International Journal of Pure and Applied Mathematics

Volume 20 No. 2 2005, 143-156

CONDITIONAL STRONG LAW OF LARGE NUMBER

Dariusz Majerek[1], Wioletta Nowak[2], Wiesław Zięba[3 §]

[1,2]Department of Mathematics
Technical University
Lublin, POLAND
[1]e-mail: majerek@antenor.pol.lublin.pl
[2]e-mail: wnowak@antenor.pol.lublin.pl
[3]Institute of Mathematics
Maria Curie-Skłodowska University
Pl. Marii Curie-Skłodowskiej 1, Lublin 20-031, POLAND
e-mail: zieba@golem.umcs.lublin.pl

Abstract: The aim of this note is to give a conditional version of Kolmogorov's strong law of large numbers. A strong law of large numbers was generalized in many ways. One of the assumptions, which was weakened, was the independence condition (for example for martingales increments).

In this paper we consider sequences of $\mathcal{F}$-independence of random variables. Note that conditional independence does not imply independence, the opposite implication is also not true, as incorrectly given in the book [3]. In the second part of this paper we prove a conditional version of the Kolmogorov's strong law of large numbers.

AMS Subject Classification: 60F15
Key Words: independence, conditional expectation, law of large numbers

1. Elementary Examples

Let $(\Omega, \mathcal{A}, \mathrm{P})$ be a probability space, and $\mathcal{F}$ a sub-σ-field $\mathcal{A}$. Events $A_1, A_2, \ldots, A_n \in \mathcal{A}$ are independent if

Received: October 25, 2004

§Correspondence author

$$\bigwedge_{1\le k\le n}\ \bigwedge_{1\le i_1<i_2<\cdots<i_k\le n} \mathrm{P}(\bigcap_{s=1}^{k} A_{i_s}) = \prod_{s=1}^{k} \mathrm{P}(A_{i_s}).$$

Definition 1.1. The random events $A_1, A_2, \ldots, A_n \in \mathcal{A}$ are called conditionally independent with respect to event B with $\mathrm{P}(B) > 0$ if

$$\bigwedge_{1\le k\le n}\ \bigwedge_{1\le i_1<i_2<\cdots<i_k\le n} \mathrm{P}(\bigcap_{s=1}^{k} A_{i_s}|B) = \prod_{s=1}^{k} \mathrm{P}(A_{i_s}|B).$$

In [3] was shown that conditional independence does not imply independence of events. Moreover the author claims that opposite implication is true. Consideration about the conditional expectation are more delicate. The following examples illustrate the problem of relation between the two concepts of mutual independence and conditional independence.

Example 1. Let $\Omega = \{1,2,3,4,5,6,7,8\}$, then $p_i = \frac{1}{8}$ for $1 \le i \le 8$, and let $A_1 = \{1,2,3,4\}$ and $A_2 = \{3,4,5,6\}$. Then $\mathrm{P}(A_1) = \mathrm{P}(A_2) = \frac{1}{2}$ and $\mathrm{P}(A_1A_2) = \frac{1}{4}$ so events A_1 and A_2 are independent. Note that if $B = \{2,3,4,5\}$, then $\mathrm{P}(B) = \frac{1}{2}$, $\mathrm{P}(A_1A_2|B) = \frac{1}{2}$, $\mathrm{P}(A_1|B) = \frac{3}{4}$ and $\mathrm{P}(A_2|B) = \frac{3}{4}$, thus $\mathrm{P}(A_1A_2|B) \neq \mathrm{P}(A_1|B)\mathrm{P}(A_2|B)$. Hence the independent events A_1 and A_2 are not conditionally independent with respect to B.

The following example shows that the opposite implication is also not true.

Example 2. Let us consider an experiment relying on tossing a coin twice. We have at our disposal n coins. Let p_o^i be probabilities of tails for i-th coin and p_i be probabilities of choosing i-th coin for $i = 1, 2, \ldots, n$.

Let $A_1 = \{\text{tail in first toss}\}$, $A_2 = \{\text{tail in second toss}\}$, $H_i = \{\text{coin } i$ was selected $, i = 1, 2, \ldots, n\}$. Then

$$\mathrm{P}(A_1A_2|H_i) = p_o^i p_o^i, \quad \mathrm{P}(A_1|H_i) = p_o^i, \quad \mathrm{P}(A_2|H_i) = p_o^i\,,$$

hence $\mathrm{P}(A_1|H_i)\mathrm{P}(A_2|H_i) = \mathrm{P}(A_1A_2|H_i)$. However the equality $\mathrm{P}(A_1A_2) = \mathrm{P}(A_1)\mathrm{P}(A_2)$ is not always true, because

$$\mathrm{P}(A_1A_2) = \sum_{i=1}^{n} \mathrm{P}(A_1A_2|H_i)\mathrm{P}(H_i) = \sum_{i=1}^{n} (p_o^i)^2 p_i,$$

and $\mathrm{P}(A_1) = \sum_{i=1}^{n} \mathrm{P}(A_1|H_i)\mathrm{P}(H_i) = \sum_{i=1}^{n} p_o^i p_i$, $\mathrm{P}(A_2)$ we obtain analogically.

Let events $A_1, A_2, H_1, H_2, \ldots, H_n$ be chosen such way that $p_o^i = \frac{1}{2^i}, p_i = \frac{1}{n}, i = 1, 2, \ldots, n$. Then

$$\sum_{i=1}^{n} (p_o^i)^2 p_i = \sum_{i=1}^{n} \left(\frac{1}{2^i}\right)^2 \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{2^{2i}} = \frac{1}{3n} \left(\frac{2^{2n}-1}{2^{2n}}\right)$$

and

$$\left(\sum_{i=1}^{n} p_o^i p_i\right) \left(\sum_{j=1}^{n} p_o^j p_j\right) = \left(\sum_{i=1}^{n} \frac{1}{2^i}\frac{1}{n}\right)\left(\sum_{j=1}^{n} \frac{1}{2^j}\frac{1}{n}\right)$$
$$= \frac{1}{n^2}\left(\sum_{i=1}^{n} \frac{1}{2^i}\right)\left(\sum_{j=1}^{n} \frac{1}{2^j}\right) = \frac{1}{n^2}(2^n-1)^2 \frac{1}{2^{2n}}.$$

If the equality $\mathrm{P}(A_1A_2) = \mathrm{P}(A_1)\mathrm{P}(A_2)$ is true, then

$$\frac{1}{3n}\left(\frac{2^{2n}-1}{2^{2n}}\right) = \frac{1}{n^2}(2^n-1)^2\frac{1}{2^{2n}}$$

and $n(2^{2n}-1) = 3(2^n-1)^2$, then

$$2^n((n-3)2^n+6) - n - 3 = 0.$$

But for $n \geq 2$ the statement $2^n((n-3)2^n+6) - n - 3$ is always non-negative.

