ON BOREL-CANTELLI LEMMAS

By Tapas K. Chandra and Subhashis Ghosal
Indian Statistical Institute, Calcutta

Abstract

In this paper, we review some extensions of the Borel-Cantelli lemmas.


The Borel-Cantelli lemmas have been found extremely useful for the derivations of many limit theorems of probability theory. They assert that

(a) if \( \{ A_n \} \) is a sequence of arbitrary events and \( \sum_{n=1}^{\infty} P(A_n) < \infty \), then \( P(\limsup A_n) = 0 \),

(b) if \( \{ A_n \} \) is a sequence of independent events satisfying \( \sum_{n=1}^{\infty} P(A_n) = \infty \), then \( P(\limsup A_n) = 1 \),

where \( \limsup A_n = \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} A_k \). These two results are obtained by Borel (1909, 1912) and Cantelli (1917). (Actually Borel (1912) provided the following criterion for \( P(\limsup A_n) = 0 \) or 1: For a sequence of events \( \{ A_n \} \), let

\[
0 < p'_n \leq P(A_n|A_1^{\varepsilon_1} \cap \cdots \cap A_{n-1}^{\varepsilon_{n-1}}) \leq p''_n < 1, \quad \forall \varepsilon_1, \ldots, \varepsilon_{n-1}, \forall n,
\]

where \( A_{n'}^i = A_i \) or \( A_{n}^{i} \). Then

\[
\sum_{n=1}^{\infty} p''_n < \infty \quad \Rightarrow \quad P(\limsup A_n) = 0,
\]

\[
\sum_{n=1}^{\infty} p'_n = \infty \quad \Rightarrow \quad P(\limsup A_n) = 1.
\]

Cantelli (1917) proved that the condition \( \sum_{n=1}^{\infty} P(A_n) < \infty \) always implies that \( P(\limsup A_n) = 0 \).

The following extension of the convergence part of the Borel-Cantelli lemma is due to Berendoff-Nielsen (1961), who also gave a nontrivial application of it.

Lemma 1. If \( \liminf_{n \to -\infty} P(A_n) = 0 \) and \( \sum_{n=1}^{\infty} P(A_n \cap A_{n+1}^c) < \infty \), then one has \( P(\limsup A_n) = 0 \).

231
Proof. Set $B_n = A_n \cap A_{n+1}^c$, $E = \limsup A_n$ and $F = \text{Emsup} A_n$. Thus we have $P(\limsup B_n) \leq P(F) = P(\liminf A_n) \leq \liminf_{n \to \infty} P(A_n) = 0$. Observe that $E \cap F \subset \limsup B_n$. To see this, fix $n \geq 1$ and $\omega \in E \cap F$. As $\omega \in E$, $\exists m \geq n$ such that $\omega \in A_m$. Put $l = \inf\{k > m : \omega \in A_k^c\}$; by 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 $P(E) \leq P(F) + P(E \cap F) = 0$.

It is now obvious that the following result is true: If $\liminf_{n \to \infty} P(A \cap A_n) = 0$ and $\sum_{n=1}^{\infty} P(A \cap A_n \cap A_{n+1}^c) < \infty$, then $P(\limsup A_n) \leq 1 - P(A)$.

A further extension of Lemma 1 is given in Barndorff-Nielsen (1961, Lemma 1**).

In many applications, the assumption of independence assumed in (b) fails to hold and needs to be replaced by more relaxed assumptions. In the rest of the paper, we review several such extensions of the second Borel-Cantelli lemma available in the literature.

**Theorem 1. (The Extended Rényi-Lamperti Lemma).** If $\sum_{n=1}^{\infty} P(A_n) = \infty$ and

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

then $P(\limsup A_n) \geq 1/c$.

(The Rényi-Lamperti lemma states that $P(\limsup A_n) \geq 2 - c$; see, e.g., Billingsley (1991, p. 87).)

