8}$$

Using the first Borel-Cantelli lemma, it follows that almost surely, $\eta_{n_K} \geq (1-\epsilon)E(\eta_{n_k})$ except for a finite number of values of k. As by supposition $\lim_{k\to\infty} E(\eta_{n_k}) = \infty$, it follows that η_{n_k} tends to ∞ with probability 1, which implies that $P(A_n, \text{i.o.}) = 1$. □

More general results have also been proved independently by Kochen and Stone [1964] and Spitzer [1964].

Erdös and Rényi [1959] also gives that the condition of pairwise independence of events can be replace by the weaker condition $P(A_k A_j) \leq P(A_k)P(A_j)$ for every k and j such that $k \neq j$.

Lamperti [1963] formulated the following proposition.

Proposition 1.6. *If $\sum_{n=1}^{\infty} P(A_n) = \infty$ and $P(A_k A_j) \leq CP(A_k)P(A_j)$ for all $k, j > N$ and some constants C and N, then $P(A_n, i.o.) > 0$.*

Petrov [2002] extends the conclusion to $P(A_n, \text{i.o.}) > 1/C$.

Ortega and Wschebor [1984] modified the condition to the following proposition.

Proposition 1.7. *If $\sum_{n=1}^{\infty} P(A_n) = \infty$ and*

$$\liminf_{n\to\infty} \frac{\sum_{1\leq j<k\leq n}(P(A_j \cap A_k) - P(A_j)P(A_k))}{(\sum_{i=1}^{n} P(A_i))^2} \leq 0$$

then $P(A_n, i.o.) = 1$.

Later, Petrov [2004] extends it more with the following proposition.

Proposition 1.8. *If $\sum_{n=1}^{\infty} P(A_n) = \infty$, let H be an arbitrary real constant, put*

$$\alpha_H = \liminf \frac{\sum_{1\le i<k\le n}(P(A_iA_k) - HP(A_i)P(A_k))}{(\sum_{k=1}^{n} P(A_k))^2}$$

Then,

$$P(A_n, i.o.) \ge \frac{1}{H + 2\alpha_H}.$$

When $H = 0$, it could generate the result of Kochen and Stone [1964] and Spitzer [1964].

When $H = 1$, then $P(A_n, \text{i.o.}) \ge 1/\gamma$ where $\gamma = 1 + 2\alpha_1$. This proposition is new, but it could be generated from the result of Ortega and Wschebor [1984].

It could also generate the result of Petrov [2002] since $H + 2\alpha_H \ge 1$ must hold, and we apply $H \ge 1$, and then $\alpha_H \le 0$ so that $H + 2\alpha_H \le H$.

2 Kolmogorov's Zero-One Law

Theorem 2.1. *Let $(X_n : n \in \mathbb{N})$ be a sequence of independent random variables, and let $\mathcal{T}$ be its tail σ-algebra. Then $\mathcal{T}$ is P-trivial:*

(i) $F \in \mathcal{T} \Longrightarrow P(F) = 0$ *or* $P(F) = 1$,

(ii) ξ *is* $\mathcal{T}$*-measurable* $\Longrightarrow \exists c \in [-\infty, \infty]$ *s.t.* $P(\xi = c) = 1$

2.1 A Normal Proof

Below is a proof of Kolmogorov's Zero-One Law using the independence property of π-system.

Proof. (i) Let $\mathcal{X}_n = \sigma(X_1, \ldots, X_n), \mathcal{T}_n = \sigma(X_{n+1}, X_{n+2}, \ldots)$

(1) We claim that $\mathcal{X}_n$ and $\mathcal{T}_n$ are independent.
Set

$$\mathcal{K}_n = \{\{\omega : X_i(\omega) \le x_i, i = 1, \ldots, n\} : x_i \in \mathbb{R} \cup \{\infty\}, i = 1, \ldots, n\}$$
$$\mathcal{J}_n = \bigcup\{\{\omega : X_j(\omega) \le x_j : j = n+1, \ldots, n+r\}x_{n+1}, \ldots x_{n+r}\}$$

Thus, $\sigma(\mathcal{K}_n) = X_n, \sigma(\mathcal{J}_n) = \mathcal{T}_n$.
Since the sequence (X_k) is independent, $\mathcal{K}$ and $\mathcal{J}$ are independent. Therefore, by the independence of π-systems, we have our claim holds.

(2) $\mathcal{X}_n$ and $\mathcal{T}$ are independent.
This is obvious since $\mathcal{T} \subseteq \mathcal{T}_n$

(3) $\mathcal{X}_\infty := \sigma(X_n : n \in \mathbb{N})$ and $\mathcal{T}$ are independent.
Since $\mathcal{X}_n \subseteq \mathcal{X}_{n+1}, \forall n$, the class $\mathcal{K}_\infty := \bigcup \mathcal{X}_n$ is a π-system which generates $\mathcal{X}_\infty$. Moreover, $\mathcal{K}_\infty$ and $\mathcal{T}$ are independent. Thus, our claim holds.

