university of groningen
faculty of mathematics and natural sciences

On Generalizations Of The Borel-Cantelli Lemmas

S. de Vos

Bachelor Thesis in Mathematics

July 9, 2010

On Generalizations Of The Borel-Cantelli Lemmas

Summary

This bachelor thesis is about the Borel-Cantelli lemmas and ways one can generalize them. I will give the original version of the lemmas and their proofs and then look at further research that has been done on these lemmas. I will try to explain how the lemmas can be generalized, give some results published in articles that are about the lemmas and provide proof of certain of these results.

Bachelor Thesis in Mathematics
Author: S. de Vos
Supervisor: A.C.D. van Enter & C. Külske
Date: July 9, 2010

Institute of Mathematics and Computing Science
P.O. Box 407
9700 AK Groningen
The Netherlands

Contents

Chapter 1

The Borel-Cantelli lemmas

1.1 About the Borel-Cantelli lemmas

Although the mathematical roots of probability are in the sixteenth century, when mathematicians tried to analyse games of chance, it wasn't until the beginning of the 1930's before there was a solid mathematical axiomatic foundation of probability theory. The beginning of the twentieth century was a time when especially French and Russian mathematicians did a lot of research on the foundations of probability theory.

The development of measure theory and its entanglement with probability theory, along with contributions by many influential mathematicians (Émile Borel, Francesco Cantelli, Paul Lévy, Maurice Fréchet, Norbert Wiener, Aleksandr Khinchin, among many others), culminated in the now famous work *Grundbegriffe der Wahrscheinlichkeitsrechnung* (1933) by Andrei Kolmogorov. It laid down the axiomatic basis of probability theory upon which mathematicians have been building ever since.

In this bachelor thesis I will discuss two particular lemmas which were proved during the time probability theory was in its infancy. They are called the Borel-Cantelli lemmas, named after mathematicians Émile Borel and Francesco Cantelli. The lemmas tell us, when given a sequence of events $\{A_n\}_{n=1}^{\infty}$ in a probability space $(\Omega, \mathcal{F}, P)$, what conditions must hold in order for a finite or infinite number of A_n to occur. The lemmas are important results in probability theory because they are used to prove a lot of other important statements. For instance, the Strong Law of Large Numbers is proved using the Borel-Cantelli lemmas. Over the last couple of decades there have been quite a few mathematicians who have done research on these lemmas. Their interests lie in finding more generalized versions of the Borel-Cantelli lemmas.

There are a number of ways in one can generalize the Borel-Cantelli lemmas, some of which we will see in this article. But first let us look at the standard version of the Borel-Cantelli lemmas.

1.2 The Standard Version Of The Borel-Cantelli Lemmas

Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of events on a probability space $(\Omega, \mathcal{F}, P)$. The Borel-Cantelli lemmas (in short: "BC-lemmas") are as follows:

First BC-lemma

If $\sum_{n=1}^{\infty} P(A_n) < \infty$ *then* $P(\limsup_{n\to\infty} A_n) = 0$

Second BC-lemma

If $\sum_{k=1}^{\infty} P(A_n) = \infty$ *and if the sequence* $\{A_n\}_{n=1}^{\infty}$ *consists of mutually independent events then* $P(\limsup_{n\to\infty} A_n) = 1$

1.2.1 On the notation

The meaning of the expression $\limsup_{n\to\infty} A_n$ is perhaps not immediately obvious. First, note that

$$\limsup_{n\to\infty} A_n = \lim_{n\to\infty} (\sup_{j\geq n} A_j)$$

What is the supremum of a collection of sets? Remember that "supremum" means "least upper bound". Because all the subsets of our sample space can be partially ordered by inclusion so we can indeed speak of an "upper bound". The supremum of a collection of elements in a partially ordered set is its least upper bound, so we know that $\sup_{j\geq n} A_j$ should be a set and it should hold that $A_j \subset \sup_{j\geq n} A_j$ for all $j \geq n$. Because the supremum should also be the smallest upper bound, it is not hard to see that $\sup_{j\geq n} A_j = \bigcup_{j=n}^{\infty} A_j$. Therefore,

$$\lim_{n\to\infty} (\sup_{j\geq n} A_j) = \lim_{n\to\infty} (\bigcup_{j=n}^{\infty} A_j)$$