The following example shows existence of independent events which are conditionally independent with respect to a complete system of events.

Example 3. Let $A_1, A_2 \in \mathcal{A}$ be events such that $\mathrm{P}(A_1) = \mathrm{P}(A_2) = \frac{1}{3}$ and $\mathrm{P}(A_1A_2) = \frac{1}{9}$ and let $H_1, H_2, \ldots, H_n$ be events which create a complete system of events. We will show that $\mathrm{P}(A_1A_2|H_i) = \mathrm{P}(A_1|H_i)$ $\mathrm{P}(A_2|H_i)$ for $i = 1, 2, \ldots, n$. Moreover let H_i be the event such that:

$$H_i = H_i^{A_1\setminus A_2} \cup H_i^{A_2\setminus A_1} \cup H_i^{A_1\cap A_2} \cup H_i^{\overline{A_1\cup A_2}} \quad \text{for } i = 1, 2, \ldots, n,$$

where $H_i^{A_1\setminus A_2}$ is obtained by dividing event $A_1\setminus A_2$ into n equal parts in measure and means i-th part, $i = 1, 2, \ldots, n$. In similar way we define other events. Then for $i = 1, 2, \ldots, n$ $\mathrm{P}(H_i^{A_1\setminus A_2}) = \frac{2}{9n}$, $\mathrm{P}(H_i^{A_2\setminus A_1}) = \frac{2}{9n}$, $\mathrm{P}(H_i^{A_1\cap A_2}) = \frac{1}{9n}$, $\mathrm{P}(H_i^{\overline{A_1\cup A_2}}) = \frac{4}{9n}$, hence $\mathrm{P}(H_i) = \frac{1}{n}$ and $\mathrm{P}(A_1A_2|H_i) = \frac{1}{9}$, $\mathrm{P}(A_1|H_i) = \frac{1}{3}$, $\mathrm{P}(A_2|H_i) = \frac{1}{3}$.

However there is a complete system of events $\{H_i, i = 1, 2, \ldots, n\}$ such that $\mathrm{P}(A_1A_2|H_i) \neq \mathrm{P}(A_1|H_i)\mathrm{P}(A_2|H_i)$, $i = 1, 2, \ldots, n$.

Example 4. Let events $A_1, A_2 \in \mathcal{A}$ be chosen such as in above example. Moreover let

$$H_i = H_i^{A_1\setminus A_2} \cup H_i^{A_2\setminus A_1} \cup H_i^{A_1\cap A_2} \cup H_i^{\overline{A_1\cup A_2}}, \quad \text{for } i = 1, 2, \ldots, n$$

and let $H_i^{A_1\setminus A_2}$, $H_i^{A_2\setminus A_1}$ be chosen such as $\mathrm{P}(H_i^{A_1\setminus A_2}) = \frac{1}{2^n}\frac{2}{9}$, $\mathrm{P}(H_i^{A_2\setminus A_1}) = \frac{3}{2^n}\frac{2}{9}$ for $i = 1, 2, \ldots, k$, however $\mathrm{P}(H_i^{A_1\setminus A_2}) = \frac{3}{2^n}\frac{2}{9}$, $\mathrm{P}(H_i^{A_2\setminus A_1}) = \frac{1}{2^n}\frac{2}{9}$ for $i = k+1, 2, \ldots, n = 2k$, $H_i^{A_1\cap A_2}$ and $H_i^{\overline{A_1\cup A_2}}$ are chosen as in previous example. Then $\mathrm{P}(A_1A_2|H_i) = \frac{1}{9}$, $\mathrm{P}(A_1|H_i) = \frac{2}{9}$, $\mathrm{P}(A_2|H_i) = \frac{4}{9}$ for $i = 1, 2, \ldots, k$ and $\mathrm{P}(A_1A_2|H_i) = \frac{1}{9}$, $\mathrm{P}(A_1|H_i) = \frac{4}{9}$, $\mathrm{P}(A_2|H_i) = \frac{2}{9}$ for $i = k+1, 2, \ldots, n = 2k$.

Hence $\mathrm{P}(A_1A_2|H_i) \neq \mathrm{P}(A_1|H_i)\mathrm{P}(A_2|H_i)$ for $i = 1, 2, \ldots, n$.

2. Conditional Independence of Random Variables

Now let $\mathcal{F}$ be a σ-field in $\mathcal{A}$.

Definition 2.1. The random events $A_1, A_2, \ldots, A_n$ are called $\mathcal{F}$-independent if

$$\bigwedge_{1\le k\le n} \ \bigwedge_{1\le i_1<i_2<\cdots<i_k\le n} \mathrm{E}^{\mathcal{F}} \prod_{s=1}^{k} I_{A_{i_s}} = \prod_{s=1}^{k} \mathrm{E}^{\mathcal{F}} I_{A_{i_s}}.$$

If $\mathcal{F} = (\emptyset, \Omega)$ we obtain the definition of independence of random events. Note that, if $\mathcal{F} = \mathcal{A}$ then all random events are $\mathcal{A}$-independent.

A sequence of families $\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_n$, where $\mathcal{G}_k \subset \mathcal{A}$ is $\mathcal{F}$-independent, if for any sequence $A_1, A_2, \ldots, A_n$ such that $A_i \in \mathcal{G}_i$, $i = 1, 2, \ldots, n$ is the sequence of $\mathcal{F}$-independent random events.

