

# Conditional independence, conditional mixing and conditional association

B. L. S. Prakasa Rao

Received: 25 July 2006 / Revised: 14 May 2007 / Published online: 21 September 2007  
© The Institute of Statistical Mathematics, Tokyo 2007

**Abstract** Some properties of conditionally independent random variables are studied. Conditional versions of generalized Borel-Cantelli lemma, generalized Kolmogorov's inequality and generalized Hájek-Rényi inequality are proved. As applications, a conditional version of the strong law of large numbers for conditionally independent random variables and a conditional version of the Kolmogorov's strong law of large numbers for conditionally independent random variables with identical conditional distributions are obtained. The notions of conditional strong mixing and conditional association for a sequence of random variables are introduced. Some covariance inequalities and a central limit theorem for such sequences are mentioned.

**Keywords** Conditional independence · Conditional mixing · Conditional association · Conditional Borel-Cantelli lemma · Generalized Kolmogorov inequality · Conditional Hájek-Rényi inequality · Conditional strong law of large numbers · Conditional central limit theorem · Conditional covariance inequality

## 1 Introduction

Our aim in this paper is to review the concept of conditional independence and introduce the notions of conditional strong mixing and conditional association for sequences of random variables. We discuss some stochastic inequalities and limit theorems for such sequences of random variables. Earlier discussions on the topic of conditional independence can be found in [Chow and Teicher \(1978\)](#) and more recently in [Majerak et al. \(2005\)](#).

---

B. L. S. Prakasa Rao (✉)  
Department of Mathematics and Statistics, University of Hyderabad, Hyderabad 500 046, India  
e-mail: blsprsm@uohyd.ernet.in

### 2 Conditional independence of events

Let  $(\Omega, \mathcal{A}, P)$  be a probability space. A set of events  $A_1, A_2, \dots, A_n$  are said to be independent if

$$P \left( \bigcap_{j=1}^k A_{i_j} \right) = \prod_{j=1}^k P(A_{i_j}) \tag{1}$$

for all  $1 \leq i_1 < i_2 < \dots < i_k \leq n, 2 \leq k \leq n$ .

**Definition 1** The set of events  $A_1, A_2, \dots, A_n$  are said to be conditionally independent given an event  $B$  with  $P(B) > 0$  if

$$P \left( \bigcap_{j=1}^k A_{i_j} | B \right) = \prod_{j=1}^k P(A_{i_j} | B) \tag{2}$$

for all  $1 \leq i_1 < i_2 < \dots < i_k \leq n, 2 \leq k \leq n$ .

The following examples (cf. Majerak et al. 2005) show that the independence of events does not imply conditional independence and that the conditional independence of events does not imply their independence.

*Example 1* Let  $\Omega = \{1, 2, 3, 4, 5, 6, 7, 8\}$  and  $p_i = 1/8$  be the probability assigned to the event  $\{i\}$ . Let  $A_1 = \{1, 2, 3, 4\}$  and  $A_2 = \{3, 4, 5, 6\}$ . Let  $B = \{2, 3, 4, 5\}$ . It is easy to see that the events  $A_1$  and  $A_2$  are independent but not conditionally independent given the event  $B$ .

*Example 2* Consider an experiment of choosing a coin numbered  $\{i\}$  with probability  $p_i = 1/n, 1 \leq i \leq n$  from a set of  $n$  coins and suppose it is tossed twice. Let  $p_0^i = 1/2^i$  be the probability for tails for the  $i$ th coin. Let  $A_1$  be the event that tail appears in the first toss,  $A_2$  be the event that tail appears in the second toss and  $H_i$  be the event that the  $i$ th coin is selected. It can be checked that the events  $A_1$  and  $A_2$  are conditionally independent given  $H_i$  but they are not independent as long as the number of coins  $n \geq 2$ .

### 3 Conditional independence of random variables

Let  $(\Omega, \mathcal{A}, P)$  be a probability space. Let  $\mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathcal{A}$  and let  $I_A$  denote the indicator function of an event  $A$ .

**Definition 2** The set of events  $A_1, A_2, \dots, A_n$  are said to be conditionally independent given  $\mathcal{F}$  or  $\mathcal{F}$ -independent if

$$E \left( \prod_{j=1}^k I_{A_{i_j}} | \mathcal{F} \right) = \prod_{j=1}^k E(I_{A_{i_j}} | \mathcal{F}) \text{ a.s.} \tag{3}$$

for all  $1 \leq i_1 < i_2 < \dots < i_k \leq n, 2 \leq k \leq n$ .

For the definition of conditional expectation of a measurable function  $X$  given a  $\sigma$ -algebra  $\mathcal{F}$ , see Doob (1953).

*Remark 1* If  $\mathcal{F} = \{\Phi, \Omega\}$ , then the above definition reduces to the usual definition of stochastic independence for random variables. If  $\mathcal{F} = \mathcal{A}$ , then the Eq. (3) reduces to the product of  $\mathcal{A}$ -measurable functions on both sides.

Let  $\{\zeta_n, n \geq 1\}$  be a sequence of classes of events. The sequence is said to be conditionally independent given  $\mathcal{F}$  if for all choices  $A_m \in \zeta_{k_m}$  where  $k_i \neq k_j$  for  $i \neq j, m = 1, 2, \dots, n$  and  $n \geq 2$ ,

$$E \left( \prod_{j=1}^k I_{A_j} | \mathcal{F} \right) = \prod_{j=1}^k E(I_{A_j} | \mathcal{F}) \text{ a.s.} \tag{4}$$

For any set of real valued random variables  $X_1, X_2, \dots, X_n$  defined on  $(\Omega, \mathcal{A}, P)$ , let  $\sigma(X_1, X_2, \dots, X_n)$  denote the smallest  $\sigma$ -algebra with respect to which they are measurable.

**Definition 3** A sequence of random variables  $\{X_n, n \geq 1\}$  defined on a probability space  $(\Omega, \mathcal{A}, P)$  is said to be *conditionally independent* given a sub- $\sigma$ -algebra  $\mathcal{F}$  or  $\mathcal{F}$ -*independent* if the sequence of classes of events  $\zeta_n = \sigma(X_n), n \geq 1$  are conditionally independent given  $\mathcal{F}$ .

It can be checked that a set of random variables  $X_1, X_2, \dots, X_n$  defined on a probability space  $(\Omega, \mathcal{A}, P)$  are conditionally independent given a sub- $\sigma$ -algebra  $\mathcal{F}$  if and only if for all  $(x_1, x_2, \dots, x_n) \in R^n$ ,

$$E \left( \prod_{i=1}^n I_{[X_i \leq x_i]} | \mathcal{F} \right) = \prod_{i=1}^n E(I_{[X_i \leq x_i]} | \mathcal{F}) \text{ a.s.}$$

*Remark 2* Independent random variables  $\{X_n, n \geq 1\}$  may lose their independence under conditioning. For instance, let  $\{X_1, X_2\}$  be Bernoulli trials with probability of success  $p$  with  $0 < p < 1$ . Let  $S_2 = X_1 + X_2$ . Then  $P(X_i = 1 | S_2 = 1) > 0, i = 1, 2$  but  $P(X_1 = 1, X_2 = 1 | S_2 = 1) = 0$ . On the other hand, dependent random variables may become independent under conditioning, that is, they become conditionally independent. This can be seen from the following discussion.

Let  $\{X_i, i \geq 1\}$  be independent positive integer-valued random variables. Then the sequence  $\{S_n, n \geq 1\}$  is a dependent sequence where  $S_n = X_1 + \dots + X_n$ . Let us consider the event  $[S_2 = k]$  with positive probability for some positive integer  $k$ . Check that

$$P(S_1 = i, S_3 = j | S_2 = k) = P(S_1 = i | S_2 = k)P(S_3 = j | S_2 = k).$$

Hence the random variables  $S_1$  and  $S_3$  are conditionally independent given  $S_2$ . If we interpret the subscript  $n$  of  $S_n$  as “time”, “past and future are conditionally independent

given the present". This property holds not only for sums  $S_n$  of independent random variables but also when the random sequence  $\{S_n, n \geq 1\}$  forms a time homogeneous Markov chain (cf. [Chow and Teicher 1978](#)).

