Borel-Cantelli's lemmas: theoretical and practical applications

The goal of this problem is to derive the first and second Borel-Cantelli lemmas and to explore various applications of these results, both within probability theory and in practical contexts. The second Borel-Cantelli lemma, sometimes referred to as *Borel's zero–one law*, is particularly noteworthy as it belongs to the class of *zero-one laws*, which are powerful tools in probability theory. (We will study Kolmogorov's zero–one law and its applications in a subsequent problem.)

Preliminaries: $\liminf$ and $\limsup$ of a sequence

Let $(v_n)_{n\geq 0}$ be a sequence of real numbers. We consider $(u_n)_{n\geq 0}$ and $(w_n)_{n\geq 0}$ the two sequences of $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$ defined by

$$\forall n \in \mathbb{N}, \quad u_n = \inf_{k\geq n} v_k, \quad w_n = \sup_{k\geq n} v_k.$$

1. Show that $(u_n)_{n\geq 0}$ and $(w_n)_{n\geq 0}$ are respectively nondecreasing and nonincreasing.
2. Show that $(u_n)_{n\geq 0}$ and $(w_n)_{n\geq 0}$ converge either to $\pm\infty$ or to a finite real number (i.e. convergence in $\overline{\mathbb{R}}$). These limits are respectively denoted by $\liminf_{n\to+\infty} v_n$ and $\limsup_{n\to+\infty} v_n$.
3. We call limit point of $(v_n)_{n\geq 0}$ an element ℓ of $\overline{\mathbb{R}}$ such that there exists a subsequence of $(v_n)_{n\geq 0}$ converging towards ℓ. Show that $\liminf_{n\to+\infty} v_n$ and $\limsup_{n\to+\infty} v_n$ are respectively the smallest and largest limit points of $(v_n)_{n\geq 0}$.
4. Show that $\liminf_{n\to+\infty} v_n = \limsup_{n\to+\infty} v_n$ if and only if $(v_n)_{n\geq 0}$ converges in $\overline{\mathbb{R}}$.

$\liminf$ and $\limsup$ of a family of events

We consider a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and a sequence of events (elements of the $\sigma-$algebra $\mathcal{A}$) $(A_n)_{n\geq 0}$. We also define the following sets:

$$\liminf_{n\to+\infty} A_n = \bigcup_{i\geq 0} \bigcap_{j\geq i} A_j, \quad \limsup_{n\to+\infty} A_n = \bigcap_{i\geq 0} \bigcup_{j\geq i} A_j.$$

5. Show that $\liminf_{n\to+\infty} A_n^c = \left(\limsup_{n\to+\infty} A_n\right)^c$.
6. Show that $\omega \in \liminf_{n\to+\infty} A_n$ if and only if $\exists n_0 \in \mathbb{N}, \forall n \geq n_0, \omega \in A_n$.
7. Show that $\omega \in \limsup_{n\to+\infty} A_n$ if and only if $\{n \in \mathbb{N}, \omega \in A_n\}$ is infinite.
8. Deduce that $\liminf_{n\to+\infty} A_n \subset \limsup_{n\to+\infty} A_n$.
9. Show that
$$\mathbb{P}\left(\liminf_{n\to+\infty} A_n\right) \leq \liminf_{n\to+\infty} \mathbb{P}(A_n) \leq \limsup_{n\to+\infty} \mathbb{P}(A_n) \leq \mathbb{P}\left(\limsup_{n\to+\infty} A_n\right).$$

First Borel-Cantelli's lemma

Let us assume that $\sum_{n=0}^{\infty} \mathbb{P}(A_n) < +\infty$.

10. Show that almost surely $\sum_{n=0}^{\infty} 1_{A_n} < +\infty$.
11. Deduce that $\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 0$.

Second Borel-Cantelli's lemma (also known as Borel's zero-one law)

Let us assume that $(A_n)_{n\geq 0}$ constitutes a family of independent events and that $\sum_{n=0}^{\infty} \mathbb{P}(A_n) = +\infty$.

12. For $i \leq J \in \mathbb{N}$, show that $\mathbb{P}\left(\bigcap_{j=i}^{J} A_j^c\right) = \prod_{j=i}^{J} (1 - P(A_j))$.
13. Deduce that $\mathbb{P}\left(\bigcap_{j\geq i} A_j^c\right) = 0$.

14. Conclude that $\mathbb{P}\left(\liminf_{n\to+\infty} A_n^c\right) = 0$ and $\mathbb{P}\left(\limsup_{n\to+\infty} A_n\right) = 1$.