We can see what this limit means because, since

$$\bigcup_{j=n}^{\infty} A_j \supset \bigcup_{j=n+1}^{\infty} A_j \supset \bigcup_{j=n+2}^{\infty} A_j \supset \ldots$$

it is just the greatest lower bound, or infimum. In an similar way as with the supremum, we can see that the greatest lower bound is the intersection of all $\bigcup_{j=n}^{\infty} A_j$ sets. So we have found

$$\lim_{n\to\infty} (\sup_{j\geq n} A_j) = \lim_{n\to\infty} (\bigcup_{j=n}^{\infty} A_j) = \bigcap_{n=1}^{\infty} \bigcup_{j=n}^{\infty} A_j$$

and this last expression is the same event as the event that the A_j occur infinitely often. This is often abbreviated to "i.o.". So saying

$$P(\limsup_{n\to\infty} A_n) = 1$$

is the same as saying $P(A_n \text{ i.o.}) = 1$, or: " infinitely many A_n occur with probability 1". Similarly, $P(A_n \text{ i.o.}) = 0$ is the same as saying "only a finite number of A_n occur".

Note that the second lemma is a partial converse of the first lemma. The second lemma is not valid without the independence criterion. This can make the second lemma sometimes hard to apply.

1.3 Proof Of The Lemmas

The proof of the first lemma is not very hard to understand. Suppose that $\sum_{n=1}^{\infty} P(A_n) < \infty$. Then

$$P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) \leq P(\bigcup_{k=n}^{\infty} A_k) \leq \sum_{k=n}^{\infty} P(A_k)$$

The expression $\sum_{k=n}^{\infty} P(A_k)$ goes to zero for $n \to \infty$ because the sum $\sum_{n=1}^{\infty} P(A_n)$ converges. So we have

$$P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) \leq \sum_{k=n}^{\infty} P(A_k) \to 0$$

which implies that

$$P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = 0$$

and this is what we wanted to prove.

The proof of the second part is a little bit more elaborate. Suppose $\sum_{n=1}^{\infty} P(A_n) = \infty$ and that the events $\{A_i\}_{i=1}^{\infty}$ are mutually independent. Note that

$$P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = 1 \text{ if and only if } P(\bigcup_{k=n}^{\infty} A_k) = 1 \text{ for all n}$$

Now, instead of trying to show that $P(\bigcup_{k=n}^{\infty} A_k) = 1$ we start by looking at the probability of its complementary event, $\bigcap_{k=n}^{\infty} A_k^c$. If we can show that the probability of this event is zero we are done.

Because the events $\{A_n\}$ are independent the events $\{A_n^c\}$ are also independent. Now we can write

$$\begin{aligned}
P(\bigcap_{k=n}^{\infty} A_k^c) &= \lim_{N\to\infty} P(\bigcap_{k=n}^{N} A_k^c) \\
&= \lim_{N\to\infty} \prod_{k=n}^{N} P(A_k^c) \\
&= \lim_{N\to\infty} \prod_{k=n}^{N} (1 - P(A_k))
\end{aligned}$$

Now we use the fact that $1 - x \leq e^{-x}$ for all $x \in \mathbb{R}$. This can easily be seen by comparing the left-hand side with the Taylor expansion of the right-hand side.

$$\begin{aligned}\lim_{N\to\infty}\prod_{k=n}^{N}(1-P(A_k)) &\leq \lim_{N\to\infty}\prod_{k=n}^{N}e^{-P(A_k)} \\ &= \lim_{N\to\infty} e^{-\sum_{k=n}^{N}P(A_k)}\end{aligned}$$

Since $\sum_{k=n}^{N} P(A_k) \to \infty$ for $N \to \infty$ it follows that

$$\lim_{n\to\infty} e^{-\sum_{k=n}^{N}P(A_k)} \to 0$$

So we have

$$P(\bigcap_{k=n}^{\infty} A_k^c) = 0$$

which implies

$$P(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k) = 1$$

and this is what we wanted to show.