*Remark 3* (i) A sequence of random variables  $\{X_n, n \geq 1\}$  defined on a probability space  $(\Omega, \mathcal{A}, P)$  is said to be *exchangeable* if the joint distribution of every finite subset of  $k$  of these random variables depends only upon  $k$  and not on the particular subset. It can be proved that the sequence of random variables  $\{X_n, n \geq 1\}$  is exchangeable if and only if the random variables are conditionally independent and identically distributed for a suitably chosen sub- $\sigma$ -algebra  $\mathcal{F}$  of  $\mathcal{A}$  (cf. [Chow and Teicher 1978](#), p. 220). If a random variable  $X$  is independent of a sequence of independent and identically distributed random variables  $\{X_n, n \geq 1\}$ , then the sequence  $\{Y_n, n \geq 1\}$  where  $Y_n = X + X_n$  forms an exchangeable sequence of random variables and hence conditionally independent.

(ii) Another example of a conditionally independent sequence is discussed in [Prakasa Rao \(1987\)](#) (cf. [Gyires 1981](#)). This can be described as follows. Let  $\theta_{h,j}^{(k)}, 1 \leq h, j \leq p, 1 \leq k \leq n$  be independent real valued random variables defined on a probability space  $(\Omega, \mathcal{A}, P)$ . Let  $\{\eta_j, j \geq 0\}$  be a homogeneous Markov chain defined on the same space with state space  $\{1, \dots, p\}$  and a nonsingular transition matrix  $A = ((a_{hj}))$ . We denote this Markov chain by  $\{A\}$ . Define  $\psi_k = \theta_{\eta_{k-1}, \eta_k}^{(k)}$  for  $1 \leq k \leq n$ . The sequence of random variables  $\{\psi_k, 1 \leq k \leq n\}$  is said to be defined on the homogeneous Markov chain  $\{A\}$ . Let  $\mathcal{F}$  be the sub- $\sigma$ -algebra generated by the sequence  $\{\eta_j, j \geq 0\}$ . It is easy to check that the random variables  $\{\psi_k, 1 \leq k \leq n\}$  are conditionally independent, in fact,  $\mathcal{F}$ -independent.

### 4 Conditional Borel-Cantelli lemma

[Majerak et al. \(2005\)](#) proved the following conditional version of Borel-Cantelli lemma.

**Theorem 1** *Let  $(\Omega, \mathcal{A}, P)$  be a probability space and let  $\mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathcal{A}$ . The following results hold.*

- (i) *Let  $\{A_n, n \geq 1\}$  be a sequence of events such that  $\sum_{n=1}^{\infty} P(A_n) < \infty$ . Then  $\sum_{n=1}^{\infty} E(I_{A_n}|\mathcal{F}) < \infty$  a.s.*
- (ii) *Let  $\{A_n, n \geq 1\}$  be a sequence of events and let  $A = \{\omega : \sum_{n=1}^{\infty} E(I_{A_n}|\mathcal{F}) < \infty\}$  with  $P(A) < 1$ . Then, only finitely many events from the sequence  $\{A_n \cap A, n \geq 1\}$  hold with probability one.*
- (iii) *Let  $\{A_n, n \geq 1\}$  be a sequence of  $\mathcal{F}$ -independent events and let  $A = \{\omega : \sum_{n=1}^{\infty} E(I_{A_n}|\mathcal{F}) = \infty\}$ . Then  $P(\limsup A_n) = P(A)$ .*

We will now obtain a conditional version of the Borel-Cantelli lemma due to [Petrov \(2004\)](#). For simplicity, we write  $P(A|\mathcal{F})$  for  $E^{\mathcal{F}}(I_A)$  in the following discussion.

Let  $A_1, A_2, \dots$  be a sequence of events and

$$A = \left\{ \omega : \sum_{n=1}^{\infty} P(A_n|\mathcal{F}) = \infty \right\}. \tag{5}$$

Let  $H$  be an arbitrary  $\mathcal{F}$ -measurable function. Define

$$\alpha_H = \liminf_{n \rightarrow \infty} \frac{\sum_{1 \leq i < k \leq n} (P(A_i A_k | \mathcal{F}) - H P(A_i | \mathcal{F}) P(A_k | \mathcal{F}))}{(\sum_{k=1}^n P(A_k | \mathcal{F}))^2}. \tag{6}$$

Observe that

$$2 \sum_{1 \leq i < k \leq n} P(A_i A_k | \mathcal{F}) = \sum_{i,k=1}^n P(A_i A_k | \mathcal{F}) - \sum_{k=1}^n P(A_k | \mathcal{F}) \text{ a.s.} \tag{7}$$

and

$$2 \sum_{1 \leq i < k \leq n} P(A_i | \mathcal{F}) P(A_k | \mathcal{F}) = \left( \sum_{k=1}^n P(A_k | \mathcal{F}) \right)^2 - \sum_{k=1}^n (P(A_k | \mathcal{F}))^2 \text{ a.s.} \tag{8}$$

Hence

$$\begin{aligned} 2\alpha_H = \liminf & \left[ \left( \sum_{i,k=1}^n P(A_i A_k | \mathcal{F}) \right) \left( \sum_{k=1}^n P(A_k | \mathcal{F}) \right)^{-2} - \left( \sum_{k=1}^n P(A_k | \mathcal{F}) \right)^{-1} - H \right. \\ & \left. + H \left( \sum_{k=1}^n (P(A_k | \mathcal{F}))^2 \right) \left( \sum_{k=1}^n P(A_k | \mathcal{F}) \right)^{-2} \right] \text{ a.s.} \end{aligned} \tag{9}$$

Since  $(P(A_k | \mathcal{F}))^2 \leq P(A_k | \mathcal{F})$  a.s. for every  $k \geq 1$ , it follows that the second and fourth terms on the right-hand side of the above equation converge to zero as  $n \rightarrow \infty$  on the set  $A$ . It is easy to see that

$$\left( \sum_{k=1}^n P(A_k | \mathcal{F}) \right)^2 \leq \sum_{i,k=1}^n P(A_i A_k | \mathcal{F}) \text{ a.s.} \tag{10}$$

by Cauchy-Schwarz inequality (cf. [Petrov 2004](#)). Hence, it follows that

$$2\alpha_H \geq 1 - H \text{ a.s.} \tag{11}$$

on the set  $A$ .

For any integer  $m \geq 1$ , let

$$\beta_H^{(m)} = \liminf_{n \rightarrow \infty} \frac{\sum_{m \leq i < k \leq n} (P(A_i A_k | \mathcal{F}) - H P(A_i | \mathcal{F}) P(A_k | \mathcal{F}))}{(\sum_{k=m}^n P(A_k | \mathcal{F}))^2}. \tag{12}$$

It can be shown that

$$\alpha_H = \beta_H^{(m)} \tag{13}$$

on the set  $A$  following the arguments given in Petrov (2004).