(4) $\mathcal{T} \subseteq \mathcal{X}_\infty$, $\mathcal{T}$ is independent of $\mathcal{T}$.

$$F \in \mathcal{T} \Longrightarrow P(F) = P(F \cap F) = P(F)P(F)$$

Thus, $P(F) = 0$ or $P(F) = 1$.

(ii) By part (i), for every $x \in \mathbb{R}$, $P(\xi \leq x) = 0$ or 1.
Let $c = \sup\{x : P(\xi \leq x) = 0\}$. Then, if $c = \pm\infty$, it's clear that $P(\xi = c) = 1$.
Suppose c is finite. Then $P(\xi \leq c - 1/n) = 0, \forall n$, so,

$$P(\cup\{\xi \leq c - 1/n\}) = P(\xi < c) = 0,$$

since, $P(\xi \leq c + 1/n) = 1, \forall n$, we have

$$P(\cap\{\xi \leq c + 1/n\}) = P(\xi \leq c) = 1.$$

Therefore, $P(\xi = c) = 1$.

□

2.2 A Martingale Proof

Below is a martingale proof of Kolmogorov's Zero-One Law. We first introduce the Lévy's 'Upward' Theorem.

Theorem 2.2 (Lévy's 'Upward' Theorem). *Let $\xi \in \mathcal{L}^1(\Omega, \mathcal{F}, P)$ and define $M_n := E(\xi|\mathcal{F}_n)$, a.s.. Then M is a UI martingale and*

$$M_n \to \eta := E(\xi|\mathcal{F}_\infty), \quad a.s.$$

and in $\mathcal{L}^1$

Now we can use this theorem to prove the first part of Kolmogorov's Zero-One Law.

Proof. Let $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$. Let $F \in \mathcal{T}$, and let $\eta = 1_F$. Since $\eta \in \mathcal{F}_\infty$, Lévy's 'Upward' Theorem shows,

$$\eta = E(\eta|\mathcal{F}_\infty) = \lim E(\eta|\mathcal{F}_n), \quad \text{a.s.}$$

Thus, η only takes the values 0 or 1.

Moreover, for each n, η is $\mathcal{T}_n$ measurable, and hence is independent of $\mathcal{F}_n$. Hence, by the Rôle of Independence,

$$E(\eta|\mathcal{F}_n) = E(\eta) = P(F), \quad \text{a.s.}$$

Thus, $P(F) = 0$ or 1. □

3 Hewitt–Savage Zero–One Law

Theorem 3.1. *Suppose that $(X_n)_{n\in\mathbb{N}}$ is a sequence of independent and identically distributed r.v. (i.i.d). Then every permutation invariant event has probability* 0 *or* 1*.*

Remark 3.2. Consider the σ-algebra $\mathcal{F}_n = \sigma(X_1, \dots, X_n)$. This means that with the growth of n, we would receive a larger $\mathcal{F}_n$, and with $\lim \mathcal{F}_n = \mathcal{F}_\infty$.

Set f as $\mathcal{F}_\infty$-measurable, from Doob–Dynkin Lemma, it equals to $f = V \circ X, X = (X_1, X_2, \dots)$, that is, $f = V(X_1, X_2, \dots)$

For arbitrary permutation p and $f \in m\mathcal{F}_\infty$,

$$f \circ p = (V \circ X) \circ p := V \circ (X \circ p) = V(X_{p(1)}, X_{p(2)}, \dots)$$

Permutation invariant means that for every finite permutation p, $f = f \circ p$ defined as above. We call an event A is permutation invariant if the r.v. 1_A is permutation invariant.

Thus this 0-1 Law means that with i.i.d $(X_n)_{n\in\mathbb{N}}$, every permutation invariant event has probability 0 or 1.

Idea of Proof. Let A be a permutation invariant event. To show $P(A) = 0$ or 1 under the condition of the theorem, we can form a sequence of independent r.v.s 1_{A_n} and let $\mathcal{T}$ be its tail σ-algebra. Thus, by Kolmogorov's 0-1 law, we can find the conclusion holds.

4 Lévy's Zero-One Law

Theorem 4.1. *If $\mathcal{F}_n \uparrow \mathcal{F}_\infty$ and $A \in \mathcal{F}_\infty$, then $E(1_A|\mathcal{F}_n) \to 1_A$ a.s..*

Idea of Proof. By applying $X = 1_A$ to the Lévy's 'Upward' Theorem (Theorem 2.2), thus $1_A = E(1_A|\mathcal{F}_\infty)$, and the conclusion follows.