1.4 An Application of the First Borel-Cantelli lemma

As previously mentioned, the BC-lemmas are being used in proofs of many mathematical statements. Let us look at one example of the first BC-lemma in action.

1.4.1 The Infinite Monkey Theorem

Another famous theorem in probability theory, credited to Émile Borel, is the so-called "infinite monkey theorem". There exist a lot of different versions of this theorem but usually it is stated as follows.

Theorem (Infinite Monkey Theorem) *A monkey who randomly presses buttons on a typewriter for an infinite amount of time will eventually produce a work of Shakespeare.*

This theorem can be proven using the second BC-lemma. We make the assumption that the buttons are picked independently of each other. Consider the infinite string of characters that the monkey produces. We are interested in finding a substring that contains a work of Shakespeare. This substring has a finite length of, say, k characters. Divide the infinite string into blocks of length k and call E_i the event that the k-th block contains the work of Shakespeare. Now, since the probability of E_i is non-zero, we have that $P(E_i) = p_i$ for some $p_i > 0$. This means that the sum $\sum_{i=1}^{\infty} P(E_i)$ is equal to

$$\sum_{i=1}^{\infty} P(E_i) = \sum_{i=1}^{\infty} p_i = \infty$$

By applying the second BC-lemma we can now conclude that $P(E_i \text{ i.o.}) = 1$ and this is what we wanted to prove. So not only have we proved that the work of Shakespeare will eventually turn up in the infinite string of characters, it will do so an infinite number of times.

Chapter 2

Research on the BC-lemmas

There is a lot of research being done on the BC-lemmas. This is because mathematicians like to be able to extend the scope of events to which the BC-lemmas can be applied to. One thing which is convenient about the lemmas is that it tells us the probability of certain events. If we are given a sequence of events $\{A_n\}_{n=1}^{\infty}$ and we know that the sum of their probabilities converges or diverges (and, perhaps, that these events are independent) then without knowing any of the probabilities of these events, we can immediately say something about $P(A_n \text{ i.o})$. The probability of $\{A_n \text{ i.o}\}$ is either one or zero. For this reason, the BC-lemmas are sometimes called "0-1 laws".

There are a number of ways one could generalize the BC-lemmas. One possible way is to try to relax the condition of mutual independence. Another way is to ask oneself when $P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k)$ is equal to some $\alpha \in [0, 1]$. We will now look at a couple of results and their proofs.

2.1 Relaxing the independence condition in the second BC-lemma

Let's look at one way one might try to strengthen the second BC-lemma. We might be interested in relaxing the condition that in order to have
$P(A_n \text{ i.o.}) = 1$ it must hold that the events in the sequence $\{A_n\}_{n=1}^{\infty}$ are all mutual independent. One article which publishes a result to this end is an article written in 1959 by famous mathematicians Alfréd Rényi (1920 - 1971) and Paul Erdös (1913 - 1996), titled *"On Cantor Series With Convergent $\sum_{n=1}^{\infty} \frac{1}{q_n}$"*. Rényi and Erdös use the mentioned generalized version of the BC-lemma as a main tool in this article. They use it to prove that certain statistical properties of Cantor digits hold.

Here is what the article initially proves:

Lemma C *Let $\{A_n\}_{n=1}^{\infty}$ be a sequence of events and assume that*

$$\sum_{n=1}^{\infty} P(A_n) = \infty$$

and

$$\liminf_{n\to\infty} \frac{\sum_{k=1}^{n}\sum_{l=1}^{n} P(A_k \cap A_l)}{(\sum_{k=1}^{n} P(A_k))^2} = 1$$

It follows that with probability 1 infinitely many among the events A_n occur simultaneously, i.e. $P(A_n$ i.o.$) = 1$

The article then gives a corollary to this lemma, namely that if the events $\{A_n\}_{n=1}^{\infty}$ are *pairwise* independent, we have that $P(A_n$ i.o.$) = 1$.

The proof of lemma C is as follows.