Following the inequality in Chung and Erdős (1952), it can be shown that, for every  $m < n$ ,

$$P\left(\bigcup_{k=m}^n A_k|\mathcal{F}\right) \geq \frac{(\sum_{k=m}^n P(A_k|\mathcal{F}))^2}{\sum_{i,k=m}^n P(A_i A_k|\mathcal{F})} \text{ a.s.} \tag{14}$$

Furthermore

$$\begin{aligned} \sum_{i,k=m}^n P(A_i A_k|\mathcal{F}) &= \sum_{k=m}^n P(A_k|\mathcal{F}) + 2 \sum_{m \leq i < k \leq n} P(A_i A_k|\mathcal{F}) \\ &= \sum_{k=m}^n P(A_k|\mathcal{F}) + T_1 + T_2 \end{aligned} \tag{15}$$

where

$$T_1 = 2 \sum_{m \leq i < k \leq n} (P(A_i A_k|\mathcal{F}) - H P(A_i|\mathcal{F})P(A_k|\mathcal{F})) \tag{16}$$

and

$$T_2 = 2H \sum_{m \leq i < k \leq n} P(A_i|\mathcal{F})P(A_k|\mathcal{F}). \tag{17}$$

It follows, by arguments given in Petrov (2004), that

$$\frac{T_2}{(\sum_{k=m}^n P(A_k|\mathcal{F}))^2} \rightarrow H \text{ a.s. as } n \rightarrow \infty \tag{18}$$

on the set  $A$ . Relations (14) and (15) imply that

$$\begin{aligned} P\left(\bigcup_{k=m}^n A_k|\mathcal{F}\right) &\geq \left[ \left(\sum_{k=m}^n P(A_k|\mathcal{F})\right)^{-1} + T_1 \left(\sum_{k=m}^n P(A_k|\mathcal{F})\right)^{-2} \right. \\ &\quad \left. + T_2 \left(\sum_{k=m}^n P(A_k|\mathcal{F})\right)^{-2} \right]^{-1} \text{ a.s.} \end{aligned} \tag{19}$$

on the set  $A$ . Let  $n \rightarrow \infty$  keeping  $m$  fixed. Then it follows that

$$P\left(\bigcup_{k=m}^{\infty} A_k|\mathcal{F}\right) \geq \left[ H + \liminf_{n \rightarrow \infty} T_1 \left(\sum_{k=m}^n P(A_k|\mathcal{F})^{-2}\right) \right]^{-1} \text{ a.s.} \tag{20}$$

on the set  $A$ . Applying the relation (13), we obtain that

$$P\left(\bigcup_{k=m}^{\infty} A_k | \mathcal{F}\right) \geq (H + 2\alpha_H)^{-1} \text{ a.s.} \tag{21}$$

on the set  $A$ . Let  $B_m = \cup_{k=m}^{\infty} A_k$ . Then we get that

$$P(B_m | \mathcal{F}) \geq (H + 2\alpha_H)^{-1} \text{ a.s.} \tag{22}$$

on the set  $A$  for every  $m \geq 1$ . Since the set  $B_m$  decreases as  $m$  increases, it follows that

$$\lim P(B_m | \mathcal{F}) = P\left(\bigcap_{m=1}^{\infty} B_m | \mathcal{F}\right) \text{ a.s.} \tag{23}$$

and

$$\limsup A_n = \bigcap_{m=1}^{\infty} \bigcup_{k=m}^{\infty} A_k = \bigcap_{m=1}^{\infty} B_m.$$

Hence

$$P(\limsup A_n | \mathcal{F}) \geq (H + 2\alpha_H)^{-1} \text{ a.s.} \tag{24}$$

on the set  $A = \{\omega : \sum_{n=1}^{\infty} P(A_n | \mathcal{F}) = \infty\}$ .

As a consequence of the above observations, we have the following theorem which is a conditional version of the generalized Borel-Cantelli lemma due to [Petrov \(2004\)](#).

**Theorem 2** *Let  $(\Omega, \mathcal{A}, P)$  be a probability space and let  $\mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathcal{A}$ . Let  $A_1, A_2, \dots$  be a sequence of events and let  $A = \{\omega : \sum_{n=1}^{\infty} P(A_n | \mathcal{F}) = \infty\}$ . Let  $H$  be any  $\mathcal{F}$ -measurable function. Define*

$$\alpha_H = \liminf_{n \rightarrow \infty} \frac{\sum_{1 \leq i < k \leq n} (P(A_i A_k | \mathcal{F}) - H P(A_i | \mathcal{F}) P(A_k | \mathcal{F}))}{(\sum_{k=1}^n P(A_k | \mathcal{F}))^2}. \tag{25}$$

Then

$$P(\limsup A_n | \mathcal{F}) \geq \frac{1}{H + 2\alpha_H} \text{ a.s.} \tag{26}$$

on the set  $A$ .

We now obtain another version of a conditional generalized Borel-Cantelli lemma due to [Kochen and Stone \(1964\)](#) following the method of [Yan \(2004\)](#).

**Theorem 3** Let  $(\Omega, \mathcal{A}, P)$  be a probability space and let  $\mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathcal{A}$ . Let  $A_1, A_2, \dots$  be a sequence of events and  $A = \{\omega : \sum_{n=1}^{\infty} P(A_n|\mathcal{F}) = \infty\}$ . Then

$$\begin{aligned} &P(A_n, n \geq 1 \text{ infinitely often } |A) \\ &\geq \limsup_{n \rightarrow \infty} \frac{(\sum_{k=1}^n P(A_k|\mathcal{F}))^2}{\sum_{i,k=1}^n P(A_i A_k|\mathcal{F})} \text{ a.s.} \\ &= \limsup_{n \rightarrow \infty} \frac{\sum_{1 \leq i \leq k \leq n} P(A_i|\mathcal{F})P(A_k|\mathcal{F})}{\sum_{1 \leq i \leq k \leq n} P(A_i A_k|\mathcal{F})} \text{ a.s.} \end{aligned} \tag{27}$$

In particular, if  $\{A_n, n \geq 1\}$  are pairwise  $\mathcal{F}$ -independent or  $\mathcal{F}$ -negatively correlated events, that is,  $P(A_i A_k|\mathcal{F}) - P(A_i|\mathcal{F})P(A_k|\mathcal{F}) \leq 0$ , for every  $1 \leq i \neq k$ , then

$$P(A_n, n \geq 1 \text{ infinitely often } |A) = 1. \tag{28}$$

*Proof* Let  $a_n = (\sum_{k=1}^n P(A_k|\mathcal{F}))^2, b_n = \sum_{i,k=1}^n P(A_i A_k|\mathcal{F})$ . Then, on the set  $A, \lim_{n \rightarrow \infty} a_n = \infty$  a.s. The inequality (14) implies that  $\lim_{n \rightarrow \infty} b_n = \infty$  a.s. on the set  $A$ . From the relation (14) and the fact that

$$\sum_{i,k=m+1}^n P(A_i A_k|\mathcal{F}) \leq b_n - b_m \text{ a.s.,} \tag{29}$$

we get that

$$\begin{aligned} P\left(\bigcup_{k=m+1}^{\infty} A_k|\mathcal{F}\right) &= \lim_{n \rightarrow \infty} P\left(\bigcup_{k=m+1}^n A_k|\mathcal{F}\right) \\ &\geq \limsup_{n \rightarrow \infty} \frac{(\sqrt{a_n} - \sqrt{a_m})^2}{b_n - b_m} = \limsup_{n \rightarrow \infty} \frac{a_n}{b_n} \text{ a.s.} \end{aligned} \tag{30}$$

on the set  $A$ . Letting  $m \rightarrow \infty$ , we get the inequality in (27). Since  $\sum_{k=1}^{\infty} P(A_k|\mathcal{F}) = \infty$  a.s on the set  $A$ , and since

$$\left(\sum_{k=1}^n P(A_k|\mathcal{F})\right)^2 \leq 2 \sum_{1 \leq i \leq k \leq n} P(A_i|\mathcal{F})P(A_k|\mathcal{F}) + \sum_{k=1}^n P(A_k|\mathcal{F}),$$

it follows that

$$\lim_{n \rightarrow \infty} \frac{\sum_{k=1}^n P(A_k|\mathcal{F})}{\sum_{1 \leq i \leq k \leq n} P(A_i|\mathcal{F})P(A_k|\mathcal{F})} = 0 \text{ a.s.} \tag{31}$$

on the set  $A$ . Thus the equality in (27) holds. □



### 5 Generalized Kolmogorov inequality

Majerak et al. (2005) proved a conditional version of Kolmogorov’s inequality and derived a conditional version of Kolmogorov’s strong law of large numbers for a sequence of conditionally independent random variables. We will now derive a conditional version of *Generalized Kolmogorov’s inequality* due to Loève (1977), p. 275. We assume that the conditional expectations exist and the conditional distributions exist as regular conditional distributions in the following discussion (cf. Chow and Teicher 1978).