5 Some Exercises and Applications

Exercise 5.1 (E4.4 of Probability with Martingales). Suppose that a coin with probability p of heads is tossed repeatedly. Let A_k be the event that a sequence of k (or more) consecutive heads occurs amongst tosses numbered $2^k, 2^k+1, 2^k+2, \dots, 2^{k+1}-1$. Prove that

$$P(A_k, \text{i.o.}) = \begin{cases} 1, & \text{if } p \geq 1/2, \\ 0, & \text{otherwise.} \end{cases}$$

Proof. (i) We first prove $P(A_k, \text{i.o.}) = 1, p \geq 1/2$.

Let $E_{i,k}$ be the event that there are k consecutive heads beginning at toss numbered $2^k + (i-1)k$ and we have $1 \leq i \leq 2^k/k$. That is, the k consecutive heads occurs at toss numbered $2^k, 2^k+k, \ldots, 2^{k+1}-k$. These events $E_{i,k}$ are pairwise independent, and $\{E_{i,k}, \text{i.o.}\} \implies \{A_k, \text{i.o.}\}$. Since,

$$\sum_k \sum_{i=1}^{\frac{2^k}{k}} P(E_{i,k}) \geq \sum_k (\frac{2^k}{k} - 1)p^k$$
$$\geq \sum_k \frac{1}{k} - \frac{p}{1-p} = \infty$$

From the second Borel-Cantelli lemma, we have $P(E_{i,k}\text{i.o.}) = 1$. Therefore, $P(A_k, \text{i.o.}) = 1, p \geq 1/2$.

(ii) We next prove $P(A_k, \text{i.o.}) = 1, p < 1/2$

Let E_i be the event that there are k consecutive heads beginning at toss numbered i. Thus, we have

$$A_k = \bigcup_{n=2^k}^{2^{k+1}-k} B_n.$$

By inclusion-exclusion principle, we have

$$P(A_k) \leq \sum_{n=2^k}^{2^{k+1}} P(B_n) \leq 2^k p^k.$$

When $p < \frac{1}{2}$, $\sum_k P(A_k) \leq \infty$. Therefor by the first Borel-Cantelli lemma, we have $P(A_k, \text{i.o.}) = 0$.

□

Exercise 5.2 (E4.6 of Probability with Martingales, Converse to SLLN). Let Z be a non-negative RV. Let Y be the integer part of Z. Show that

$$Y = \sum_{n \in \mathbb{N}} 1_{\{Z \geq n\}},$$

and deduce that

$$\sum_{n \in \mathbb{N}} P[Z \geq n] \leq E(Z) \leq 1 + \sum_{n \in \mathbb{N}} P[Z \geq n] \tag{5.1}$$

Let (X_n) be a sequence of IID RVs with $E(|X_n|) = \infty, \forall n$. Prove that

$$\sum_n P[|X_n| > kn] = \infty (k \in \mathbb{N}) \text{ and } \limsup \frac{|X_n|}{n} = \infty, \quad \text{a.s.}$$

Deduce that if $S_n = X_1 + X_2 + \cdots + X_n$, then

$$\limsup \frac{|S_n|}{n} = \infty \quad \text{a.s.}$$

Proof. First, by the definition of $1_{\{Z\geq n\}}$, we can find that it has value 1 only if $n \leq Z$. Thus take the sum of the numbers of such n will have the integer part of Z.

Then, since,

$$P[Z \geq n] = 1_{\{Z\geq n\}}$$

then we have

$$\begin{aligned}
\sum_{n\in\mathbb{N}} P[Z \geq n] &= \sum_{n\in\mathbb{N}} 1_{\{Z\geq n\}} = Y \\
\sum_{n\in\mathbb{N}} P[Z \geq n] &= 1 + Y \\
E(Z) &= Z
\end{aligned}$$

Since Y is the integer part of Z, $Y \leq Z \leq Y + 1$ holds naturally and Eq. (5.1) holds.

It's clear that $\sum_n P[|X_n| > kn] \geq E(|X_n|)/k - 1$ which is also ∞ since $E(|X_n|) = \infty$, thus

$$\sum_n P[|X_n| > kn] = \infty$$

holds naturally.