If $X : \Omega \to \mathbb{R}$ is a random variable then by $\mathcal{A}_X$ we define the smallest σ-field with respect to the random variable X is measurable. Obviously if $\mathcal{A}_X = X^{-1}(\mathcal{B})$, where $\mathcal{B}$ is Borel σ-field on real. We will say, that $\mathcal{A}_X$ is σ-field generated by random variable X.

Random variables $X_1, X_2, \ldots, X_n$ are $\mathcal{F}$-independent if σ-fields $\mathcal{A}_{X_1}$, $\mathcal{A}_{X_2}, \ldots, \mathcal{A}_{X_n}$ are $\mathcal{F}$-independent.

Example 5. Let Ω, A_1, A_2, B be defined such as in example 1 and $\mathcal{F} = \sigma(I_B)$, where

$$I_B(\omega) = \begin{cases} 1 & \text{for } \omega \in B\,, \\ 0 & \text{for } \omega \notin B, \end{cases}$$

then

$$\mathrm{E}^{\mathcal{F}} I_{A_1A_2}(\omega) = \begin{cases} \mathrm{P}(A_1A_2|B) & \text{for } \omega \in B\,, \\ \mathrm{P}(A_1A_2|\overline{B}) & \text{for } \omega \notin B\,, \end{cases}$$

$$\mathrm{E}^{\mathcal{F}} I_{A_i}(\omega) = \begin{cases} \mathrm{P}(A_i|B) & \text{for } \omega \in B\,, \\ \mathrm{P}(A_i|\overline{B}) & \text{for } \omega \notin B\,, \end{cases}$$

for $i = 1, 2$. By Example 1 for $\omega \in B$

$$\mathrm{E}^{\mathcal{F}} I_{A_1A_2} = \mathrm{P}(A_1A_2|B) \neq \mathrm{P}(A_1|B)\mathrm{P}(A_2|B) = \mathrm{E}^{\mathcal{F}} I_{A_1} \mathrm{E}^{\mathcal{F}} I_{A_2}\,,$$

hence independent random events A_1 and A_2 are not $\mathcal{F}$-independent.

The opposite implication is also not true too.

Example 6. (see [3]) Let us consider an experiment relying on tossing a coin twice. We have at our disposal two coins a and b. Let p_a and p_b be probabilities of heads for coin a and b respectively, and $p_a \neq p_b$. We select a coin at random and toss it twice. Let $A_1 = \{\text{head in the first toss}\}$, $A_2 = \{\text{head in the second toss}\}$, $B = \{\text{coin } a \text{ is selected}\}$, $\overline{B} = \{\text{coin } b \text{ is selected}\}$.

Then

$$\mathrm{P}(A_1A_2|B) = p_a p_a, \quad \mathrm{P}(A_1|B) = p_a, \quad \mathrm{P}(A_2|B) = p_a$$

and

$$\mathrm{P}(A_1A_2|B) = \mathrm{P}(A_1|B)\mathrm{P}(A_2|B) \quad \text{and} \quad \mathrm{P}(A_1A_2|\overline{B}) = p_b p_b,$$
$$\mathrm{P}(A_1|\overline{B}) = p_b, \quad \mathrm{P}(A_2|\overline{B}) = p_b\,,$$

hence

$$\mathrm{P}(A_1A_2|\overline{B}) = \mathrm{P}(A_1|\overline{B})\mathrm{P}(A_2|\overline{B}).$$

Let $\mathcal{F} = \sigma(B)$, then

$$\begin{aligned}\mathrm{P}(A_1A_2|\mathcal{F}) &= I_B \mathrm{P}(A_1A_2|B) + I_{\overline{B}} \mathrm{P}(A_1A_2|\overline{B}) \\ &= I_B \mathrm{P}(A_1|B)\mathrm{P}(A_2|B) + I_{\overline{B}} \mathrm{P}(A_1|\overline{B})\mathrm{P}(A_2|\overline{B}),\end{aligned}$$

$$\begin{aligned}&\mathrm{P}(A_1|\mathcal{F})\mathrm{P}(A_2|\mathcal{F}) \\ &\quad = [I_B \mathrm{P}(A_1|B) + I_{\overline{B}} \mathrm{P}(A_1|\overline{B})][I_B \mathrm{P}(A_2|B) + I_{\overline{B}} \mathrm{P}(A_2|\overline{B})] \\ &\quad = I_B I_B \mathrm{P}(A_1|B)\mathrm{P}(A_2|B) + I_{\overline{B}} I_B \mathrm{P}(A_1|\overline{B})\mathrm{P}(A_2|B) \\ &\quad + I_B I_{\overline{B}} \mathrm{P}(A_1|B)\mathrm{P}(A_2|\overline{B}) + I_{\overline{B}} I_{\overline{B}} \mathrm{P}(A_1|\overline{B})\mathrm{P}(A_2|\overline{B})\end{aligned}$$

$$= I_B \mathrm{P}(A_1|B)\mathrm{P}(A_2|B) + I_{\overline{B}} \mathrm{P}(A_1|\overline{B})\mathrm{P}(A_2|\overline{B}).$$

Hence we obtain that $\mathrm{P}(A_1A_2|\mathcal{F}) = \mathrm{P}(A_1|\mathcal{F})\mathrm{P}(A_2|\mathcal{F})$ therefore events A_1 and A_2 are $\mathcal{F}$-independent. However $\mathrm{P}(A_1A_2) = \frac{1}{2}p_a^2 + \frac{1}{2}p_b^2$, $\mathrm{P}(A_1) = \frac{1}{2}(p_a + p_b)$, $\mathrm{P}(A_2) = \frac{1}{2}(p_a + p_b)$ and because $p_a \neq p_b$ then equality $\mathrm{P}(A_1A_2) = \mathrm{P}(A_1)\mathrm{P}(A_2)$ is not true, hence events A_1 and A_2 are not independent.

Moreover we note, that if random variable X and Y are $\mathcal{F}$-independent then equality $\mathrm{E}XY = \mathrm{E}X\mathrm{E}Y$ does not always hold.