**Theorem 4** *Let  $\{X_k, 1 \leq k \leq n\}$  be a set of  $\mathcal{F}$ -independent random variables with  $E^{\mathcal{F}}[|X_k|^r] < \infty, 1 \leq k \leq n$  for some  $r \geq 1$  where  $E^{\mathcal{F}}(Z)$  denotes the conditional expectation of a random variable  $Z$  given a sub- $\sigma$ -algebra  $\mathcal{F}$ . For an arbitrary  $\mathcal{F}$ -measurable random variable  $\epsilon > 0$  a.s., let  $S_k = X_1 + \dots + X_k, k \geq 1$  and let*

$$C = \left[ \max_{1 \leq k \leq n} |S_k - E^{\mathcal{F}} S_k| \geq \epsilon \right].$$

Then

$$\epsilon^r P(C|\mathcal{F}) \leq E^{\mathcal{F}}[|S_n - E^{\mathcal{F}} S_n|^r I_C] \leq E^{\mathcal{F}}[|S_n - E^{\mathcal{F}} S_n|^r] \text{ a.s.} \tag{32}$$

We will prove now a lemma which will be used in the proof of the above theorem.

**Lemma 1** *Let  $X$  and  $Y$  be  $\mathcal{F}$ -independent random variables such that  $E^{\mathcal{F}}|X|^r < \infty$  and  $E^{\mathcal{F}}|Y|^r < \infty$  where  $r \geq 1$ . Then, for any event  $A \in \sigma(X)$ ,*

$$E^{\mathcal{F}}[|X - E^{\mathcal{F}} X + Y - E^{\mathcal{F}} Y|^r I_A] \geq E^{\mathcal{F}}[|X - E^{\mathcal{F}} X|^r I_A].$$

*Proof* Let the conditional distribution of  $X - E^{\mathcal{F}} X$  given  $\mathcal{F}$  be denoted by  $F_{X-E^{\mathcal{F}} X}^{\mathcal{F}}$  and the conditional distribution of  $Y - E^{\mathcal{F}} Y$  given  $\mathcal{F}$  be denoted by  $F_{Y-E^{\mathcal{F}} Y}^{\mathcal{F}}$ . Observe that

$$\begin{aligned} |x - E^{\mathcal{F}} X|^r &= |E^{\mathcal{F}}(x - E^{\mathcal{F}} X + Y - E^{\mathcal{F}} Y)|^r \\ &\leq E^{\mathcal{F}}[|x - E^{\mathcal{F}} X + Y - E^{\mathcal{F}} Y|^r] \end{aligned} \tag{33}$$

by Jensen’s inequality for conditional expectations. Hence

$$\begin{aligned} &E^{\mathcal{F}}[|X - E^{\mathcal{F}} X + Y - E^{\mathcal{F}} Y|^r I_A] \\ &= \int_A dF_{X-E^{\mathcal{F}} X}^{\mathcal{F}}(x) \int_{-\infty}^{\infty} |x - E^{\mathcal{F}} X + y - E^{\mathcal{F}} Y|^r dF_{Y-E^{\mathcal{F}} Y}^{\mathcal{F}}(y) \\ &\geq \int_A |x - E^{\mathcal{F}} X|^r dF_{X-E^{\mathcal{F}} X}^{\mathcal{F}}(x) = E^{\mathcal{F}}[|X - E^{\mathcal{F}} X|^r I_A]. \end{aligned} \tag{34}$$

□

*Remark 4* It is easy to see that the above lemma holds if  $X$  is a sum of  $\mathcal{F}$ -independent random variables  $X_1, X_2, \dots, X_k, Y$  is a sum of  $\mathcal{F}$ -independent random variables  $X_{k+1}, \dots, X_n$  and  $A \in \sigma(X_1, X_2, \dots, X_k)$ . We will use the above lemma in this form.

*Proof of Theorem 4* Let

$$\tau = \inf\{k : |S_k - E^{\mathcal{F}} S_k| \geq \epsilon, 1 \leq k \leq n\},$$

$$A_k = C \cap [\tau = k],$$

and

$$R_k = X_{k+1} + \dots + X_n.$$

Note that

$$\begin{aligned} E^{\mathcal{F}}[|S_n - E^{\mathcal{F}} S_n|^r I_{A_k}] &= E^{\mathcal{F}}[|S_k + R_k - E^{\mathcal{F}} S_k - E^{\mathcal{F}} R_k|^r I_{A_k}] \\ &= E^{\mathcal{F}}[|S_k - E^{\mathcal{F}} S_k + R_k - E^{\mathcal{F}} R_k|^r I_{A_k}] \\ &\geq E^{\mathcal{F}}[|S_k - E^{\mathcal{F}} S_k|^r I_{A_k}] \quad (\text{by Lemma 1}) \\ &\geq \epsilon^r P(A_k|\mathcal{F}). \end{aligned} \tag{35}$$

Summing over  $k = 1, 2, \dots, n$  on both sides of the above inequality, we get that

$$\begin{aligned} E^{\mathcal{F}}[|S_n - E^{\mathcal{F}} S_n|^r] &\geq E^{\mathcal{F}}[|S_n - E^{\mathcal{F}} S_n|^r I_C] \\ &= E^{\mathcal{F}} \left[ |S_n - E^{\mathcal{F}} S_n|^r \sum_{k=1}^n I_{A_k} \right] \\ &= \sum_{k=1}^n E^{\mathcal{F}}[|S_n - E^{\mathcal{F}} S_n|^r I_{A_k}] \\ &\geq \sum_{k=1}^n \epsilon^r P(A_k|\mathcal{F}) \quad (\text{by Eq. 35}) \\ &= \epsilon^r P(C|\mathcal{F}). \end{aligned} \tag{36}$$

Hence

$$\epsilon^r P(C|\mathcal{F}) \leq E^{\mathcal{F}}[|S_n - E^{\mathcal{F}} S_n|^r I_C] \leq E^{\mathcal{F}}[|S_n - E^{\mathcal{F}} S_n|^r].$$

As a consequence of Theorem 4, we have the following corollary.

**Corollary 1** *Let  $\{X_{nk}, 1 \leq k \leq k_n, n \geq 1\}$  be a triangular array of random variables such that the random variables  $\{X_{nk}, 1 \leq k \leq k_n\}$  are  $\mathcal{F}$ -independent for every*

$n \geq 1$ . Let  $S_{nk} = X_{n1} + \dots + X_{nk}$ ,  $1 \leq k \leq k_n$ . Suppose that  $E^{\mathcal{F}}[|X_{nk}|^r] < \infty$  for  $1 \leq k \leq k_n$ ,  $n \geq 1$  for some  $r \geq 1$ . Further suppose that

$$E^{\mathcal{F}} \left[ |S_{nk_n} - E^{\mathcal{F}} S_{nk_n}|^r \right] \rightarrow 0$$

almost surely as  $n \rightarrow \infty$ . Then, for every  $\mathcal{F}$ -measurable positive random variable  $\epsilon$ ,

$$P \left( \max_{1 \leq k \leq k_n} |S_{nk} - E^{\mathcal{F}} S_{nk}| \geq \epsilon | \mathcal{F} \right) \rightarrow 0$$

almost surely as  $n \rightarrow \infty$ .

### Conditional Hájek-Rényi Inequality

The following result gives a conditional version of the Hájek-Rényi inequality (cf. Hájek and Rényi 1955) which in turn generalizes the conditional Kolmogorov inequality in Majerak et al. (2005).

