CONDITIONAL CENTRAL LIMIT THEOREMS FOR A SEQUENCE OF CONDITIONAL INDEPENDENT RANDOM VARIABLES

De-Mei Yuan, Li-Ran Wei, and Lan Lei

ABSTRACT. A conditional version of the classical central limit theorem is derived rigorously by using conditional characteristic functions, and a more general version of conditional central limit theorem for the case of conditionally independent but not necessarily conditionally identically distributed random variables is established. These are done anticipating that the field of conditional limit theory will prove to be of significant applicability.

1. Introduction

Let $(\Omega, \mathcal{A}, P)$ be a probability space and let $\mathcal{F}$ be a sub-σ-algebra of $\mathcal{A}$. A finite sequence of random variables $\{X_k, 1 \leq k \leq n\}$ is said to be conditionally independent with respect to $\mathcal{F}$ ($\mathcal{F}$-independent, in short) if, for every $B_k \in \mathcal{B}$ (the Borel σ-algebra in $\mathbb{R}$),

$$(1.1) \qquad P^{\mathcal{F}} \left\{ \bigcap_{k=1}^{n} (X_k \in B_k) \right\} = \prod_{k=1}^{n} P^{\mathcal{F}} (X_k \in B_k) \text{ a.s.}$$

Here and in the sequel, $P^{\mathcal{F}}(A)$ denotes the conditional probability of an event $A \in \mathcal{A}$ relative to $\mathcal{F}$. An infinite sequence $\{X_n, n \geq 1\}$ is said to be $\mathcal{F}$-independent if every finite subsequence is $\mathcal{F}$-independent. Note that an equivalent condition to (1.1), which was proved by Roussas [8], is

$$P^{\mathcal{F}} \left\{ \bigcap_{k=1}^{n} (X_k \leq x_k) \right\} = \prod_{k=1}^{n} P^{\mathcal{F}} (X_k \leq x_k) \text{ a.s.}$$

for every $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$.

If $\mathcal{F} = \{\Omega, \emptyset\}$, then $\mathcal{F}$-independence reduces to the ordinary (unconditional) independence. In Prakasa Rao [7], concrete examples were given, where independent random variables lose their independence under conditioning, and dependent random variables become independent under conditioning.

The random nature of many problems arising in the applied sciences leads to mathematical models where conditioning is present. For example, martingale sequences are well-known cases of stochastic processes defined through conditioning. Markov processes are another example of stochastic processes in which conditioning (specifically, conditional independence) is essential. A more extensive enumeration of models such as statistical inference and engineering literature is given by Roussas [8] in which conditioning plays a key role.

A typical example of statistical application of conditional limit theorems is in the study of statistical inference for some branching processes, such as the Galton-Watson process (see, e.g. Basawa and Prakasa Rao [1]). Let $\{Z_0 = 1, \ Z_n, \ n \geq 1\}$ be a Galton-Watson process with mean offspring Θ. This process can be studied by means of the following autoregressive type model:

$$Z_{n+1} = \Theta Z_n + Z_n^{\frac{1}{2}} U_{n+1}, \ n \geq 0,$$

where $\{U_k, k \geq 1\}$ is the sequence of error random variables. In order to estimate the mean offspring Θ from a realization $\{Z_0 = 1, Z_1, \ldots, Z_n\}$, the maximum likelihood estimator of Θ is $\hat{\Theta}_n = \left(\sum_{k=1}^{n} Z_{k-1}\right)^{-1} \left(\sum_{k=1}^{n} Z_k\right)$, which coincides with the "least-square" estimator of Θ obtained by minimizing $\sum_{k=0}^{n} U_k^2$ with respect to Θ. The study of asymptotic properties of $\hat{\Theta}_n$ leads to a conditional limit theorem since, as it is detailed in Basawa and Prakasa Rao [1], these asymptotic properties of $\hat{\Theta}_n$ depend on the event of non-extinction of the process.

As it was pointed out by Prakasa Rao [7], 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 setup.