Proof Let I_n be the indicator random variable, that is, $I_n = 1$ if A_n occurs and $I_n = 0$ if A_n does not occur. We now have that $P(A_n) = E(I_n)$ and $P(A_j \cap A_k) = E(I_j I_k)$. Then we define $\eta_n := \sum_{k=1}^{n} I_k$ so that we can write

$$\sum_{j=1}^{n}\sum_{k=1}^{n} P(A_j \cap A_k) = E(\eta_n^2)$$

This is because $\eta_j\eta_k = (I_1 + \ldots\ I_j)(I_1 + \ldots + I_k)$ and all the combinations $I_r I_s$ correspond to a certain $P(A_r \cap A_s)$. This means that

$$\liminf_{n\to\infty} \frac{\sum_{j=1}^{n}\sum_{k=1}^{n} P(A_j \cap A_k)}{(\sum_{k=1}^{n} P(A_k))^2} = \liminf_{n\to\infty} \frac{E(\eta_n^2)}{E^2(\eta_n)}$$

And so, by assumption, we have that

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

If we denote the standard deviation of a random variable X by $\sigma^2(X)$ we can use the fact that $E(\eta_n^2) = \sigma^2(\eta_n) + E^2(\eta_n)$ to write

$$\frac{E(\eta_n^2)}{E^2(\eta_n)} = \frac{\sigma^2(\eta_n) + E^2(\eta_n)}{E^2(\eta_n)} = \frac{\sigma^2(\eta_n)}{E^2(\eta_n)} + 1$$

This means that

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

is the same as saying that

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

At this moment, we use the Chebyshev-inequality, which states that for a random variable X the following inequality holds:

$$P(|X - E(X)| \geq \lambda\sigma(X)) \leq \frac{1}{\lambda^2}, \qquad \lambda > 1$$

Now, choose $\lambda := \frac{\epsilon E(\eta_n)}{\sigma(\eta_n)}$ (the reason for picking this λ will be clear in a moment), then we see from plugging λ into the inequality that

$$P(|\eta_n - E(\eta_n)| \geq \epsilon E(\eta_n)) \leq \frac{\sigma^2(\eta_n)}{\epsilon^2 E^2(\eta_n)}$$

which implies that

$$P(\eta_n \leq (1-\epsilon)E(\eta_n)) \leq \frac{\sigma^2(\eta_n)}{\epsilon^2 E^2(\eta_n)}$$

Now, because we know that $\liminf_{n\to\infty} \frac{\sigma^2(\eta_n)}{E^2(\eta_n)} = 0$, we know that we can certainly find a subsequence $\eta_{n_k}(\eta_{n_1} < \eta_{n_2} < \ldots)$ such that

$$\sum_{k=1}^{\infty} \frac{\sigma^2(\eta_{n_k})}{\epsilon^2 E^2(\eta_{n_k})} < \infty$$

and since $P(\eta_n \leq (1-\epsilon)E(\eta_n)) \leq \dfrac{\sigma^2(\eta_n)}{\epsilon^2 E^2(\eta_n)}$ we have

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

but this means that, by the first BC-lemma, $\eta_{n_k} \geq (1-\epsilon)E(\eta_{n_k})$ with probability 1. Note that our "pick" for λ was such that we could make this deduction. We are now almost done. Because, as by supposition, we had

$$\lim_{k\to\infty} E(\eta_{n_k}) = \infty$$

we can now state that $P(\eta_{n_k} \to \infty) = 1$ and because the η_n are indicator random variables of the events $\{A_1, \ldots, A_n\}$ it can be shown that $P(A_n \text{ i.o.}) = 1$ and this is what we wanted to show.

Now, note that if the events $\{A_n\}$ are pairwise independent and if

$$\sum_{n=1}^{\infty} P(A_n) = \infty$$

then the conditions in lemma C are satisfied because then we have

$$\sum_{k=1}^{n}\sum_{l=1}^{n} P(A_k \cap A_l) = (\sum_{k=1}^{n} P(A_k))^2 + \sum_{k=1}^{n} P(A_k)(1 - P(A_k))$$

and now that we have proved lemma C we can immediately state the following:

Corollary *If the $\{A_n\}$ are pairwise independent and $\sum_{n=1}^{\infty} P(A_n) = \infty$ then $P(A_n \text{ i.o.}) = 1$*

We have now formulated a generalization of the second Borel-Cantelli lemma. The condition of mutual independence has been replaced with with pairwise independence.