Example 7. Let events A_1, A_2 and B be defined such as in previous example, moreover let $X = I_{A_1}$, $Y = I_{A_2}$ then

$$\sigma(X) = \sigma(I_{A_1}) = \sigma(A_1) \quad \text{and} \quad \sigma(Y) = \sigma(I_{A_2}) = \sigma(A_2).$$

Now we will show that σ-fields $\sigma(A_1)$ and $\sigma(A_2)$ are $\mathcal{F}$-independent.
By Example 6 we have, that

$$\mathrm{P}(A_1A_2|\mathcal{F}) = \mathrm{P}(A_1|\mathcal{F})\mathrm{P}(A_2|\mathcal{F}).$$

We note that

$$\mathrm{P}(A_1\overline{A}_2|B) = p_a(1-p_a), \quad \mathrm{P}(A_1|B) = p_a, \quad \mathrm{P}(\overline{A}_2|B) = 1 - p_a\,,$$

hence

$$\mathrm{P}(A_1\overline{A}_2|B) = \mathrm{P}(A_1|B)\mathrm{P}(\overline{A}_2|B).$$

Similarly we can show that

$$\mathrm{P}(A_1\overline{A}_2|\overline{B}) = \mathrm{P}(A_1|\overline{B})\mathrm{P}(\overline{A}_2|\overline{B}).$$

Thus

$$\begin{aligned}\mathrm{P}(A_1\overline{A}_2|\mathcal{F}) &= I_B\mathrm{P}(A_1\overline{A}_2|B) + I_{\overline{B}}\mathrm{P}(A_1\overline{A}_2|\overline{B}) \\ &= I_B\mathrm{P}(A_1|B)\mathrm{P}(\overline{A}_2|B) + I_{\overline{B}}\mathrm{P}(A_1|\overline{B})\mathrm{P}(\overline{A}_2|\overline{B})\,,\end{aligned}$$

$$\begin{aligned}&\mathrm{P}(A_1|\mathcal{F})\mathrm{P}(\overline{A}_2|\mathcal{F}) \\ &\quad= [I_B\mathrm{P}(A_1|B) + I_{\overline{B}}\mathrm{P}(A_1|\overline{B})][I_B\mathrm{P}(\overline{A}_2|B) + I_{\overline{B}}\mathrm{P}(\overline{A}_2|\overline{B})] \\ &\quad= I_B\mathrm{P}(A_1|B)\mathrm{P}(\overline{A}_2|B) + I_{\overline{B}}\mathrm{P}(A_1|\overline{B})\mathrm{P}(\overline{A}_2|\overline{B})\,,\end{aligned}$$

hence

$$\mathrm{P}(A_1\overline{A}_2|\mathcal{F}) = \mathrm{P}(A_1|\mathcal{F})\mathrm{P}(\overline{A}_2|\mathcal{F}) \text{ a.s.}$$

In similar way we obtain $\mathcal{F}$-independence of random events $\overline{A}_1$, A_2 and $\overline{A}_1$, $\overline{A}_2$. Therefore we have $\mathcal{F}$-independence of random variables X and Y, but

$$\mathrm{E}XY = \mathrm{E}I_{A_1}I_{A_2} = \mathrm{P}(A_1A_2) \neq \mathrm{P}(A_1)\mathrm{P}(A_2) = \mathrm{E}I_{A_1}\mathrm{E}I_{A_2} = \mathrm{E}X\mathrm{E}Y$$

by the previous example.

The independence of random variables X and Y does not imply the following equality $\mathrm{E}^{\mathcal{F}}XY = \mathrm{E}^{\mathcal{F}}X\mathrm{E}^{\mathcal{F}}Y$.

Example 8. Let events A_1, A_2 and B be defined such as in Example 1. Moreover, let $X = I_{A_1}$, $Y = I_{A_2}$. From independence of random events A_1 and A_2 we obtain the independence of random variable X and Y. Let $\mathcal{F} = \sigma(B)$, then

$$\begin{aligned}\mathrm{E}^{\mathcal{F}}XY = \mathrm{E}^{\mathcal{F}}I_{A_1}I_{A_2} = \mathrm{P}(A_1A_2|\mathcal{F}) \neq \mathrm{P}(A_1|\mathcal{F})\mathrm{P}(A_2|\mathcal{F})\\ = \mathrm{E}^{\mathcal{F}}I_{A_1}\mathrm{E}^{\mathcal{F}}I_{A_2} = \mathrm{E}^{\mathcal{F}}X\mathrm{E}^{\mathcal{F}}Y\end{aligned}$$

by Example 5.

By above examples we know that there is a possible situation that every observer state independence of random variables which are not independent and vice versa.

3. Conditional Version of Kolmogorov's Strong Law of Large Numbers

Let $(\Omega, \mathcal{A}, \mathrm{P})$ be a probability space and $\mathcal{F}$ nonempty sub-σ-field $\mathcal{A}$. The first lemma of Borel-Cantelli states that if $\{A_n, n \geq 1\}$ is an arbitrary sequence $\sum_{n=1}^{\infty} \mathrm{P}(A_n) < \infty$, then with probability one only finite number of events of the sequence $\{A_n, n \geq 1\}$ holds. The second lemma of Borel-Cantelli shows that if the sequence of independent random events such that $\sum_{n=1}^{\infty} \mathrm{P}(A_n) = \infty$, then probability that infinity many events of the sequence $\{A_n, n \geq 1\}$ holds equals one.

Theorem 3.1. *Let $\{A_n, n \geq 1\}$ be a sequence of random events such that $\sum_{n=1}^{\infty} \mathrm{P}(A_n) < \infty$, then $\sum_{n=1}^{\infty} \mathrm{P}(A_n|\mathcal{F}) < \infty$* a.s.