In the field of conditional limit theorems, in addition to the above example, a lot of efforts have been carried out. For example, Majerek et al. [5] proved a conditional version of the Kolmogorov strong law of large numbers, Prakasa Rao [7] obtained conditional versions of the generalized Borel-Cantelli lemma, generalized Kolmogorov inequality and generalized Hájek-Rényi inequality, Yuan et al. [10] extended many results from negative association to conditional negative association, Yuan and Yang [12] generalized many results from association to conditional association, Ordóñez Cabrera et al. [6] derived some results on conditional mean convergence theorem and conditional almost sure convergence, Yuan and Lei [11] established conditional results for conditionally strong mixing sequences, Yuan and Xie [13] studied conditionally linearly negative quadrant dependence, and all that. But all these efforts are not nearly

enough in comparison with what should be dedicated to the field of conditional limit theorems.

Next we pay our attention to a special case of limit theorems. Let $\{X_n, n \geq 1\}$ be a sequence of independent and identically distributed random variables having zero means and variance one, and let $S_n = \sum_{k=1}^{n} X_k$. The classical central limit theorem asserts that $S_n/\sqrt{n}$ converges in distribution to the standard normal distribution $\Phi(x)$. It may be of some interest, however, whether or not this assertion is possible when independence is replaced by conditional independence for some models described earlier. So we will restrict ourselves to conditional versions of the central limit theorem and anticipate that conditional central limit theorems will prove to be of significant applicability.

Throughout the remaining part of this paper, all events and random variables are defined on the same probability space $(\Omega, \mathcal{A}, P)$, $\mathcal{F}$ is a sub-σ-algebra of $\mathcal{A}$, $E^{\mathcal{F}} X$ denotes the conditional expectation of a random variable X relative to $\mathcal{F}$. In addition, we assume that the conditional distribution function F_X of X exists as regular conditional distribution, and this assumption is workable since the space of values of all random variables in this paper is the real line (cf. Shiryaev [9]).

2. Definitions and basic results on conditioning

We first recall the concept of conditional characteristic function, which was originally introduced by Loève [4] and generalized by Roussas [8] and Grzenda and Zieba [2] independently.

The conditional characteristic function of a random variable X with respect to $\mathcal{F}$ ($\mathcal{F}$-characteristic function, in short) is defined by

$$\varphi_{X,\mathcal{F}}(t) = E^{\mathcal{F}} e^{itX} \left(= E^{\mathcal{F}} \cos tX + i E^{\mathcal{F}} \sin tX\right) \text{ a.s., } t \in \mathbb{R}.$$

Similarly for the joint conditional characteristic function of any random vector $\underline{\mathbf{X}} = (X_1, X_2, \ldots, X_n)$. If there is no danger of confusion we will write simply $\varphi_{\mathcal{F}}(t)$ or $\varphi_{\mathcal{F}}(t_1, t_2, \ldots, t_n)$ without reference to X or $\underline{\mathbf{X}}$.

The statistical perspective of conditional characteristic function is that of a Bayesian. A problem begins with a parameter Θ with its prior probability distribution that exists only in mind of the investigator. The statistical model that is most commonly in use is that of a sequence $\{X_n, n \geq 1\}$ of observable random variables that is independent and identically distributed for each given value of Θ. In this case, we may express the (unconditional) characteristic function of $\underline{\mathbf{X}} = (X_1, X_2, \ldots, X_n)$ in terms of the conditional characteristic function of $\underline{\mathbf{X}}$ given Θ.

Let us consider a beta-Bernoulli process $\{X_n, n \geq 1\}$ with parameters $a > 0$ and $b > 0$. In this case, $\{X_n, n \geq 1\}$ is a sequence of conditional independent indicator random variables given Θ with

$$P(X_n = 1 | \Theta = \theta) = \theta, \ 0 < \theta < 1, \ n \geq 1,$$

where Θ is a beta random variable with left parameter a and right parameter b. Thus, Θ has a probability density function f given by

$$f(\theta) = \frac{1}{B(a,b)}\theta^{a-1}(1-\theta)^{b-1},\ 0 < \theta < 1.$$

Let $\mathcal{F} = \sigma(\Theta)$. Some simple calculations show that the $\mathcal{F}$-characteristic function corresponding to $(X_1, X_2, \ldots, X_n)$