**Theorem 5** Let  $\{X_k, 1 \leq k \leq m\}$  be a set of  $\mathcal{F}$ -independent random variables with  $E^{\mathcal{F}}[X_k^2] < \infty$ ,  $1 \leq k \leq m$  where  $E^{\mathcal{F}}(Z)$  denotes the conditional expectation of a random variable  $Z$  given a sub- $\sigma$ -algebra  $\mathcal{F}$ . For an arbitrary  $\mathcal{F}$ -measurable random variable  $\epsilon > 0$  a.s., let  $S_k = X_1 + \dots + X_k$ ,  $k \geq 1$  and let

$$C = \left[ \max_{n \leq k \leq m} c_k |S_k - E^{\mathcal{F}} S_k| \geq \epsilon \right]$$

where  $c_k, 1 \leq k \leq m$  is a non-decreasing sequence of positive  $\mathcal{F}$ -measurable random variables a.s. Then, for any positive integers  $1 \leq n \leq m$ ,

$$\epsilon^2 P(C | \mathcal{F}) \leq c_n^2 \sum_{k=1}^n E^{\mathcal{F}} [X_k - E^{\mathcal{F}}(X_k)]^2 + \sum_{k=n+1}^m c_k^2 E^{\mathcal{F}} [X_k - E^{\mathcal{F}}(X_k)]^2 \quad \text{a.s.} \quad (37)$$

*Proof* Without loss of generality, assume that  $E^{\mathcal{F}}(X_k) = 0$ ,  $1 \leq k \leq m$ . Let  $\epsilon > 0$  a.s. be an arbitrary  $\mathcal{F}$ -measurable random variable and let  $A = \{\max_{n \leq k \leq m} c_k |S_k| \geq \epsilon\}$ . Let

$$\tau = \inf\{k : c_k |S_k| \geq \epsilon, n \leq k \leq m\}.$$

Observe that the random variable  $\tau$  takes the values  $n, \dots, m$  on the set  $A$ . Let  $A_r = A \cap [\tau = r]$ ,  $n \leq r \leq m$ . Let

$$\zeta = \sum_{k=n}^{m-1} S_k^2 (c_k^2 - c_{k+1}^2) + c_m^2 S_m^2. \quad (38)$$

Then

$$\begin{aligned}
 E^{\mathcal{F}}[\zeta] &= \sum_{k=n}^{m-1} (c_k^2 - c_{k+1}^2) E^{\mathcal{F}}(S_k^2) + c_m^2 E^{\mathcal{F}}(S_m^2) \\
 &= c_n^2 \sum_{k=1}^n E^{\mathcal{F}}[X_k^2] + \sum_{k=n+1}^m c_k^2 E^{\mathcal{F}}[X_k^2].
 \end{aligned}
 \tag{39}$$

The last equality follows from the  $\mathcal{F}$ -independence of the random sequence  $X_i, 1 \leq k \leq m$ . We have to prove that

$$\epsilon^2 \sum_{r=n}^m P(A_r | \mathcal{F}) \leq E^{\mathcal{F}}[\zeta].
 \tag{40}$$

Observe that

$$E^{\mathcal{F}}[\zeta] \geq E^{\mathcal{F}}[\zeta I_A] = \sum_{r=n}^m E^{\mathcal{F}}[\zeta I_{A_r}].
 \tag{41}$$

Furthermore

$$E^{\mathcal{F}}[\zeta I_{A_r}] = \sum_{k=n}^{m-1} (c_k^2 - c_{k+1}^2) E^{\mathcal{F}}(S_k^2 I_{A_r}) + c_m^2 E^{\mathcal{F}}(S_m^2 I_{A_r}).
 \tag{42}$$

In addition, we note that, for any  $k$  such that  $r \leq k \leq m$ ,

$$\begin{aligned}
 E^{\mathcal{F}}[S_k^2 I_{A_r}] &= E^{\mathcal{F}}[(S_r + S_k - S_r)^2 I_{A_r}] \\
 &\geq E^{\mathcal{F}}[S_r^2 I_{A_r}] \\
 &\geq \frac{\epsilon^2}{c_r^2} P(A_r | \mathcal{F}).
 \end{aligned}
 \tag{43}$$

The second inequality in the above follows from the fact that the random variables  $S_r I_{A_r}$  and  $S_k - S_r$  are  $\mathcal{F}$ -independent for  $r \leq k \leq m$ . Combining the inequalities in (41) to (43), we obtain (40) proving the conditional Hájek-Rényi inequality.  $\square$

*Remark 5* As corollaries to Theorem 5, we can obtain the following results.

(i) If we choose  $n = 1$  and  $c_1 = c_2 = \dots = c_m = 1$ , then we get the conditional version of Kolmogorov’s inequality proved in Theorem 3.4 of [Majerak et al. \(2005\)](#). If we choose  $c_k = 1/k, k = n + 1, \dots, m$ , then we obtain the inequality

$$\begin{aligned}
 &\epsilon^2 P\left(\left[\max_{n \leq k \leq m} \frac{|S_k - E^{\mathcal{F}}(S_k)|}{k} \geq \epsilon\right] | \mathcal{F}\right) \\
 &\leq \frac{\sum_{k=1}^n E^{\mathcal{F}}[X_k - E^{\mathcal{F}}(X_k)]^2}{n^2} + \sum_{k=n+1}^m \frac{E^{\mathcal{F}}[X_k - E^{\mathcal{F}}(X_k)]^2}{k^2} \text{ a.s.}
 \end{aligned}
 \tag{44}$$

(ii) Letting  $m \rightarrow \infty$  in Theorem 5, it can be seen that

$$\begin{aligned} &\epsilon^2 P \left( \left[ \sup_{n \leq k} c_k |S_k - E^{\mathcal{F}}(S_k)| \geq \epsilon \right] | \mathcal{F} \right) \\ &\leq c_n^2 \sum_{k=1}^n E^{\mathcal{F}} [X_k - E^{\mathcal{F}}(X_k)]^2 + \sum_{k=n+1}^{\infty} c_k^2 E^{\mathcal{F}} [X_k - E^{\mathcal{F}}(X_k)]^2 \text{ a.s.} \end{aligned} \tag{45}$$

Choosing  $c_k = \frac{1}{k}, k \geq 1$  in the above inequality, we get that

$$\begin{aligned} &\epsilon^2 P \left( \left[ \sup_{n \leq k} \frac{|S_k - E^{\mathcal{F}}(S_k)|}{k} \geq \epsilon \right] | \mathcal{F} \right) \\ &\leq \frac{\sum_{k=1}^n E^{\mathcal{F}} [X_k - E^{\mathcal{F}}(X_k)]^2}{n^2} + \sum_{k=n+1}^{\infty} \frac{E^{\mathcal{F}} [X_k - E^{\mathcal{F}}(X_k)]^2}{k^2} \text{ a.s.} \end{aligned} \tag{46}$$

(iii) Let  $\{X_k, 1 \leq k\}$  be a set of  $\mathcal{F}$ -independent random variables with  $E^{\mathcal{F}} [X_k^2] < \infty, k \geq 1$ . Suppose that

$$\sum_{k=1}^{\infty} \frac{E^{\mathcal{F}} [X_k - E^{\mathcal{F}} X_k]^2}{k^2} < \infty \text{ a.s.} \tag{47}$$

From the inequality obtained in (45), it follows that for any  $\mathcal{F}$ -measurable random variable  $\epsilon > 0$  a.s.,

$$\epsilon^2 P \left( \left[ \sup_{n \leq k} \frac{|S_k - E^{\mathcal{F}}(S_k)|}{k} \geq \epsilon \right] | \mathcal{F} \right) \rightarrow 0 \text{ a.s.} \tag{48}$$

as  $n \rightarrow \infty$ . Hence

$$P \left( \lim_{n \rightarrow \infty} \frac{S_n - E^{\mathcal{F}}(S_n)}{n} = 0 | \mathcal{F} \right) = 1 \text{ a.s.} \tag{49}$$

This result is derived in Theorem 3.5 of Majerak et al. (2005) as a consequence of the conditional Kolmogorov inequality.

### 6 Conditional strong law of large numbers

We now obtain a generalized version of the conditional Kolmogorov’s strong law of large numbers proved in Majerak et al. (2005).