Proof. Recall that if $EX > 0$ and $0 \leq \epsilon < 1$, then $P(X > \epsilon X) \geq (1 - \epsilon)^2(EX)^2/EX^2$, which can be proved by applying the Cauchy-Schwarz inequality to $E(XI(X > \epsilon EX))$. Put $J_n = \sum_{j=1}^{n} I(A_j)$ and $s_n = \sum_{j=1}^{n} P(A_j) = EJ_n$. Fix $0 < \epsilon < 1$ and put $B_n = \{J_n \geq \epsilon s_n\}$. As $s_n \to \infty$, $\limsup B_n \subset \limsup A_n$. Now $P(B_n) \geq (1 - \epsilon)^2(EJ_n)^2/EJ_n^2$, and so $\limsup_{n \to \infty} P(B_n) \geq (1 - \epsilon)^2/c$. As $\epsilon > 0$ is arbitrary, the proof is complete.

The above result contains a result of Erdős and Rényi (1959) who considered the case $c = 1$. It also contains the following corollary which unifies results of Lamperti (1963) and Kochen and Stone (1964) (see also Chow and Teicher (1988, p. 103)). The present derivation is much simpler.

**Corollary 1.** If $\sum_{n=1}^{\infty} P(A_n) = \infty$ and there exists nonnegative reals $c_1$, $c_2$ and $c_3$, $c_1 + c_2 + c_3 > 0$, such that

$$P(A_i \cap A_j) \leq (c_1 P(A_i) + c_2 P(A_j))P(A_{j-i}) + c_3 P(A_i)P(A_j), \quad i < j. \tag{2}$$
Then \( P(\limsup A_n) \geq 1/(2c_1 + 2c_2 + c_3) \).

We have the following extension, due to Chandra and Ghosal (1993), of Remark 1 of Etemadi (1983).

**Theorem 2.** Suppose \( \{A_n\} \) is a sequence of events satisfying
\[
P(A_i \cap A_j) - P(A_i)P(A_j) \leq q(j-i)P(A_j), \quad i < j,
\]
where \( q(m) \geq 0 \ \forall m \geq 1 \) and
\[
\sum_{m=1}^{\infty} q(m)/(\sum_{j=1}^{m} P(A_j)) < \infty.
\]
If \( \sum_{n=1}^{\infty} P(A_n) = \infty \), then \( P(\limsup A_n) = 1 \).

**Proof.** Without loss of generality, we can assume that \( P(A_1) > 0 \). Set \( X_k = I(A_k) \) and \( S_n = \sum_{k=1}^{n} X_k \). We shall use the following result (Theorem 4 of Chandra and Ghosal, 1993):

Let \( \{X_n\} \) be a sequence of nonnegative square integrable random variables. Assume that

(i) \( \sup_{n \geq 1} (\sum_{k=1}^{n} EX_k)/f(n) < \infty \),

(ii) There exists a double sequence \( \{\rho_{ij}\} \) of nonnegative reals with \( \rho_{ij} = 0 \) for \( i > j \) such that

(a) \( \text{var}(S_n) \leq \sum_{j=1}^{n} \sum_{j=1}^{n} \rho_{ij} \ \forall n \geq 1 \),

(b) \( \rho_{j-m,j} \leq q(m)c_j \) for \( m = 1, \ldots, j-1, j \geq 1 \),

where \( q(m) \geq 0, \ c_j \geq 0 \) and \( \{f(n)\} \) is a positive sequence increasing to infinity. If
\[
\sum_{j=1}^{\infty} \rho_{ij}/(f(j))^2 < \infty, \quad \sum_{m=1}^{\infty} q(m) \sum_{j=m+1}^{\infty} c_j/(f(j))^2 < \infty,
\]
then \( \sum_{k=1}^{n} (X_k - EX_k)/f(n) \to 0 \ a.s. \)

We apply the above result with \( f(n) = ES_n \) and derive the stronger fact that \( S_n/ES_n \to 1 \) a.s. (Note that \( \limsup A_n = \{S_n \to \infty\} \) and \( ES_n \to \infty \) by the hypothesis.) Let \( \rho_{ij} = 2\text{cov}^+(X_i, X_j), \ i \leq j, \) and zero otherwise. Clearly Assumption 1(a) holds; also \( \rho_{ij} \leq 2q(j-i)P(A_j), \ i < j, \) and \( \rho_{jj} \leq 2P(A_j) \). Now
\[
\sum_{j=2}^{\infty} (ES_j)^{-2}P(A_j) = \sum_{j=2}^{\infty} (ES_j)^{-2}(ES_j - ES_{j-1})
\]
\[
\leq \sum_{j=2}^{\infty} \int_{ES_{j-1}}^{ES_j} x^{-2}dx \leq 1/P(A_1) < \infty.
\]
Also by the same arguments,
\[
\sum_{m=1}^{\infty} q(m) \sum_{j=m+1}^{\infty} (ES_j)^{-2} P(A_j) \leq \sum_{m=1}^{\infty} q(m)/ES_m < \infty. \tag{8}
\]

Theorem 2 has the following useful consequences.