Proof. Obviously

$$\sum_{n=1}^{\infty} \mathrm{E}I_{A_n} = \sum_{n=1}^{\infty} \mathrm{E}(\mathrm{E}^{\mathcal{F}}I_{A_n}) = \lim_{n\to\infty} \sum_{k=1}^{n} \mathrm{E}(\mathrm{E}^{\mathcal{F}}I_{A_k}) = \lim_{n\to\infty} \mathrm{E}(\sum_{k=1}^{n} \mathrm{E}^{\mathcal{F}}I_{A_k}).$$

By the Lebesgue's Monotone Convergence Theorem we have that

$$\lim_{n\to\infty} \mathrm{E}(\sum_{k=1}^{n} \mathrm{E}^{\mathcal{F}} I_{A_k}) = \mathrm{E}(\lim_{n\to\infty} \sum_{k=1}^{n} \mathrm{E}^{\mathcal{F}} I_{A_k}) = \mathrm{E}(\sum_{n=1}^{\infty} \mathrm{E}^{\mathcal{F}} I_{A_n}) = \infty$$

if

$$\mathrm{P}\left(\sum_{n=1}^{\infty} \mathrm{E}^{\mathcal{F}} I_{A_n} = \infty\right) > 0. \qquad \square$$

The opposite implication is not true by Example 1 in [2].

Lemma 3.2. (Conditional Lemma of Borel-Cantelli I) *Let* $\{A_n, n \geq 1\}$ *be a sequence of random events such that* $A = \{\omega : \sum_{n=1}^{\infty} \mathrm{P}(A_n|\mathcal{F}) < \infty\}$, $\mathrm{P}(A) < 1$, *then only finitely many events from the sequence* $\{A_n \cap A, n \geq 1\}$ *hold with probability one.*

Proof. Let

$$U = \bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k \cap A,$$

then

$$\mathrm{P}(U|\mathcal{F}) = \mathrm{P}\left(\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} [A_k \cap A]|\mathcal{F}\right) = \lim_{n\to\infty} \mathrm{P}\left(\bigcup_{k=n}^{\infty} [A_k \cap A]|\mathcal{F}\right)$$

$$= \lim_{n\to\infty} \mathrm{E}^{\mathcal{F}} I_{\bigcup_{k=n}^{\infty} [A_k \cap A]} \leq \lim_{n\to\infty} \left(\sum_{k=n}^{\infty} \mathrm{E}^{\mathcal{F}} I_{[A_k \cap A]}\right) = 0 \quad \text{a.s.}$$

Hence $\mathrm{P}(U) = \mathrm{E}[\mathrm{P}(U|\mathcal{F})] = 0$ a.s. $\square$

Lemma 3.3. (Conditional Lemma of Borel-Cantelli II) *Let* $\{A_n, n \geq 1\}$ *be a sequence of* $\mathcal{F}$*- independent events and let* $A = \{\omega : \sum_{n=1}^{\infty} \mathrm{P}(A_n|\mathcal{F}) = \infty\}$. *Then* $\mathrm{P}(\limsup A_n) = \mathrm{P}(A)$.

Proof. Let $E = (\bigcap_{n=1}^{\infty} \bigcup_{k=n}^{\infty} A_k)^c = \bigcup_{n=1}^{\infty} \bigcap_{k=n}^{\infty} A_k^c$. Properties of conditional expectation imply

$$\mathrm{P}(E|\mathcal{F}) = \lim_{n\to\infty} \mathrm{P}(\bigcap_{k=n}^{\infty} A_k^c|\mathcal{F}) = \lim_{n\to\infty} (\lim_{k\to\infty} \mathrm{P}(\bigcap_{i=n}^{k} A_i^c|\mathcal{F}))$$

$$= \lim_{n\to\infty} (\lim_{k\to\infty} \prod_{i=n}^{k} \mathrm{P}(A_i^c|\mathcal{F})) = \lim_{n\to\infty} \lim_{k\to\infty} [\prod_{i=n}^{k} (1 - \mathrm{P}(A_i|\mathcal{F}))]$$

$$= \lim_{n\to\infty} \prod_{i=n}^{\infty} (1 - \mathrm{P}(A_i|\mathcal{F})) \leq \lim_{n\to\infty} \exp(-\sum_{i=n}^{\infty} \mathrm{P}(A_i|\mathcal{F})) \quad \text{a.s.}$$

Thus for almost every $\omega \in A$ we have

$$0 \leq \mathrm{P}(E|\mathcal{F})(\omega) \leq \lim_{n\to\infty} \exp(-\sum_{i=n}^{\infty} \mathrm{P}(A_i|\mathcal{F})(\omega)) = 0 \quad \text{a.s.}$$

Thus

$$\mathrm{P}(E) = \int_{\Omega} \mathrm{P}(E|\mathcal{F})d\mathrm{P} = \int_{A} \mathrm{P}(E|\mathcal{F})d\mathrm{P} + \int_{A^c} \mathrm{P}(E|\mathcal{F})d\mathrm{P} \leq \mathrm{P}(A^c),$$

so $\mathrm{P}(E^c) \geq \mathrm{P}(A)$.

On the other hand, following the reasoning given in Lemma 3.2, we state that on the set A^c only finitely many events from the sequence $\{A_n, n \geq 1\}$ hold, so $\mathrm{P}(E^c) \leq \mathrm{P}(A)$. □

Theorem 3.4. (Kolmogorov's Inequality) *If $\{X_n, n \geq 1\}$ is a sequence of $\mathcal{F}$-independent random variables belonging to $L^2_{\mathcal{F}}$ then for an arbitrary $\mathcal{F}$-measurable random variable $\varepsilon > 0$ a.s. we have*

$$\varepsilon^2 \mathrm{P}[\max_{1\leq k\leq n} |S_k - \mathrm{E}^{\mathcal{F}} S_k| \geq \varepsilon \ |\mathcal{F}] \leq \sum_{k=1}^{n} \sigma^2_{\mathcal{F}} X_k \quad \text{a.s.},$$

where $S_n = X_1 + X_2 + \cdots + X_n$.