$$\begin{aligned}
\varphi_{\mathcal{F}}(t_1, t_2, \ldots, t_n) &= \prod_{j=1}^{n} E^{\mathcal{F}} e^{it_j X_j} \\
&= \prod_{j=1}^{n}\left[(1-\Theta) + \Theta e^{it_j}\right] \\
&= (1-\Theta)^n + (1-\Theta)^{n-1}\Theta \sum_{j=1}^{n} e^{it_j} \\
&\quad + (1-\Theta)^{n-2}\Theta^2 \sum_{1\le j<k\le n} e^{i(t_j+t_k)} + \cdots + \Theta^n e^{i\sum_{j=1}^{n} t_j}\ \text{a.s.}
\end{aligned}$$

and consequently

$$\begin{aligned}
\varphi(t_1, t_2, \ldots, t_n) &= E\varphi_{\mathcal{F}}(t_1, t_2, \ldots, t_n) \\
&= \frac{1}{B(a,b)}\left[B(a,b+n) + B(a+1,b+n-1)\sum_{j=1}^{n} e^{it_j}\right. \\
&\quad + B(a+2,b+n-2)\sum_{1\le j<k\le n} e^{i(t_j+t_k)} \\
&\quad \left. + \cdots + B(a+n,b)e^{i\sum_{j=1}^{n} t_j}\right].
\end{aligned}$$

From the uniqueness theorem for characteristic functions, we have the finite dimensional distributions

$$P(X_1=x_1, X_2=x_2, \ldots, X_n=x_n) = \frac{B(a+k,b+n-k)}{B(a,b)} = \frac{a^{(1,k)}b^{(1,n-k)}}{(a+b)^{(1,n)}},$$

where $(x_1, x_2, \ldots, x_n) \in \{0,1\}^n$, $k = x_1 + x_2 + \cdots + x_n$, and $r^{(s,j)} = r(r+s)(r+2s)\cdots(r+(j-1)s)$. As usual, we adopt the convention that a product over an empty index set is 1. Hence $r^{(s,0)} = 1$ for every r and s.

Compared with the unconditional characteristic function, conditional characteristic function should also have many properties, here we establish the relation between the conditional characteristic function of a random variable and its conditional moments which will be used in subsequent sections.

Lemma 2.1. *Let $\varphi_{\mathcal{F}}(t)$ be the $\mathcal{F}$-characteristic function of X. If $E^{\mathcal{F}}|X|^n < \infty$ a.s. for some $n \geq 1$, then for every positive integer r satisfying $r \leq n$, $\varphi_{\mathcal{F}}^{(r)}(t)$ exists almost surely and*

$$(2.1) \qquad \varphi_{\mathcal{F}}^{(r)}(t) = E^{\mathcal{F}}\left[(iX)^r e^{itX}\right] \ \ a.s.,$$

$$(2.2) \qquad E^{\mathcal{F}} X^r = \frac{\varphi_{\mathcal{F}}^{(r)}(0)}{i^r} \ \ a.s.$$

and

$$(2.3) \qquad \varphi_{\mathcal{F}}(t) = \sum_{r=0}^{n} \frac{(it)^r}{r!} E^{\mathcal{F}} X^r + \frac{(it)^n}{n!} \varepsilon_n^{\mathcal{F}}(t) \ \ a.s.,$$

where $\left|\varepsilon_n^{\mathcal{F}}(t)\right| \leq 3 E^{\mathcal{F}}|X|^n$ a.s. and $\varepsilon_n^{\mathcal{F}}(t) \to 0$ a.s. as $t \to 0$.

Proof. Consider the difference quotient

$$\frac{\varphi_{\mathcal{F}}(t+h) - \varphi_{\mathcal{F}}(t)}{t} = E^{\mathcal{F}}\left[e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right].$$

Since $\left|\left(e^{ihx}-1\right)/h\right| \leq |x|$ and $E^{\mathcal{F}}|X| < \infty$ a.s., it follows from the conditionally dominated convergence theorem that the limit

$$\lim_{h \to 0} E^{\mathcal{F}}\left[e^{itX}\left(\frac{e^{ihX}-1}{h}\right)\right]$$

exists and equals

$$E^{\mathcal{F}}\left[e^{itX}\lim_{h \to 0}\left(\frac{e^{ihX}-1}{h}\right)\right] = E^{\mathcal{F}}\left(iXe^{itX}\right).$$

Hence $\varphi_{\mathcal{F}}'(t)$ exists and $\varphi_{\mathcal{F}}'(t) = E^{\mathcal{F}}\left(iXe^{itX}\right)$.