**Corollary 2.** Assume that (3) holds, \(\sum_{n=1}^{\infty} P(A_n) = \infty\) and \(q(m)\) is decreasing.

(a) If \(\sum_{m=1}^{\infty} q(m) < \infty\), then \(P(\lim \sup A_n) = 1\).

(b) If \(\lim sup_{n \to \infty} n^\alpha P(A_n) > 0\) and \(\sum_{m=1}^{\infty} q(m)m^{\alpha-1} < \infty\) for some \(0 \leq \alpha < 1\), then \(P(\lim \sup A_n) = 1\).

(c) If \(\lim sup_{n \to \infty} n^\alpha P(A_n) > 0\) and \(\sum_{m=1}^{\infty} q(m)/\log m < \infty\), then we have \(P(\lim \sup A_n) = 1\).

**Proof.** Part (a) follows trivially since \(\sum_{m=1}^{\infty} q(m)/ES_m \leq \sum_{m=1}^{\infty} q(m)/P(A_1) < \infty\). To prove (b), get a subsequence \(\{n_k\}\) and \(\epsilon > 0\) such that \(P(A_n) > \epsilon n^{-\alpha}\) \(\forall n = n_1, n_2, \ldots\). Set \(B_k = A_{n_k}\) and it suffices to show that \(P(\lim \sup B_k) = 1\). Note that for all \(i < j\),

\[
P(B_i \cap B_j) - P(B_i)P(B_j) \leq q(n_j - n_i)P(A_{n_i}) - q(j - i)P(B_j);
\]

here we have used the fact that \(q(m)\) is decreasing. Also, \(\sum_{j=1}^{k} P(B_j) \geq \delta k^{1-\alpha}\) for some \(\delta > 0\). Thus \(\sum_{m=1}^{\infty} q(m)/\sum_{j=m+1}^{\infty} P(B_j) \geq \delta^{-1} \sum_{m=1}^{\infty} m^{\alpha-1} q(m) < \infty\). Applying Theorem 2 to \(\{B_k\}\), we get the result. The proof of (c) is similar.

**Remark 1.** If \(\{f(A_n)\}\) is a pairwise \(m\)-dependent sequence, i.e., \(A_i\) and \(A_j\) are independent if \(|i - j| > m\) (a fortiori, a pairwise independent sequence), then (3) and (4) are trivially satisfied with \(q(k) = 0\) for all \(k > m\).

Recall that a sequence \(\{X_n\}\) is called pairwise negative quadrant dependent (pairwise NQD) if \(\forall i \neq j, s, t \in \mathcal{R},\)

\[
P(X_i > s, X_j > t) \leq P(X_i > s)P(X_j > t).
\]

If for a sequence of events \(\{A_n\}, \{f(A_n)\}\) is pairwise NQD, (i.e., \(P(A_i \cap A_j) \leq P(A_i)P(A_j)\)) \(\forall i \neq j\), then (3) and (4) hold trivially with \(q(m) = 0 \forall m \geq 1\).

**Definition 1.** A sequence \(\{X_n\}\) is called \(*\)-mixing if there exist \(N \geq 1\) and a sequence \(f_m \downarrow 0\) such that for all \(m \geq N\), we have

\[
|P(A \cap B) - P(A)P(B)| \leq f_m P(A)P(B), A \in \sigma(X_1, \ldots, X_n), B \in \sigma(X_{m+n}).
\]
We have the following Borel-Cantelli lemma, due to Blum, Hanson and Koopmans (1963), for \( \ast \)-mixing sequences. Some related results are obtained by Cohn (1972) and Yoshihara (1979).