Proof. Let $\varepsilon > 0$ a.s. be an arbitrary $\mathcal{F}$-measurable random variable, let $A = \{\max_{1\leq k\leq n} |S_k - \mathrm{E}^{\mathcal{F}} S_k| \geq \varepsilon\}$, where $S_n = X_1 + X_2 + \cdots + X_k$. This event we can divide into parts, according to which k the inequality $|S_k - \mathrm{E}^{\mathcal{F}} S_k| \geq \varepsilon$ a.s. holds the first time. Let $\tau = \inf\{k : |S_k - \mathrm{E}^{\mathcal{F}} S_k| \geq \varepsilon, 1 \leq k \leq n\}$. On the set A random variable τ takes values $1, 2, \ldots n$. We denote $R_k = X_{k+1} + \ldots X_n$ and $A_k = A \cap \{\tau = k\}$.

Then

$$\begin{aligned} \mathrm{E}^{\mathcal{F}}[(S_n - \mathrm{E}^{\mathcal{F}} S_n)^2 I_{A_k}] &= \mathrm{E}^{\mathcal{F}}[((S_k - \mathrm{E}^{\mathcal{F}} S_k) + (R_k - \mathrm{E}^{\mathcal{F}} R_k))^2 I_{A_k}] \\ &= \mathrm{E}^{\mathcal{F}}[(S_k - \mathrm{E}^{\mathcal{F}} S_k)^2 I_{A_k}] + \mathrm{E}^{\mathcal{F}}[(R_k - \mathrm{E}^{\mathcal{F}} R_k)^2 I_{A_k}] \\ &+ 2\mathrm{E}^{\mathcal{F}}[(S_k - \mathrm{E}^{\mathcal{F}} S_k)(R_k - \mathrm{E}^{\mathcal{F}} R_k) I_{A_k}] \geq \mathrm{E}^{\mathcal{F}}[(S_k - \mathrm{E}^{\mathcal{F}} S_k)^2 I_{A_k}] \text{ a.s.} \end{aligned} \quad (1)$$

because

$$\mathrm{E}^{\mathcal{F}}[(R_k - \mathrm{E}^{\mathcal{F}} R_k)^2 I_{A_k}] \geq 0 \text{ a.s.}$$

and

$$\mathrm{E}^{\mathcal{F}}[S_k R_k I_{A_k} - R_k \mathrm{E}^{\mathcal{F}} S_k I_{A_k} - S_k \mathrm{E}^{\mathcal{F}} R_k I_{A_k} + \mathrm{E}^{\mathcal{F}} S_k \mathrm{E}^{\mathcal{F}} R_k I_{A_k}] = 0 \text{ a.s.}$$

because random variable $S_k I_{A_k}$ and R_k are $\mathcal{F}$-independent.

By (1) we have

$$\mathrm{E}^{\mathcal{F}}[(S_n - \mathrm{E}^{\mathcal{F}} S_n)^2 I_{A_k}] \geq \varepsilon^2 \mathrm{E}^{\mathcal{F}} I_{A_k} = \varepsilon^2 \ \mathrm{P}[A_k|\mathcal{F}] \text{ a.s.} \tag{2}$$

If we add both side of the inequality (2) for $k = 1, 2, \dots n.$ we obtain

$$\begin{aligned}\mathrm{E}^{\mathcal{F}}(S_n - \mathrm{E}^{\mathcal{F}} S_n)^2 &\geq \mathrm{E}^{\mathcal{F}}[(S_n - \mathrm{E}^{\mathcal{F}} S_n)^2 I_{A_k}] = \mathrm{E}^{\mathcal{F}}[(S_n - \mathrm{E}^{\mathcal{F}} S_n)^2 \sum_{k=1}^{n} I_{A_k}] \\ &= \sum_{k=1}^{n} \mathrm{E}^{\mathcal{F}}(S_n - \mathrm{E}^{\mathcal{F}} S_n)^2 I_{A_k} \geq \sum_{k=1}^{n} \varepsilon^2 \mathrm{P}(A_k|\mathcal{F}) = \varepsilon^2 \sum_{k=1}^{n} \mathrm{P}(A_k|\mathcal{F}) \\ &= \varepsilon^2 \sum_{k=1}^{n} \mathrm{E}^{\mathcal{F}} I_{A_k} = \varepsilon^2 \sum_{k=1}^{n} \mathrm{E}^{\mathcal{F}} I_{A \cap \{\tau = n\}} = \varepsilon^2 \sum_{k=1}^{n} \mathrm{E}^{\mathcal{F}} I_A I_{\{\tau = k\}} \\ &= \varepsilon^2 \mathrm{E}^{\mathcal{F}} I_A = \varepsilon^2 \mathrm{P}(A|\mathcal{F}) = \varepsilon^2 \mathrm{P}[\max_{1 \leq k \leq n} |S_k - \mathrm{E}^{\mathcal{F}} S_k| \geq \varepsilon \ |\mathcal{F}] \text{ a.s.} \quad \square\end{aligned}$$

Theorem 3.5. (Conditional Kolmogorov's Strong Law of Large Numbers) *If $\{X_n, n \geq 1\}$ is a sequence of $\mathcal{F}$-independent random variables such that* $\sum_{k=1}^{\infty} \frac{\sigma^2_{\mathcal{F}} X_k}{k^2} < \infty$ *a.s. then* $\frac{S_n - \mathrm{E}^{\mathcal{F}} S_n}{n} \longrightarrow 0$ *a.s. as* $n \to \infty$.