The existence of the derivatives $\varphi_{\mathcal{F}}^{(r)}(t)$, $1 < r \leq n$, and the validity of (2.1), follow by induction. Relation (2.2) follows immediately from (2.1). Let us now establish (2.3). Since

$$e^{iy} = \cos y + i \sin y = \sum_{k=0}^{n-1} \frac{(iy)^k}{k!} + \frac{(iy)^n}{n!}\left[\cos\theta_1 y + i\sin\theta_2 y\right]$$

for real y, with $|\theta_1| \leq 1$ and $|\theta_2| \leq 1$, we have

$$e^{itX} = \sum_{k=0}^{n-1} \frac{(itX)^k}{k!} + \frac{(itX)^n}{n!}\left[\cos(\theta_1(\omega)tX) + i\sin(\theta_2(\omega)tX)\right]$$

and

$$E^{\mathcal{F}} e^{itX} = \sum_{k=0}^{n} \frac{(it)^k}{k!} E^{\mathcal{F}} X^k + \frac{(it)^n}{n!} \varepsilon_n^{\mathcal{F}}(t),$$

where

$$\varepsilon_n^{\mathcal{F}}(t) = E^{\mathcal{F}}\left\{X^n\left[\cos(\theta_1(\omega)tX) + i\sin(\theta_2(\omega)tX) - 1\right]\right\}.$$

It is clear that $\left|\varepsilon_n^{\mathcal{F}}(t)\right| \leq 3E^{\mathcal{F}}|X|^n$ and the conditionally dominated convergence theorem applies and gives $\varepsilon_n^{\mathcal{F}}(t) \to 0$ as $t \to 0$. $\square$

Next we discuss the concept of conditionally identical distribution, which was proposed by Majerek et al. [5].

Two random variables X and Y are said to be conditionally identically distributed with respect to $\mathcal{F}$ ($\mathcal{F}$-identically distributed, in short) if

$$P^{\mathcal{F}}(X \in B) = P^{\mathcal{F}}(Y \in B) \text{ a.s. for all } B \in \mathcal{B}.$$

An equivalent statement concerning the concept of conditional identical distribution is the following.

Lemma 2.2. *Two random variables X and Y are $\mathcal{F}$-identically distributed if and only if*

$$P^{\mathcal{F}}(X \leq a) = P^{\mathcal{F}}(Y \leq a) \text{ a.s. for all } a \in \mathbb{R}.$$

Proof. We need only to prove sufficiency. Let $\mathcal{C} = \{(-\infty, a] : a \in \mathbb{R}\}$ be a class of sets. Obviously, $\mathcal{C}$ is closed under intersections. Let

$$\mathcal{D} = \left\{B \in \mathcal{B} : P^{\mathcal{F}}(X \in B) = P^{\mathcal{F}}(Y \in B)\right\}$$

be another class of sets, then $\mathcal{D}$ is a d-system and $\mathcal{D} \supset \mathcal{C}$. By Shiryaev [9, Chapter 2, Section 2, Theorem 2],

$$\mathcal{D} \supset d(\mathcal{C}) = \sigma(\mathcal{C}) = \mathcal{B},$$

where $d(\mathcal{C})$ denotes the smallest d-system containing $\mathcal{C}$. $\square$

If $\mathcal{F} = \{\Omega, \emptyset\}$, then $\mathcal{F}$-identical distribution turns into the usual identical distribution. Furthermore, if X and Y are $\mathcal{F}$-identically distributed, then by a simple property of conditional expectation

$$P(X \in B) = E\left[P^{\mathcal{F}}(X \in B)\right] = E\left[P^{\mathcal{F}}(Y \in B)\right] = P(Y \in B)$$

for all $B \in \mathcal{B}$, showing that X and Y are identically distributed.