Application of the first Borel-Cantelli's lemma: longest sequence of ones in a sequence of Bernoulli variables

We consider in this section a sequence $(X_n)_{n\geq 1}$ of i.i.d. Bernoulli variables with parameter $\frac{1}{2}$. For $n \geq 1$, we define L_n the longest sequence of ones in $X_1, \ldots, X_n$, i.e.

$$L_n = \sup\{l \in \{1,\ldots,n\} | \exists i \in \{1,\ldots,n-l+1\}, X_i = \ldots = X_{i+l-1} = 1\}.$$

15. For $k \in \{1,\ldots,n\}$, show that

$$\{L_n \geq k\} = \bigcup_{i=1}^{n-k+1} \{X_i = \ldots = X_{i+k-1} = 1\}.$$

16. Deduce that $\forall k \in \mathbb{N}^*, \mathbb{P}(L_n \geq k) \leq \frac{n}{2^k}$.
17. Let $\epsilon > 0$. For $j \geq 1$, let us consider $n_j = 1 + \left\lfloor j^{\frac{2}{\epsilon}} \right\rfloor$ and $k_j = 1 + \left\lfloor (1+\epsilon)\frac{\log(n_j)}{\log(2)} \right\rfloor$. What is the nature of the series $\sum_{j\geq 1} \mathbb{P}(L_{n_j} \geq k_j)$?
18. Use the first Borel-Cantelli's lemma to deduce that almost surely $\limsup_{j\to+\infty} \frac{L_{n_j}}{\log(n_j)} \leq \frac{1+\epsilon}{\log(2)}$.
19. Prove then that almost surely $\limsup_{n\to+\infty} \frac{L_n}{\log(n)} \leq \frac{1}{\log(2)}$.
20. For $m \in \{1,\ldots,n\}$, show that

$$\{L_n < m\} \subset \bigcap_{i=0}^{\lfloor \frac{n}{m} \rfloor - 1} \{X_{im+1} = \ldots = X_{(i+1)m} = 1\}^c.$$

21. Deduce that $\forall m \in \{1,\ldots,n\}, \mathbb{P}(L_n < m) \leq \left(1 - \frac{1}{2^m}\right)^{\lfloor \frac{n}{m} \rfloor}$.
22. Let $\epsilon > 0$. For $n \geq 1$, let us consider $m_n = 1 + \left\lfloor (1-\epsilon)\frac{\log(n)}{\log(2)} \right\rfloor$. Prove that $m_n \leq n$ and study the nature of the series $\sum_{n\geq 1} \mathbb{P}(L_n < m_n)$.
23. Use the first Borel-Cantelli's lemma to deduce that almost surely $\liminf_{n\to+\infty} \frac{L_n}{\log(n)} \geq \frac{1}{\log(2)}$.
24. Conclude that almost surely $\lim_{n\to+\infty} \frac{L_n}{\log(n)} = \frac{1}{\log(2)}$.
25. (Bonus) – Illustrate this result with a graph obtained thanks to a Python code.

(Bonus) – Application of the second Borel-Cantelli's lemma: there is no "natural" probability measure on $\mathbb{N}$

The goal of this bonus section is to prove a surprising result: there is no probability measure $\mathbb{P}$ on $\mathbb{N}$ such that the probability of the set of multiples of d is $\frac{1}{d}$ for all $d \in \mathbb{N}^*$, i.e. $\mathbb{P}(d\mathbb{N}) = \frac{1}{d}$.

For this purpose, let us assume that there is such a probability measure. In what follows, we denote by $(p_n)_{n\geq 1}$ the sequence of prime numbers.

26. For $m \geq 1$, let us define $n_m = \sup\{n | p_n \leq m\}$. Prove that

$$\{1,\ldots,m\} \subset \left\{p_1^{\epsilon_1} \ldots p_{n_m}^{\epsilon_{n_m}} k^2 | k \in \mathbb{N}^*, \epsilon_1, \ldots, \epsilon_{n_m} \in \{0,1\}\right\}.$$

27. Prove that

$$\sum_{i=1}^{m} \frac{1}{i} \leq \prod_{n=1}^{n_m} \left(1 + \frac{1}{p_n}\right) \sum_{k=1}^{m} \frac{1}{k^2}$$

28. Deduce that

$$\exists C \in \mathbb{R}, \forall m \geq 1, \sum_{n=1}^{n_m} \frac{1}{p_n} \geq \log(\log(m)) - C.$$

29. Conclude on the nature of the series $\sum_{n\geq 1} \frac{1}{p_n}$.

30. Prove that the events $(p_n\mathbb{N})_{n\geq 1}$ are independent, i.e.

$$\forall k \geq 1, \forall 1 \leq n_1 < \ldots < n_k, \mathbb{P}\left(\bigcap_{j=1}^{k} p_{n_j}\mathbb{N}\right) = \prod_{j=1}^{k} \mathbb{P}(p_{n_j}\mathbb{N}).$$

31. What is the nature of the series $\sum_{n\geq 1} \mathbb{P}(p_n\mathbb{N})$?

32. Conclude using the second Borel-Cantelli's lemma.