Proof. Let $S_n = X_1 + X_2 + \cdots + X_n$, $\sigma^2_{\mathcal{F}} X_n$ be a conditional variance of X_n for $n \geq 1$. For an arbitrary $\mathcal{F}$-measurable random variable $\varepsilon > 0$ a.s. let $B_k = \{\omega : |S_n - \mathrm{E}^{\mathcal{F}} S_n| > n\varepsilon, n \in [2^k, 2^{k+1})\}$. From the character of the set B_k we have $|S_n - \mathrm{E}^{\mathcal{F}} S_n| > \varepsilon 2^k$ for some $n \leq 2^{k+1}$, therefore by conditional Kolmogorov's inequality we obtain

$$\mathrm{P}(B_k|\mathcal{F}) \leq \frac{1}{\varepsilon^2 2^{2k}} \sum_{n=1}^{2^{k+1}} \sigma^2_{\mathcal{F}} X_n \quad \text{a.s.}$$

Thus for $k = 0, 1, \dots$ we have

$$\sum_{k=0}^{\infty} \mathrm{P}(B_k|\mathcal{F}) \leq \frac{1}{\varepsilon^2} \sum_{k=0}^{\infty} \frac{1}{2^{2k}} \sum_{n=1}^{2^{k+1}} \sigma^2_{\mathcal{F}} X_n = \frac{1}{\varepsilon^2} \sum_{n=1}^{\infty} \sigma^2_{\mathcal{F}} X_n \sum_{2^{k+1} \geq n} \frac{1}{2^{2k}} \quad \text{a.s.}$$

If k_n is the smallest negative integer such as $2^{k+1} \geq n$, then

$$\sum_{2^{k+1} \geq n} \frac{1}{2^{2k}} = \frac{1}{2^{2k_n}} \left(\frac{1}{1 - \frac{1}{4}} \right) = \frac{1}{(2^{k_n+1})^2} 2^2 \frac{4}{3} \leq \frac{16}{3} \frac{1}{n^2}.$$

Thus $\sum_{k=0}^{\infty} \mathrm{P}(B_k|\mathcal{F}) \leq \frac{16}{3}\frac{1}{\varepsilon^2}\sum_{n=1}^{\infty}\frac{\sigma^2_{\mathcal{F}}X_n}{n^2} < \infty$ a.s. by assumption. The conclusion now follows immediately from conditional lemma of Borel-Cantelli, because then we obtain that for an arbitrary $\mathcal{F}$-measurable random variable $\varepsilon > 0$ a.s. probability of event that $|S_n - \mathrm{E}^{\mathcal{F}} S_n| > n\varepsilon$ holds infinitely often is equal to zero. □

Example 9. Let $(\Omega, \mathcal{A}, \mathrm{P})$ be probability space, where $\Omega = [0,1]$, and P is Lebesgue measure,

$$X = \begin{cases} \frac{1}{\omega} & \text{for } \omega \in (0,1]\,, \\ 0 & \text{for } \omega = 0. \end{cases}$$

Then

$$\mathrm{E}X = \int_\Omega X d\mathrm{P} = \int_0^1 \frac{1}{\omega} d\omega = \ln \omega|_0^1 = \infty\,,$$

therefore mathematical expectation does not exist. Now let $\mathcal{F} = \sigma\left((\frac{1}{n+1}, \frac{1}{n}]\right)$ for $n = 0, 1, \ldots$, then $\mathrm{E}^{\mathcal{F}} X = n(n+1)\ln\frac{n+1}{n}$, for $\omega \in [\frac{1}{n+1}, \frac{1}{n}]$, $n = 1, 2, \ldots$, hence $\mathrm{E}^{\mathcal{F}} X < \infty$ a.s.

Similarly $\mathrm{E}^{\mathcal{F}} X^2 = n(n+1)$, for $\omega \in [\frac{1}{n+1}, \frac{1}{n}]$, $n = 1, 2, \ldots$, so $\mathrm{E}^{\mathcal{F}} X^2 < \infty$ a.s.

Now we will give the generalization of conditional Kolmogorov's law for such sequences $\{YX_n, n \geq 1\}$.

Let notice that if

$$\begin{matrix} X_{11} & X_{12} & \ldots \\ X_{21} & X_{22} & \ldots \\ \ldots & \ldots & \ldots \end{matrix}$$

is a matrix of $\mathcal{F}$-independent random variables and $\mathcal{F}_i = \sigma(X_{i1}, X_{i2}, \ldots)$, $i = 1, 2, \ldots$ then σ-fields $\mathcal{F}_1, \mathcal{F}_2, \ldots$ are $\mathcal{F}$-independent.

Let $Y_i = \phi_i(Z_i)$, $i = 1, 2, \ldots, n$, where ϕ_i - Borel functions, $Z_i = (X_{i1}, X_{i2}, \ldots, X_{ik_i})$, $i = 1, 2, \ldots, n$.

Then for an arbitrary Borel set B

$$\{Y_i \in B\} = \{Z_i \in \phi_i^{-1}(B)\} \in \sigma(Z_i),$$

thus $Y_1, Y_2, \ldots, Y_n$ are $\mathcal{F}$-independent.

Let a sequence of random variables $\{X_n, n \geq 1\}$ fulfills assumptions of Theorem 3.5, let Y be a random variable such that $\mathcal{F} = \sigma(Y)$.

By the above and from Theorem 6 in [2] sequence $\{YX_n, n \geq 1\}$ is a sequence of $\mathcal{F}$-independent random variables, moreover

$$\sigma^2_{\mathcal{F}}(YX_k) = \mathrm{E}^{\mathcal{F}}(YX_k)^2 - [\mathrm{E}^{\mathcal{F}}(YX_k)]^2 = \mathrm{E}^{\mathcal{F}}(Y^2X_k^2) - (\mathrm{E}^{\mathcal{F}}Y\mathrm{E}^{\mathcal{F}}X_k)^2$$
$$= \mathrm{E}^{\mathcal{F}}Y^2\mathrm{E}^{\mathcal{F}}X_k^2 - (\mathrm{E}^{\mathcal{F}}Y)^2(\mathrm{E}^{\mathcal{F}}X_k)^2 = Y^2\mathrm{E}^{\mathcal{F}}X_k^2 - Y^2(\mathrm{E}^{\mathcal{F}}X_k)^2 = Y^2\sigma^2_{\mathcal{F}}X_k.$$

Therefore the sequence of random variables $\{YX_n, n \geq 1\}$ fulfills assumptions of Theorem 3.5.

4. Conditional Version of Kolmogorov's Strong Law of Large Numbers for Identically Conditional Distribution

Let $(\Omega, \mathcal{A}, \mathrm{P})$ be a probability space and $\mathcal{F}$ nonempty sub-σ-field $\mathcal{A}$.