Let  $\{X_n, n \geq 1\}$  be a sequence of  $\mathcal{F}$ -independent random variables with  $E^{\mathcal{F}} |X_n - E^{\mathcal{F}} X_n|^r < \infty, n \geq 1$  for some  $r \geq 1$ . Let  $S_n = X_1 + \dots + X_n, n \geq 1$ . From the elementary inequality

$$\left| \sum_{k=1}^n a_k^2 \right|^r \leq n^{r-1} \sum_{k=1}^n |a_k|^{2r} \tag{50}$$

for any sequence of real numbers  $a_k, 1 \leq k \leq n$  whenever  $r \geq 1$ , it follows that

$$E^{\mathcal{F}} |S_n - E^{\mathcal{F}} S_n|^{2r} \leq n^{r-1} \sum_{k=1}^n E^{\mathcal{F}} |X_k - E^{\mathcal{F}} X_k|^{2r} \text{ a.s.} \tag{51}$$

whenever  $r \geq 1$ .

**Theorem 6** *If  $\{X_n, n \geq 1\}$  is a sequence of  $\mathcal{F}$ -independent random variables such that*

$$\sum_{n=1}^{\infty} \frac{E^{\mathcal{F}} |X_n - E^{\mathcal{F}} X_n|^{2r}}{n^{r+1}} < \infty \text{ a.s} \tag{52}$$

for some  $r \geq 1$ . Then, conditionally on  $\mathcal{F}$ ,

$$\frac{S_n - E^{\mathcal{F}} S_n}{n} \rightarrow 0 \text{ a.s. as } n \rightarrow \infty. \tag{53}$$

*Proof* Without loss of generality, we assume that  $E^{\mathcal{F}} X_k = 0, k \geq 1$ . Note that  $E^{\mathcal{F}}(S_n) = 0$ . For an arbitrary  $\mathcal{F}$ -measurable random variable  $\epsilon > 0$ , let

$$D_k = \{\omega : |S_n| > n\epsilon, n \in [2^k, 2^{k+1})\}. \tag{54}$$

From the definition of the set  $D_k$ , it follows that  $|S_n| > \epsilon 2^k$  for some  $2^k \leq n < 2^{k+1}$ . Hence, by the Generalized Kolmogorov’s inequality proved in Theorem 4, it follows that

$$\begin{aligned} (\epsilon 2^k)^{2r} P(D_k | \mathcal{F}) &\leq E^{\mathcal{F}} |S_{2^{k+1}}|^{2r} \\ &\leq (2^{k+1})^{r-1} \sum_{n=1}^{2^{k+1}} E^{\mathcal{F}} |X_n|^{2r} \end{aligned} \tag{55}$$

for every  $k \geq 1$ . Hence

$$\begin{aligned} \sum_{k=0}^{\infty} P(D_k | \mathcal{F}) &\leq \frac{1}{\epsilon^{2r}} \sum_{k=0}^{\infty} \frac{2^{(k+1)(r-1)}}{2^{2kr}} \sum_{n=1}^{2^{k+1}} E^{\mathcal{F}} |X_n|^{2r} \text{ a.s} \\ &= \frac{1}{\epsilon^{2r}} \sum_{n=1}^{\infty} E^{\mathcal{F}} |X_n|^{2r} \sum_{k: 2^{k+1} \geq n} \frac{2^{(k+1)(r-1)}}{2^{2kr}} \text{ a.s.} \end{aligned} \tag{56}$$

Let  $k_n$  be the smallest integer  $k$  such that  $2^{k+1} \geq n$ . It is easy to check that

$$\sum_{k:2^{k+1} \geq n} \frac{2^{(k+1)(r-1)}}{2^{2kr}} \leq c_r \frac{1}{n^{r+1}} \tag{57}$$

for some positive constant  $c_r$ . Therefore

$$\sum_{k=0}^{\infty} P(D_k|\mathcal{F}) \leq \frac{c_r}{\epsilon^{2r}} \sum_{n=1}^{\infty} \frac{E^{\mathcal{F}}|X_n|^{2r}}{n^{r+1}} \text{ a.s.} \tag{58}$$

and the last term is finite a.s. by assumption. Hence, by the conditional Borel-Cantelli lemma stated earlier (cf. Theorem 1) due to Majerak et al. (2005), it follows that, for an arbitrary  $\mathcal{F}$ -measurable random variable  $\epsilon > 0$  a.s., the conditional probability of the event that  $|S_n| > n\epsilon$  holds infinitely often given  $\mathcal{F}$  is equal to zero. Hence, conditional on  $\mathcal{F}$ ,

$$\frac{S_n}{n} \rightarrow 0 \text{ a.s. as } n \rightarrow \infty.$$

Suppose  $X$  and  $Y$  are two random variables defined on a probability space  $(\Omega, \mathcal{A}, P)$  and  $\mathcal{F}$  is a nonempty sub- $\sigma$ -algebra of  $\mathcal{A}$ . The random variables  $X$  and  $Y$  are said to have identical conditional distributions if

$$E^{\mathcal{F}}(I_{X \leq a}) = E^{\mathcal{F}}(I_{Y \leq a}) \text{ a.s.}$$

for  $-\infty < a < \infty$ . As a special case of the above theorem, we obtain the following conditional version of the strong law of large numbers for  $\mathcal{F}$ -independent random variables with identical conditional distributions. This was also proved in Majerak et al. (2005). □

**Theorem 7** *Let  $\{X_n, n \geq 1\}$  be a sequence of  $\mathcal{F}$ -independent random variables with identical conditional distributions. Then, conditional on  $\mathcal{F}$ ,*

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

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

### 7 Conditional central limit theorem

The following theorem holds for  $\mathcal{F}$ -conditionally independent random variables with identical conditional distributions.

**Theorem 8** Let  $\{X_n, n \geq 1\}$  be a sequence of  $\mathcal{F}$ -independent random variables with identical conditional distributions with almost surely positive finite conditional variance  $\sigma_{\mathcal{F}}^2 = E^{\mathcal{F}}[X_1 - E^{\mathcal{F}}X_1]^2$ . Let  $S_n = X_1 + \dots + X_n$ . Then

$$E^{\mathcal{F}} \left( \exp \left\{ it \left[ \frac{S_n - E^{\mathcal{F}}S_n}{\sqrt{n}\sigma_{\mathcal{F}}} \right] \right\} \right) \rightarrow \exp(-t^2/2) \text{ a.s.} \tag{59}$$

as  $n \rightarrow \infty$  for every  $t \in R$ .

Proof of this result follows by standard arguments. We omit the proof. A more general version of the conditional central limit theorem for conditionally independent random variables can be proved under conditional versions of the Lindeberg condition.

### 8 Conditional strong-mixing

The concept of strong-mixing (or  $\alpha$ -mixing ) for sequences of random variables was introduced by Rosenblatt (1956) to study short range dependence. Let  $\{X_n, n \geq 1\}$  be a sequence of random variables defined on a probability space  $(\Omega, \mathcal{A}, P)$ . Suppose there exists a sequence  $\alpha_n > 0$  such that

$$|P(A \cap B) - P(A)P(B)| \leq \alpha_n \tag{60}$$

for all  $A \in \sigma(X_1, \dots, X_k), B \in \sigma(X_{k+n}, X_{k+n+1}, \dots)$  and  $k \geq 1, n \geq 1$ . Suppose that  $\alpha_n \rightarrow 0$  as  $n \rightarrow \infty$ . Then the sequence is said to be *strong-mixing*. Note that the set  $(X_1, X_2, \dots, X_k)$  and  $(X_{k+n}, X_{k+n+1}, \dots)$  are approximately independent for large  $n$ . For a survey on mixing sequences and their properties, see Roussas and Ioannides (1987). Suppose the sequence  $\{X_n, n \geq 1\}$  is observed at random integer-valued times  $\{\tau_n, n \geq 1\}$ . It is not necessarily true that the sequence  $\{X_{\tau_n}, n \geq 1\}$  be strong-mixing even if the original sequence  $\{X_n, n \geq 1\}$  is strong-mixing. In order to see when the property of strong-mixing of a sequence is inherited by such subsequences, we have introduced the notion of *mixing strongly* in Prakasa Rao (1990) and obtained some moment inequalities.