Since we have $\sum_n P[|X_n| > kn] = \infty, (k \in \mathbb{N})$, we can deduce that,

$$\sum_n P[\frac{|X_n|}{n} > k] = \infty, \quad k \in \mathbb{N}$$

From the second Borel-Cantelli lemma, we have $P(\frac{|X_n|}{n} > k, \text{i.o.}) = 1$. Since k is arbitrary, we have $P(\frac{|X_n|}{n}, \text{i.o.}) = 1$, that is

$$\limsup \frac{|X_n|}{n} = \infty, \text{a.s.}$$

Then, from

$$\frac{X_n}{n} = \frac{S_n}{n} - (\frac{n-1}{n})\frac{S_{n-1}}{n-1},$$

we have

$$\frac{|X_n|}{n} \leq \frac{|S_n|}{n} + (\frac{n-1}{n})\frac{|S_{n-1}|}{n-1},$$

and thus, we have

$$\limsup_n \frac{|X_n|}{n} \leq 2 \limsup_n \frac{|S_n|}{n}.$$

Therefore, $\limsup_n |X_n|/n = \infty$ implies $\limsup_n |S_n|/n = \infty$. □

Exercise 5.3 (E4.7 of Probability with Martingales, What's fair about a fair game). Let $X_1, X_2, \ldots$ be independent RVs such that

$$X_n = \begin{cases} n^2 - 1 & \text{with probability } n^{-2} \\ -1 & \text{with probability } 1 - n^{-2} \end{cases}$$

Prove that $E(X_n) = 0, \forall n$, but that if $S_n = X_1 + X_2 + \cdots + X_n$, then

$$\frac{S_n}{n} \to -1, \quad \text{a.s.}$$

Proof. First,

$$\begin{aligned} E(X_n) &= (n^2 - 1) \times n^{-2} + (-1) \times (1 - n^{-2}) \\ &= \frac{n^2 - 1}{n^2} - \frac{n^2 - 1}{n^2} \\ &= 0 \end{aligned}$$

Then, since

$$\begin{aligned} \sum_n P(X_n \neq -1) &= \sum_n P(X_n = n^2 - 1) \\ &= \sum_n n^{-2} \\ &< 2 < \infty, \end{aligned}$$

from the first Borel-Cantelli lemma, we have $P(X_n \neq -1, \text{i.o.}) = 0$. Thus, almost surely, there exists some finite N such that $X_n = -1$ for every $n \geq N$. Every such sequence (X_n) is such that $S_n/n \to -1$ hence, $P(S_n/n \to -1) = 1$.

□

Example 5.4 (Monkey typing Shakespeare). Let us agree that correctly typing WS, the Collected Works of Shakespeare, amounts to typing a particular sequence of N symbols on a type-writer.

A monkey types symbols at random, one per unit time, producing a finite sequence (X_n) of IID RVs with values in the set of all possible symbols. We agree that,

$$\epsilon := \inf\{P(X_1 = x) : x \text{ is a symbol } \} > 0.$$

Let H be the event that the monkey produces infinitely many copies of WS.

Use the Second Borel-Cantelli Lemma to prove that $P(H) = 1$.

Proof. Since WS has a finite many symbols, let it be N. Assume that there are c possible characters that the monkey could type and there are no delete keys on the typewriter.

If we pick an arbitrary starting point in the finite string, the probability that this is the beginning of the full text of WS is

$$P(H) = (\frac{1}{c})^N = \epsilon > 0$$

Consider a sequence of events $E_1, E_{N+1}, E_{2N+1}, \dots$ where E_i is the event that the i-th character is the star of WS-length substring of our infinite string with no overlap. Clearly, $P(E_i) = \epsilon$. Since,

$$\sum_{i=1}^{\infty} P(E_{Ni+1}) = \sum_{i=1}^{\infty} \epsilon = \infty$$

Thus, by the second Borel-Cantelli lemma, we have $P(S_i, \text{i.o.}) = 1$. In other word, $P(H) = 1$, that is the monkey will eventually type the WS. $\square$

References

P. Erdös and A. Rényi. On cantor's series with convergent $\sum 1/q_n$. *Ann. Univ. Sci. Budapest. Eötvös Sect. Math.*, 2:93–109, 1959.

Simon Kochen and Charles Stone. A note on the borel-cantelli lemma. *Illinois J. Math.*, 8(2):248–251, 06 1964. doi: 10.1215/ijm/1256059668.

J. Lamperti. Wiener's test and markov chains. *Journal of Mathematical Analysis and Applications*, 6(1):58–66, 1963. doi: 10.1016/0022-247X(63)90092-1.

J. Ortega and M. Wschebor. On the sequence of partial maxima of some random sequences. *Stochastic Processes and their Applications*, 16(1):85–98, 1984. doi: 10.1016/0304-4149(84)90177-7.

V.V. Petrov. A note on the borel-cantelli lemma. *Statistics and Probability Letters*, 58 (3):283–286, 2002. doi: 10.1016/S0167-7152(02)00113-X.

V.V. Petrov. A generalization of the borel-cantelli lemma. *Statistics and Probability Letters*, 67(3):233–239, 2004. doi: 10.1016/j.spl.2004.01.008.

Frank Spitzer. *Principles of Random Walk.* Springer-Verlag New York, 1964. ISBN 0072-5285. doi: 10.1007/978-1-4757-4229-9.