Remark 2.3. Identically distributed random variables need not always be conditionally identically distributed.

To illustrate this, we give a counter-example. Let $\Omega = \{1, 2, 3, 4, 5, 6\}$ and let $p_i = 1/6$ be the probability assigned to the event $\{i\}$. If events A_1, A_2 are defined by $A_1 = \{1, 2\}$, $A_2 = \{3, 4\}$ and random variables X, Y are defined by $X = I_{A_1} + 2I_{A_2}$, $Y = 2I_{A_1} + I_{A_2}$, where I_A denotes the indicator function of an event A, then X and Y are identically distributed.

Suppose that $B = \{4, 5\}$ and $\mathcal{F} = \{\Omega, B, B^c, \emptyset\}$ is the sub-σ-algebra generated by B, some simple calculations show that

$$P^{\mathcal{F}}(X \leq 1) = \left\{ \begin{array}{l} P(X \leq 1 | B), \ \omega \in B \\ P(X \leq 1 | B^c), \ \omega \in B^c \end{array} \right. = \left\{ \begin{array}{l} 1/2, \ \omega \in B, \\ 3/4, \ \omega \in B^c \end{array} \right.$$

and

$$P^{\mathcal{F}}(Y \leq 1) = \left\{ \begin{array}{l} P(Y \leq 1 | B), \ \omega \in B \\ P(Y \leq 1 | B^c), \ \omega \in B^c \end{array} \right. = \left\{ \begin{array}{l} 1, \ \omega \in B, \\ 1/2, \ \omega \in B^c. \end{array} \right.$$

So that X and Y are not $\mathcal{F}$-identically distributed by Lemma 2.2.

We now give a result on conditional independence of random variables.

Lemma 2.4. *Let X and Y be $\mathcal{F}$-independent random variables, and let ξ and η be $\mathcal{F}$-measurable random variables. Then for arbitrary $\mathcal{B}^2$-measurable functions $f(x,y)$ and $g(x,y)$,*

$$(2.4) \qquad E^{\mathcal{F}}[f(X,\xi)g(Y,\eta)] = E^{\mathcal{F}}f(X,\xi) \cdot E^{\mathcal{F}}g(Y,\eta) \ \ a.s.,$$

and similarly for any finite number of random variables.

Proof. First suppose that $f = I_{A_1 \times B_1}$, $g = I_{A_2 \times B_2}$, where $A_i, B_i \in \mathcal{B}$, $i = 1, 2$, then (2.4) follows by assumptions. Let $\mathcal{C} = \{A \times B : A \in \mathcal{B}, \ B \in \mathcal{B}\}$ be a class of sets. Obviously, $\mathcal{C}$ is closed under intersections. For fixed $A_2, B_2 \in \mathcal{B}$, define

$$\mathcal{D}_1 = \{C_1 : C_1 \in \mathcal{B}^2, E^{\mathcal{F}} I_{C_1}(X_1, \xi_1) I_{A_2 \times B_2}(X_2, \xi_2)$$
$$= E^{\mathcal{F}} I_{C_1}(X_1, \xi_1) \cdot E^{\mathcal{F}} I_{A_2 \times B_2}(X_2, \xi_2)\},$$

then $\mathcal{D}_1$ is a d-system and $\mathcal{D}_1 \supset \mathcal{C}$ from what has just now been proved. By Shiryaev [9, Chapter 2, Section 2, Theorem 2],

$$\mathcal{D}_1 \supset d(\mathcal{C}) = \sigma(\mathcal{C}) = \mathcal{B}^2.$$

For fixed $C_1 \in \mathcal{B}^2$, define

$$\mathcal{D}_2 = \left\{ C_2 : C_2 \in \mathcal{B}^2, E^{\mathcal{F}} I_{C_1}(X, \xi) I_{C_2}(Y, \eta) = E^{\mathcal{F}} I_{C_1}(X, \xi) \cdot E^{\mathcal{F}} I_{C_2}(Y, \eta) \right\},$$