In analogy with the concept of conditional independence discussed earlier, we will now investigate the concept of *conditional strong mixing* for a sequence of random variables.

**Definition 4** Let  $(\Omega, \mathcal{A}, P)$  be a probability space and let  $\mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathcal{A}$ . Let  $\{X_n, n \geq 1\}$  be a sequence of random variables defined on  $(\Omega, \mathcal{A}, P)$ . The sequence of random variables  $\{X_n, n \geq 1\}$  is said to be *conditionally strong-mixing* ( $\mathcal{F}$ -strong-mixing) if there exists a nonnegative  $\mathcal{F}$ -measurable random variable  $\alpha_n^{\mathcal{F}}$  converging to zero almost surely as  $n \rightarrow \infty$  such that

$$|P(A \cap B|\mathcal{F}) - P(A|\mathcal{F})P(B|\mathcal{F})| \leq \alpha_n^{\mathcal{F}} \text{ a.s.} \tag{61}$$

for all  $A \in \sigma(X_1, \dots, X_k), B \in \sigma(X_{k+n}, X_{k+n+1}, \dots)$  and  $k \geq 1, n \geq 1$ .



It is clear that this concept generalizes the concept of conditional independence. From the remarks made above, the property of conditional strong mixing does not imply strong mixing for a sequence of random variables and vice versa.

An example of a conditionally strong-mixing sequence can be constructed in the following manner. Let  $\{X_n = f(Y_n^{(N)}), n \geq 1\}$  be a sequence of random variables defined on a probability space  $(\Omega, \mathcal{A}, P)$  such that, given  $N = j$ , the sequence  $\{Y_n^{(j)}, n \geq 1\}$  is a homogeneous Markov chain with finite state space and positive transition probabilities where  $N$  is a positive integer-valued random variable defined on the probability space  $(\Omega, \mathcal{A}, P)$ . For any fixed  $j \geq 1$ , the Markov chain  $\{Y_n^{(j)}, n \geq 1\}$  has a stationary distribution. Suppose the initial distribution of the random variable  $Y_1^{(j)}$  is this stationary distribution. Then the process  $\{Y_n^{(j)}, n \geq 1\}$  is a stationary Markov chain for every  $j \geq 1$ . Let  $\mathcal{F}$  be the  $\sigma$ -algebra generated by the random variable  $N$ . Then the sequence  $\{X_n, n \geq 1\}$  is conditionally strong-mixing, that is,  $\mathcal{F}$ -strong-mixing. This can be seen from the fact that, given the event  $[N = j]$ , the sequence  $\{Y_n^{(j)}, n \geq 1\}$  is strong-mixing with mixing coefficient  $\alpha_n^{(j)} = r_j \rho_j^n$  for some  $r_j \geq 0$  and  $0 < \rho_j < 1$  (cf. Billingsley 1986, p. 375).

The following covariance inequalities hold for  $\mathcal{F}$ -strong-mixing sequence of random variables.

**Theorem 9** *Let  $\{X_n, n \geq 1\}$  be  $\mathcal{F}$ -strong mixing sequence of random variables with mixing coefficient  $\alpha_n^{\mathcal{F}}$  defined on a probability space  $(\Omega, \mathcal{A}, P)$ .*

- (i) *Suppose that a random variable  $Y$  is measurable with respect to the  $\sigma(X_1, \dots, X_k)$  and bounded by an  $\mathcal{F}$ -measurable function  $C$  and  $Z$  is another random variable measurable with respect to the  $\sigma(X_{k+n}, X_{k+n+1}, \dots)$  bounded by an  $\mathcal{F}$ -measurable function  $D$ . Then*

$$|E^{\mathcal{F}}[YZ] - E^{\mathcal{F}}[Y]E^{\mathcal{F}}[Z]| \leq 4CD\alpha_n^{\mathcal{F}} \text{ a.s.} \tag{62}$$

- (ii) *Suppose that a random variable  $Y$  is measurable with respect to the  $\sigma(X_1, \dots, X_k)$  and  $E^{\mathcal{F}}[Y^4]$  is bounded by a  $\mathcal{F}$ -measurable function  $C$  and  $Z$  is another random variable measurable with respect to the  $\sigma(X_{k+n}, X_{k+n+1}, \dots)$  and  $E^{\mathcal{F}}[Z^4]$  is bounded by a  $\mathcal{F}$ -measurable function  $D$ . Then*

$$|E^{\mathcal{F}}[YZ] - E^{\mathcal{F}}[Y]E^{\mathcal{F}}[Z]| \leq 8(1 + C + D)[\alpha_n^{\mathcal{F}}]^{1/2} \text{ a.s.} \tag{63}$$

**Definition 5** Let  $(\Omega, \mathcal{A}, P)$  be a probability space and let  $\mathcal{F}$  be a sub- $\sigma$ -algebra of  $\mathcal{A}$ . A sequence of random variables  $\{X_n, n \geq 1\}$  defined on  $(\Omega, \mathcal{A}, P)$  is said to be  $\mathcal{F}$ -stationary or conditionally stationary if the joint distribution of  $(X_{t_1}, \dots, X_{t_k})$  conditioned on  $\mathcal{F}$  is the same as the joint distribution of  $(X_{t_1+r}, \dots, X_{t_k+r})$  conditioned on  $\mathcal{F}$  a.s. for all  $1 \leq t_1 < \dots < t_k \leq \infty, r \geq 1$ .

As a consequence of the above theorem, the following central limit theorem can be proved.

**Theorem 10** Let  $\{X_n, n \geq 1\}$  be a  $\mathcal{F}$ -stationary and  $\mathcal{F}$ -strong mixing sequence of random variables, with mixing coefficient  $\alpha_n^{\mathcal{F}} = O(n^{-5})$  almost surely, defined on a probability space  $(\Omega, \mathcal{A}, P)$  with  $E^{\mathcal{F}}[X_1] = 0$  a.s. and  $E^{\mathcal{F}}[X_n^{12}] < \infty$  a.s. Let  $S_n = X_1 + \dots + X_n$ . Then

$$n^{-1} E^{\mathcal{F}}[S_n]^2 \rightarrow \sigma_{\mathcal{F}}^2 \equiv E^{\mathcal{F}}[X_1^2] + 2 \sum_{k=1}^{\infty} E^{\mathcal{F}}[X_1 X_{k+1}] \text{ a.s.} \tag{64}$$

and the series converges absolutely almost surely. If  $\sigma_{\mathcal{F}} > 0$  almost surely, then

$$E^{\mathcal{F}}[e^{itS_n/(\sigma_{\mathcal{F}}\sqrt{n})}] \rightarrow e^{-t^2/2} \text{ a.s. as } n \rightarrow \infty. \tag{65}$$

Proofs of Theorems 9 and 10 in the conditional framework follow the proofs given in Billingsley (1986) subject to necessary modifications. We omit the proofs.

*Remark 6* The conditions in Theorems 9 and 10 can be weakened further but we do not go into the details here. The notion of  $\phi$ -mixing can also be generalized to a conditional framework in an analogous way.