One other corollary of lemma C is that if

$$P(A_k \cap A_l) \leq P(A_k)P(A_l)$$

and

$$\liminf_{n\to\infty} \frac{\sum_{k=1}^{n}\sum_{l=1}^{n} P(A_k \cap A_l)}{(\sum_{k=1}^{n} P(A_k))^2} = 1$$

then

$$P(A_n \text{ i.o.}) = 1$$

This result will be expanded upon later on.

2.2 Providing A Lower Bound For $P(\bigcap_{n=1}^{\infty}\bigcup_{k=n}^{\infty} A_k)$

In 1970 an article written by J. Shuster entitled *"On The Borel-Cantelli Problem"* was published. In this article two theorems are presented which provide a generalized version of the first and the second BC-lemma. The first theorem is the main result of the article:

Theorem
(a) If there exists an $A \in \mathcal{F}$ such that

$$\sum_{k=1}^{\infty} P(A \cap A_k) < \infty$$

then $P(\{A_n \text{ i.o.}\}) \leq 1 - P(A)$.
(b)If for every set $A \in \mathcal{F}$ such that $P(A) > 0$ it holds that

$$\sum_{k=1}^{\infty} P(A \cap A_k) = \infty$$

then $P(A_n \text{ i.o.}) = 1$.

How is this theorem a generalization of the original BC-lemmas? First, note that if we pick $A = \Omega$ we have that $P(A_n \cap A) = P(A_n \cap \Omega) = P(A_n)$ and $P(A_n \text{ i.o.}) = 1 - P(\Omega) = 0$ and this is precisely what the original first BC-lemma stated. Secondly, if we compare (b) with the original second BC-lemma we can see that the independence criterion is dropped altogether, so this theorem is definitely an improvement over the original lemmas.

Proof of (a) To see that (a) theorem holds, we can look at the indicator function I_k of the event $A_k \cap A$. Now, define $T = \sum_{k=1}^{\infty} I_k$. Consider the expected value of T. Because $E(T) = \sum_{k=1}^{\infty} P(A_k \cap A)$ and, by the hypothesis of (b), $\sum_{k=1}^{\infty} P(A_k \cap A) < \infty$, we have $E(T) < \infty$. But this means that only a finite number of $A_k \cap A$ occur. This means that the probability $P(A_n \text{ i.o.})$ is *at most* equal to $P(A^c) = 1 - P(A)$, which is what we wanted to prove.

Proof of (b) To see why (b) is true, define the set

$$B_n := \bigcap_{k=n}^{\infty} A_k^c = A_n^c \cap A_{n+1}^c \cap A_{n+2}^c \cap \dots$$

If we look at the event $A_k \cap B_n$ we can see that if $k \geq n$ then $P(A_k \cap B_n) = 0$. This means that

$$\sum_{k=1}^{\infty} P(A_k \cap B_n) = \sum_{k=1}^{n-1} P(A_k \cap B_n)$$

Note that the sum on the right-hand side is finite. This implies that we must have that $P(B_n) = 0$ and this means that

$$P(\bigcup_{n=1}^{\infty} B_n) = P(\bigcup_{n=1}^{\infty} A_k^c) = 0 \Longrightarrow P(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k) = 1$$

and this is what we wanted to show.

At this moment we have been able to drop the independence criterion altogether and we have found a more general set of conditions for the A_k to occur infinitely often or only a finite number of times. A question we can ask ourself is "can we refine the necessary and sufficient conditions for $P(A_n \text{ i.o.}) = \alpha$ where $\alpha \in [0, 1]$"? It turns out that we can.

2.3 $P(A_n$ i.o.$) = \alpha$ And Equivalent Statements

In 1994 the Russian mathematicians V.V. Petrov and A.I. Martikainen wrote an article about the Borel-Cantelli lemma which extended the results found by Shuster. Their intention was to formulate necessary and sufficient conditions for $P(A_n \text{ i.o.}) = \alpha$ with $\alpha \in [0, 1]$ and also to provide other necessary and sufficient conditions stated in different terms. They also were interested in convenient ways to calculate α. Their main result is as follows.