Definition 4.1. Random variables X and Y have the same conditional distribution if

$$\bigwedge_{a\in\mathbb{R}} \mathrm{E}^{\mathcal{F}}I_{[X\leq a]} = \mathrm{E}^{\mathcal{F}}I_{[Y\leq a]} \quad \text{a.s.}$$

Theorem 4.2. (Conditional Version of Kolmogorov's Strong Law of Large Numbers for Identically Conditional Distribution) *If $\{X_n, n \geq 1\}$ is a sequence of $\mathcal{F}$-independent random variables with the same conditional distribution. Then*

$$\lim_{n\to\infty} \frac{S_n}{n} = Y \quad \text{a.s.}$$

if and only if $\mathrm{E}^{\mathcal{F}}X = Y$ a.s.

Proof. Let $Y_n = X_nI_{[|X_n|\leq n]}$, then by the use that X_n have the same conditional distribution, we have

$$\sum_{n=1}^{\infty} \mathrm{P}\left(X_n \neq Y_n|\mathcal{F}\right) = \sum_{n=1}^{\infty} \mathrm{P}\left(|X_n| > n|\mathcal{F}\right)$$
$$= \sum_{n=1}^{\infty} \mathrm{P}\left(|X| > n|\mathcal{F}\right) = \sum_{n=1}^{\infty} \mathrm{E}^{\mathcal{F}}I_{[|X|>n]} \leq \mathrm{E}^{\mathcal{F}}|X| < \infty,$$
$$\text{because} \quad \sum_{n=1}^{\infty} \mathrm{P}\left(X \geq n|\mathcal{F}\right) \leq \mathrm{E}^{\mathcal{F}}X \leq 1 + \sum_{n=1}^{\infty} \mathrm{P}\left(X \geq n|\mathcal{F}\right).$$

By Lemma 3.2, we conclude that

$$\lim_{n\to\infty} \frac{S_n}{n} = \mathrm{E}^{\mathcal{F}}X \quad \text{a.s. iff} \quad \lim_{n\to\infty} \frac{Y_1 + Y_2 + \ldots Y_N}{n} = \mathrm{E}^{\mathcal{F}}X \text{ a.s.}$$

Since

$$\lim_{n\to\infty} \mathrm{E}^{\mathcal{F}}Y_n = \lim_{n\to\infty} \mathrm{E}^{\mathcal{F}}X_nI_{[|X_n|\leq n]} = \lim_{n\to\infty} \mathrm{E}^{\mathcal{F}}XI_{[|X|\leq n]} = \mathrm{E}^{\mathcal{F}}X,$$

then

$$\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathrm{E}^{\mathcal{F}} Y_k = \mathrm{E}^{\mathcal{F}} X,$$

therefore we had to show that $\lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \left(Y_k - \mathrm{E}^{\mathcal{F}} Y_k\right) = 0$ a.s., but by Theorem 3.5, we only need to prove that $\sum_{n=1}^{\infty} \frac{\sigma^2_{\mathcal{F}} Y_n}{n^2} < \infty$.

Thus,

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{\sigma^2_{\mathcal{F}} Y_n}{n^2} &\leq \sum_{n=1}^{\infty} \frac{1}{n^2}\mathrm{E}^{\mathcal{F}}\left[X_n I_{[|X_n|\leq n]}\right]^2 = \sum_{n=1}^{\infty} \frac{1}{n^2}\mathrm{E}^{\mathcal{F}}\left[X^2 I_{[|X|\leq n]}\right] \\
&= \sum_{n=1}^{\infty} \frac{1}{n^2}\sum_{k=1}^{n} \mathrm{E}^{\mathcal{F}}\left[X^2 I_{[k-1<|X|\leq k]}\right] = \sum_{k=1}^{\infty} \mathrm{E}^{\mathcal{F}}\left[X^2 I_{[k-1<|X|\leq k]}\right]\sum_{n=k}^{\infty}\frac{1}{n^2} \\
&\leq 2\sum_{k=1}^{\infty} \frac{1}{k}\mathrm{E}^{\mathcal{F}}\left[X^2 I_{[k-1<|X|\leq k]}\right] \leq 2\sum_{k=1}^{\infty} \mathrm{E}^{\mathcal{F}}\left[|X| I_{[k-1<|X|<k]}\right] = 2\mathrm{E}^{\mathcal{F}}|X| \\
&< \infty.
\end{aligned}$$

To prove the opposite implication we will see that

$$\frac{X_n}{n} = \frac{S_n}{n} - \frac{n-1}{n}\cdot\frac{S_{n-1}}{n-1} \to 0 \quad \text{a.s.},$$

so the event $\left\{\left|\frac{X_n}{n}\right| > 1\right\}$ hold infinity many times with probability zero. By Lemma 3.2 we have

$$\sum_{n=1}^{\infty} \mathrm{P}^{\mathcal{F}}\left(\left|\frac{X_n}{n}\right| > 1\right) < \infty \quad \text{a.s.},$$

so

$$\sum_{n=1}^{\infty} \mathrm{P}^{\mathcal{F}}\left(|X| > n\right) < \infty \quad \text{a.s.}$$

Hence

$$\begin{aligned}
\mathrm{E}^{\mathcal{F}}|X| \leq \sum_{j=1}^{\infty}(j+1)\mathrm{P}^{\mathcal{F}}(j<|X|\leq j+1) &= 1 + \sum_{j=1}^{\infty} j\mathrm{P}^{\mathcal{F}}(j<|X|\leq j+1) \\
&= 1 + \sum_{j=1}^{\infty} \mathrm{P}^{\mathcal{F}}(j<|X|) < \infty \quad \text{a.s.} \quad \square
\end{aligned}$$

References

[1] P. Billingsley, *Convergence of Probability Measure*, Wiley, New York (1968).

[2] Ł. Kruk, W. Zięba, On almost sure convergence of asymptotic martingales, *Annales Universitas Marie Curie-Skłodowska*, **47**, No. 8 (1993), 82-89.

[3] J.M. Stoyanov, *Couterexamples in Probability*, John Wiley and Sons, Chichester, New York-Brisbane-Toronto-Singapore (1989).