### 9 Conditional association

Let  $X$  and  $Y$  be random variables defined on a probability space  $(\Omega, \mathcal{A}, P)$  with  $E(X^2) < \infty$  and  $E(Y^2) < \infty$ . Let  $\mathcal{F}$ - be a sub- $\sigma$ -algebra of  $\mathcal{A}$ . We define the conditional covariance of  $X$  and  $Y$  given  $\mathcal{F}$  or  $\mathcal{F}$ -covariance as

$$\text{Cov}^{\mathcal{F}}(X, Y) = E^{\mathcal{F}}[(X - E^{\mathcal{F}}X)(Y - E^{\mathcal{F}}Y)].$$

It easy to see that  $\mathcal{F}$ -covariance reduces to the ordinary concept of covariance when  $\mathcal{F} = \{\Phi, \Omega\}$ . A set of random variables  $\{X_k, 1 \leq k \leq n\}$  is said to be  $\mathcal{F}$ -associated if, for any coordinatewise non-decreasing functions  $h, g$  defined on  $R^n$ ,

$$\text{Cov}^{\mathcal{F}}(h(X_1, \dots, X_n), g(X_1, \dots, X_n)) \geq 0 \text{ a.s.}$$

A sequence of random variables  $\{X_n, n \geq 1\}$  is said to be  $\mathcal{F}$ -associated if every finite subset of the sequence  $\{X_n, n \geq 1\}$  is  $\mathcal{F}$ -associated.

An example of a  $\mathcal{F}$ -associated sequence  $\{X_n, n \geq 1\}$  is obtained by defining  $X_n = Z + Y_n, n \geq 1$  where  $Z$  and  $Y_n, n \geq 1$  are  $\mathcal{F}$ -independent random variables as defined in Sect. 3. It can be shown by standard arguments that

$$\text{Cov}^{\mathcal{F}}(X, Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} H^{\mathcal{F}}(X, Y) dx dy \text{ a.s.}$$

where

$$H^{\mathcal{F}}(x, y) = E^{\mathcal{F}}[I_{(X \leq x, Y \leq y)}] - E^{\mathcal{F}}[I_{(X \leq x)}]E^{\mathcal{F}}[I_{(Y \leq y)}].$$

Let  $X$  and  $Y$  be  $\mathcal{F}$ -associated random variables. Suppose that  $f$  and  $g$  are almost everywhere differentiable functions such that  $\text{ess sup}_x |f'(x)| < \infty$ ,  $\text{ess sup}_x |g'(x)| < \infty$ ,  $E^{\mathcal{F}}[f^2(X)] < \infty$  a.s. and  $E^{\mathcal{F}}[g^2(Y)] < \infty$  a.s. where  $f'$  and  $g'$  denotes the derivatives of  $f$  and  $g$  respectively whenever they exist. Then it can be shown that

$$\text{Cov}^{\mathcal{F}}(f(X), g(Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f'(x)g'(y)H^{\mathcal{F}}(x, y)dx dy \quad \text{a.s.} \quad (66)$$

and hence

$$\left| \text{Cov}^{\mathcal{F}}(f(X), g(Y)) \right| \leq \text{ess sup}_x |f'(x)|\text{ess sup}_y |g'(y)|\text{Cov}^{\mathcal{F}}(X, Y) \quad \text{a.s.} \quad (67)$$

Proofs of these results can be obtained following the methods used for the study of *associated* random variables. As a consequence of these covariance inequalities, it should be possible to obtain a central limit theorem for conditionally associated sequences of random variables following the methods in Newman (1984). Note that  $\mathcal{F}$ -association does not imply association and vice versa. For results on associated random variables, see Prakasa Rao and Dewan (2001) and Roussas (1999).

### 10 Remarks

As it was pointed out earlier, conditional independence does not imply independence and vice versa. Hence one does have to derive limit theorems under conditioning if there is a need for such results even though the results and proofs of such results may be analogous to those under the non-conditioning set up. This was one of the reasons for developing results for conditional sequences in the earlier sections. We have given some examples of stochastic models where such results are useful such as exchangeable sequences and sequences of random variables defined on a homogeneous Markov chain discussed in Remark 3. Another example of a conditionally strong-mixing sequence is described in Sect. 8. A concrete example where conditional limit theorems are useful is in the study of statistical inference for non-ergodic models as discussed in Basawa and Prakasa Rao (1980) and Basawa and Scott (1983). For instance, if one wants to estimate the mean off-spring  $\theta$  for a Galton-Watson Branching process, the asymptotic properties of the maximum likelihood estimator depend on the set of non-extinction . Conditional limit theorems under a martingale set up have been used for the study of asymptotic properties for maximum likelihood estimators of parameters in a stochastic process framework where the Fisher information need not be additive and might not increase to infinity or might even increase to infinity at an exponential rate (cf. Basawa and Prakasa Rao 1980; Prakasa Rao 1999a,b; Guttorp 1991).

**Acknowledgment** The author thanks both the referees for their comments and suggestions.

## References

- Basawa, I. V., Prakasa Rao, B. L. S. (1980). *Statistical inference for stochastic processes*. London: Academic.
- Basawa, I. V., Scott, D. (1983). *Asymptotic optimal inference for non-ergodic models*. Lecture Notes in Statistics, Vol. 17. New York: Springer.
- Billingsley, P. (1986). *Probability and measure*. New York: Wiley.
- Chow, Y. S., Teicher, H. (1978). *Probability theory: independence, interchangeability, martingales*. New York: Springer.
- Chung, K. L., Erdős, P. (1952). On the application of the Borel-Cantelli lemma. *Transactions of American Mathematical Society*, 72, 179–186.
- Doob, J. L. (1953). *Stochastic processes*. New York: Wiley.
- Guttorp, P. (1991). *Statistical inference for branching processes*. New York: Wiley.
- Gyires, B. (1981). Linear forms in random variables defined on a homogeneous Markov chain, In P. Revesz et al. (Eds.), *The first pannonian symposium on mathematical statistics*, Lecture Notes in Statistics, Vol. 8. New York: Springer. (pp. 110–121).
- Hájek, J., Rényi, A. (1955). Generalization of an inequality of Kolmogorov. *Acta Mathematica Academiae Scientiarum Hungaricae*, 6, 281–283.
- Kochen, S., Stone, C. (1964). A note on the Borel-Cantelli lemma. *Illinois Journal of Mathematics*, 8, 248–251.
- Loève, M. (1977). *Probability theory I* (4th Ed.). New York: Springer.
- Majerak, D., Nowak, W., Zieba, W. (2005). Conditional strong law of large number. *International Journal of Pure and Applied Mathematics*, 20, 143–157.
- Newman, C. (1984). Asymptotic independence and limit theorems for positively and negatively dependent random variables. In Y. L. Tong (Ed.), *Inequalities in statistics and probability* (pp. 127–140). Hayward:IMS.
- Petrov, V. V. (2004). A generalization of the Borel-Cantelli lemma. *Statistics and Probability Letters*, 67, 233–239.
- Prakasa Rao, B. L. S. (1987). Characterization of probability measures by linear functions defined on a homogeneous Markov chain. *Sankhyā, Series A*, 49, 199–206.
- Prakasa Rao, B. L. S. (1990). On mixing for flows of  $\sigma$ -algebras, *Sankhyā. Series A*, 52, 1–15.
- Prakasa Rao, B. L. S. (1999a). *Statistical inference for diffusion type processes*. London: Arnold and New York: Oxford University Press
- Prakasa Rao, B. L. S. (1999b). *Semimartingales and their statistical inference*. Boca Raton: CRC Press.
- Prakasa Rao, B. L. S., Dewan, I. (2001). Associated sequences and related inference problems, In C. R. Rao, D. N. Shanbag (Eds.), *Handbook of statistics, 19, stochastic Processes: theory and methods* (pp. 693–728). Amsterdam: North Holland.
- Rosenblatt, M. (1956). A central limit theorem and a strong mixing condition. *Proceedings National Academy of Sciences U.S.A*, 42, 43–47.
- Roussas, G. G. (1999). Positive and negative dependence with some statistical applications. In S. Ghosh (Ed.), *Asymptotics, nonparametrics and time series* (pp. 757–788). New York: Marcel Dekker.
- Roussas, G. G., Ioannides, D. (1987). Moment inequalities for mixing sequences of random variables. *Stochastic Analysis and Applications*, 5, 61–120.
- Yan, J.A. (2004). A new proof of a generalized Borel-Cantelli lemma (Preprint), Academy of Mathematics and Systems Science. Chinese Academy of Sciences, Beijing.