Theorem *Let $0 < \alpha \leq 1$. Then the following statements are equivalent:*

1. $P(A_n \text{ i.o}) \geq \alpha$

2. $\sum_{n=1}^{\infty} P(A_n \cap B) = \infty$ for any $B \in \mathcal{F}$ with $P(B) > 1 - \alpha$

3. for any $B \in \mathcal{F}$ with $P(B) > 1 - \alpha$, the sequence $\{P(A_n \cap B)\}$ contains an infinite number of positive numbers.

We now have to show that (1) ⇒ (2) ⇒ (3) ⇒ (1). First, note that (2) ⇒ (3). If it didn't, it would mean only a finite number of $P(A_n \cap B)$ would be bigger than zero and then the sum $\sum_{n=1}^{\infty} P(A_n \cap B)$ could never go to infinity.

Does (1) ⇒ (2)? Suppose $P(A_n \text{ i.o.}) \geq \alpha$ but assume that the event B is such that $\sum_{n=1}^{\infty} P(A_n \cap B) < \infty$. Then,

$$P(A_n \text{ i.o.}) + P(B) - 1 \leq P(A_n \cap B \text{ i.o.}) = 0$$

according to the first Borel-Cantelli lemma. This implies that $P(B) \leq 1 - \alpha$ and this proves (1) ⇒ (2).

Now for the (3) ⇒ (1) case. Suppose that (3) holds but that we have that $P(A_n \text{ i.o.}) < \alpha$. This would imply that $P(\bigcup_{k=n}^{\infty} A_n) < \alpha$ for some n. If we define $C := \bigcup_{k=n}^{\infty} A_k$ and $B = C^c$ then, obviously, $P(B) > 1-\alpha$ and $B \cap C = B \cap \bigcup_{k=n}^{\infty} A_k = \emptyset$. This means that $P(A_k \cap B) = 0$ for all $k \geq n$. This is a contradiction and thus we have proved (3) ⇒ (1).
We've established that the above-mentioned statements are the same. However, it can be quite tricky to calculate the $P(A_n \cap B)$ so to obtain a value for α is not always easy.

2.4 On Computing A Lower Bound For $P(A_n$ i.o.)

Petrov has (co-)written numerous articles about the Borel-Cantelli lemma. In later articles he tries to formulate more statements equivalent to those mentioned in the theorem above and also to find an explicit expression for α. In his article *"A Note On The Borel-Cantelli Lemma"*, a result mentioned in [ErdRe59] is extended and used to give a lower bound for $P(A_n \text{ i.o.})$. The result of this article is stated as follows.

Theorem *If $\{A_n\}_{n=1}^{\infty}$ is such that*

$$P(A_k \cap A_j) \leq C P(A_k) P(A_j)$$

for $k, j > L$ for some L and $C \geq 1$ then

$$P(A_n \text{ i.o.}) \geq \frac{1}{C}$$

We have seen a similar statement in the previous section, however this result gives a less elaborate way of finding the lower bound. Also, a corollary of this theorem is that if $C = 1$ then $P(A_j \cap A_j) \leq P(A_j)P(A_k)$ implies $P(A_n \text{ i.o.}) = 1$. This is a generalization of the pairwise-independent case of the Borel-Cantelli lemma discussed in section 2. A proof of this theorem makes use of an inequality discovered by Chung and Erdös in 1952. It states that if

$A_1, A_2, \ldots, A_n$ are events then

$$P(\bigcup_{k=1}^{n} A_k) \geq \frac{(\sum_{k=1}^{n} P(A_k))^2}{\sum_{k,j=1}^{n} P(A_k \cap A_j)}$$