**Theorem 3.** Let \( \{A_n\} \) be a sequence of events such that \( \{I(A_n)\} \) is \( \ast \)-mixing and \( \sum_{n=1}^{\infty} P(A_n) = \infty \). Then \( P(\limsup A_n) = 1 \).

**Proof.** Let \( 0 < \delta < 1 \) and get \( k \geq N \) such that \( f_k < \delta \). Get \( 1 \leq j \leq k \) such that \( \sum_{n=0}^{\infty} P(A_{nk+j}) = \infty \). Set \( B_n = A_{nk+j} \). It suffices to show that \( P(\limsup B_n) = 1 \). If not, there exists \( m \geq 1 \) such that \( P(\bigcup_{i=1}^{m} B_i) < 1 \). Hence

\[
P(\cap_{i=1}^{m} B_i) = P(B_m) + \sum_{t=1}^{\infty} P(B_m \cap \cap_{s=0}^{t-1} B_{m+s}^c)
\geq (1 - \delta)[P(B_m) + \sum_{t=1}^{\infty} P(B_{m+t})P(\cap_{s=0}^{t-1} B_{m+s}^c)]
\geq (1 - \delta)P(\cap_{s=0}^{t-1} B_{m+s}^c) \sum_{t=0}^{\infty} P(B_{m+t}) = \infty,
\]

which is a contradiction.

**Remark 2.** Note that in the proof of Theorem 3, we have only used the fact that \( \{f_m\} \) is bounded away from 1.

The following version of the Borel-Cantelli lemma is due to Serfling (1975) and extends an earlier result of Iosifescu and Theodorescu (1969).

**Theorem 4.** Let \( \{A_n\} \) be a sequence of events such that \( \sum_{n=1}^{\infty} P(A_n) = \infty \). Let \( \mathcal{F}_n = \sigma(A_1, \ldots, A_n) \). If \( \sum_{n=2}^{\infty} E[P(A_n | \mathcal{F}_{n-1}) - P(A_n)] < \infty \), then we have \( P(\limsup A_n) = 1 \).

**Proof.** Serfling (1975) has shown that for all \( m < N \),

\[
P(\cap_{n=m}^{N} A_n) \leq \exp[- \sum_{n=m}^{N} P(A_n)] + \sum_{n=m}^{N} E[P(A_n | \mathcal{F}_{n-1}) - P(A_n)].
\]

Letting \( N \to \infty \) and then \( m \to \infty \), the result follows.

We now turn to the martingale version of the Borel-Cantelli lemmas due to Lévy (1937).

**Theorem 5.** Let \( \{A_n\} \) be a sequence of events and let \( \mathcal{F}_n = \sigma(A_1, \ldots, A_n) \). Then

\[
P\{\limsup A_n \Delta \{\sum_{n=1}^{\infty} P(A_n | \mathcal{F}_{n-1}) = \infty\}\} = 0,
\]

where \( \Delta \) stands for the symmetric difference.
A proof of Theorem 5 may be found in Breiman (1968, p. 96). Some related results are given in Dubins and Freedman (1965) and Freedman (1973).

There have been some results on the necessary and sufficient conditions for $P(\limsup A_n) = 0$ or $P(\limsup A_n) = 1$. These are respectively due to Loève (1951) and Nash (1954) (see also Chung and Erdős (1952)). Note that Theorem 5 is also a result of this kind. A quantitative form of the Borel-Cantelli lemma has been obtained by Philipp (1967).

2. Applications.

For applications of the Borel-Cantelli lemmas in the strong law of large numbers and the law of iterated logarithm, the reader may consult the references given in the previous section. Here we present some illustrations which are elementary in nature.

Example 1. Let $\{X_n\}$ be a sequence of random variables such that there is a positive sequence $q(m)$ with $\sum_{m=1}^{\infty} q(m) < \infty$ and

$$P\{X_i > s, X_j > t\} - P\{X_i > s\}P\{X_j > t\} \leq q(j-i)P\{X_j > t\},$$

$$P\{X_i < s, X_j < t\} - P\{X_i < s\}P\{X_j < t\} \leq q(j-i)P\{X_j < t\},$$

(13)

for all $i < j, s, t \in \mathbb{R}$. Then for any $a \in \mathbb{R}$, $P\{X_n \to a\} = 0$ or 1. (If $X_n$'s are pairwise $m$-dependent or pairwise NQD, then the above condition is trivially satisfied.)