To prove the theorem, again let $\{A_n\}_{n=1}^{\infty}$ be a sequence of events with the conditions mentioned in the theorem. First, note that

$$P(A_k \cap A_j) \leq CP(A_k)P(A_j)$$

implies that

$$\sum_{k,j=n}^{N} P(A_k \cap A_j) \leq C(\sum_{k,j=n,k\neq j}^{N} P(A_k)P(A_j) + \sum_{n=1}^{\infty} P(A_n))$$

if $n > L$. Because we've assumed that $C \geq 1$, this implies that

$$\sum_{k,j=n}^{N} P(A_k \cap A_j) \leq C(\sum_{k,j=n}^{N} P(A_k)P(A_j) + \sum_{n=1}^{\infty} P(A_n))$$

for $k \neq j$. Now, note that

$$\sum_{k,j=n,k\neq j}^{N} P(A_k)P(A_j) = (\sum_{k=n}^{N} P(A_k))^2 - \sum_{k=n}^{N} (P(A_k))^2$$

therefore

$$\sum_{k,j=n,k\neq j}^{N} P(A_k)P(A_j) + \sum_{n=1}^{\infty} P(A_n) \leq (\sum_{k=n}^{N} P(A_k))^2 + \sum_{k=n}^{N} P(A_k)$$

Now we use the Chung-Erdös inequality and plug in the inequalities above to obtain

$$P(\bigcup_{k=n}^{N} A_k) \geq \frac{(\sum_{k=n}^{N} P(A_k))^2}{C(\sum_{k=n}^{N} P(A_k)^2 + \sum_{k=n}^{N} P(A_k))} \geq \frac{1}{C}(1 + (\sum_{k=n}^{N} P(A_k)^{-1})^{-1}$$

Now, if we take n fixed and let N go to infinity, we find that

$$1 + (\sum_{k=n}^{N} P(A_k))^{-1} \longrightarrow 1$$

because, by assumption, $\sum_{k=n}^{\infty} P(A_k) = \infty$. So in other words,

$$\liminf_{N\to\infty} P(\bigcup_{k=n}^{N} A_k) \geq \frac{1}{C}$$

and

$$P(\bigcup_{k=n}^{\infty} A_k) \geq \frac{1}{C}$$

We are now almost done. If we define $B_n := \bigcup_{k=n}^{\infty} A_k$ and note that $\bigcap_{n=1}^{\infty} B_n = \limsup_{n\to\infty} A_n$ (because the B_n form a increasing sequence, i.e. $B_1 \supset B_2 \supset \ldots$) we can see that

$$\liminf_{n\to\infty} P(B_n) = P(\bigcap_{n=1}^{\infty} B_n) = P(\bigcap_{k=1}^{\infty} \bigcup_{k=n}^{\infty} A_n) = P(\limsup_{n\to\infty} A_n) \geq \frac{1}{C}$$

and this is what we wanted to prove.

2.5 Conclusion

We have seen a few examples of how the BC-lemmas can be generalized. Ofcourse, this is not the end of the road. There has been and still is a lot of effort being put into the investigation of these lemmas. For some examples of recent results on generalizations of the BC-lemmas, see [BCGen] and [STBC].

References

- "*On The Borel-Cantelli Lemma*" - A.I. Martikainen & V. V. Petrov, Journal of Mathematical Sciences Volume 68, Number 4 - February, 1994 [MartPe94]
- "*A note on the Borel-Cantelli lemma*" - V.V. Petrov, Statistics & Probability Letters Volume 58, Issue 3, 1 July 2002, Pages 283-286 [Pe01]
- "*On the Borel-Cantelli problem*" - J. Shuster, Canadian Mathematical Bulletin 13 (1970), pages 273 - 275[Shu70]
- "*A generalization of the Borel-Cantelli lemma* - V. V. Petrov, Statistics & Probability Letters Volume 67, Issue 3, 15 April 2004, Pages 233-239 [Pe03]
- "*On Cantor's series with convergent* $\sum_{n=1}^{\infty} \frac{1}{q_n}$" - P. Erdös & A. Rényi (`www.renyi.hu/~p_erdos/1959-09.pdf`) [ErdRe59]
- "*On The Borel-Cantelli Lemma And Its Generalization*" - Chunrong Feng, Liangpan Li, Jian Shen (`http://arxiv.org/abs/0910.0067`)[BCGen]
- "*A Simple Proof Of Two Generalized Borel-Cantelli Lemmas*" - Jia-An Yan, Lecture Notes in Mathematics Volume 1874/2006 [STBC]