For a proof, note that it is sufficient to show

$$P(\limsup_{n \to \infty} X_n \leq a) = 0 \text{ or } 1, \quad P(\liminf_{n \to \infty} X_n \geq a) = 0 \text{ or } 1. \quad (14)$$

Observe that $\limsup_{n \to \infty} X_n \leq a = (\bigcup_{m=1}^{\infty} \{X_n > a + 1/m\})^c$.

If $\forall m \geq 1$, $\sum_{m=1}^{\infty} P\{X_n > a + 1/m\} < \infty$, then $P(\limsup_{n \to \infty} X_n \leq a) = 1$.

If $\exists m \geq 1$ such that $\sum_{n=1}^{\infty} P\{X_n > a + 1/m\} = \infty$, then the events $\{X_n > a + 1/m\}$, $n \geq 1$, satisfy the hypotheses of Corollary 2(a) and hence $P(\limsup_{n \to \infty} X_n \leq a) = 0$. The other part of (14) can be proved by replacing $X_n$ by $-X_n$ and $a$ by $-a$.

In a similar manner, if $X_n$'s are $\ast$-mixing, then one can apply Theorem 3 and arrive at the same conclusion.

Example 2. Assume that $X_n$'s satisfy (13). Then $P\{X_n \text{ converges to a finite limit}\} = 0$ or 1.

To see this, let, with the usual convention regarding the empty set,

$$x_0 = \inf\{r \in \mathbb{R} : P\{X_n > r \text{ i.o.}\} = 0\}, \quad -\infty \leq x_0 \leq \infty,$$

$$y_0 = \inf\{r \in \mathbb{R} : P\{X_n < r \text{ i.o.}\} = 0\}, \quad -\infty \leq y_0 \leq \infty.$$
Then, by elementary real analysis, it follows that

$$P\left( \limsup_{n \to \infty} X_n \leq x_0 \right) = 1, \quad P\left( \liminf_{n \to \infty} X_n \geq y_0 \right) = 1$$ (15)

(and hence \( y_0 \leq x_0 \)). If \( x_0 = -\infty \) or \( y_0 = \infty \), (15) implies that \( P\{X_n \text{ converges}\} = 0 \). If \( y_0 = x_0 \in \mathcal{R} \), \( P\{X_n \text{ converges}\} = 1 \). If \( y_0 \leq x_0 \), find \( u, v \in \mathcal{R} \) such that \( y_0 < u < v < x_0 \).

Then, by definition of \( x_0 \) and \( y_0 \), we obtain

$$P\{X_n < u \text{ i.o.}\} > 0, \quad P\{X_n > v \text{ i.o.}\} > 0. \quad (16)$$

Applying Corollary 2(a) to the events \( \{X_n < u\} \) and \( \{X_n > v\} \), \( n \geq 1 \), respectively, we get

$$P\{X_n < u \text{ i.o.}\} = 1, \quad P\{X_n > v \text{ i.o.}\} = 1. \quad (17)$$

Clearly (17) implies that \( P\{X_n \text{ converges}\} = 0 \).

A similar remark applies to the *-mixing case.

**Example 3.** Let \( \{Y_n\} \) be a sequence of strictly positive iid integer valued random variables and let \( E_i = \{Y_1 + \cdots + Y_k = i \text{ for some } k\} \). Then clearly

$$P(E_i \cap E_j) = P(E_i)P(E_{j-i}), \quad 1 \leq i < j.$$ Let \( \{m_n\} \) be any subsequence with \( \sum_{n=1}^{\infty} P(E_{m_n}) = \infty \). Then, applying Corollary 1 with \( c_1 = 1 \) and \( c_2 = c_3 = 0 \), \( P(E_{m_n} \text{ occurs i.o.}) \geq 1/2(> 0) \). Now, with the aid of the Hewitt-Savage zero-one law, one can actually conclude that

$$P(E_{m_n} \text{ occurs i.o.}) = 1.$$ For further illustrations of the Borel-Cantelli lemmas, the reader may consult Breiman (1968), Billingsley (1991) and Chow and Teicher (1988).

**REFERENCES**