then $\mathcal{D}_2$ is a d-system and $\mathcal{D}_2 \supset \mathcal{C}$ from what has just now been proved. Again, by Shiryaev [9, Chapter 2, Section 2, Theorem 2],

$$\mathcal{D}_2 \supset d(\mathcal{C}) = \sigma(\mathcal{C}) = \mathcal{B}^2.$$

Up to now, we have proved that

$$(2.5) \qquad E^{\mathcal{F}} I_{C_1}(X, \xi) I_{C_2}(Y, \eta) = E^{\mathcal{F}} I_{C_1}(X, \xi) \cdot E^{\mathcal{F}} I_{C_2}(Y, \eta)$$

for all $C_1, C_2 \in \mathcal{B}^2$.

Next suppose that $f \geq 0$ and $g \geq 0$. Then there exist nondecreasing sequences $\{f_n\}$ and $\{g_n\}$ of nonnegative $\mathcal{B}^2$-measurable simple functions such that $0 \leq f_n \uparrow f$ and $0 \leq g_n \uparrow g$. In this case equality (2.4) follows for $f, g \geq 0$ from (2.5) by using the conditional monotone convergence theorem.

Finally, in the general case relation (2.4) is proved by setting $f = f^+ - f^-$ and $g = g^+ - g^-$ and applying what has just now been proved separately to $f^+ g^+$, $f^+ g^-$, $f^- g^+$ and $f^- g^-$. $\qquad \square$

Using the basic method appearing in the proof of Lemma 2.4, we can prove:

Lemma 2.5. *Let X and Y be $\mathcal{F}$-identically distributed random variables, and let ξ be an $\mathcal{F}$-measurable random variable. Then for arbitrary $\mathcal{B}^2$-measurable function $f(x, y)$,*

$$E^{\mathcal{F}} f(X, \xi) = E^{\mathcal{F}} f(Y, \xi) \quad a.s.,$$

and similarly for any finite number of random variables.

3. Conditional version of the classical central limit theorem

The following conditional version of the classical central limit theorem has been stated in Prakasa Rao [7] without a proof. Although Grzenda and Zieba [2] gave a proof, it was not completely stringent. To overcome these, here we provide a rigorous proof based on lemmas in the previous section. A more general result will be postponed until Section 4.

Theorem 3.1. *Let $\{X_n, n \geq 1\}$ be a sequence of $\mathcal{F}$-independent and $\mathcal{F}$-identically distributed random variables with $\sigma_{\mathcal{F}}^2 = E^{\mathcal{F}} \left(X_1 - E^{\mathcal{F}} X_1 \right)^2 < \infty$ a.s. Then*

$$(3.1) \qquad E^{\mathcal{F}} \exp \left(it \frac{S_n - E^{\mathcal{F}} S_n}{\sqrt{n} \sigma_{\mathcal{F}}} \right) \to e^{-\frac{t^2}{2}} \quad a.s.$$

as $n \to \infty$ for every $t \in \mathbb{R}$. In particular,

$$(3.2) \qquad \frac{S_n - E^{\mathcal{F}} S_n}{\sqrt{n} \sigma_{\mathcal{F}}} \to N(0, 1) \quad in \; distribution.$$

Proof. Relation (3.2) follows from (3.1) by using the dominated convergence theorem and the continuity theorem for characteristic functions. So we need only to prove (3.1). In view of the $\mathcal{F}$-independence of $X_1, X_2, \ldots, X_n$ and Lemma 2.4, we conclude that

$$
\begin{aligned}
E^{\mathcal{F}} \exp \left(it \frac{S_n - E^{\mathcal{F}} S_n}{\sqrt{n} \sigma_{\mathcal{F}}} \right) &= E^{\mathcal{F}} \left\{ \prod_{k=1}^{n} \exp \left[\frac{it}{\sqrt{n} \sigma_{\mathcal{F}}} \left(X_k - E^{\mathcal{F}} X_k \right) \right] \right\} \\
&= \prod_{k=1}^{n} E^{\mathcal{F}} \exp \left[\frac{it}{\sqrt{n} \sigma_{\mathcal{F}}} \left(X_k - E^{\mathcal{F}} X_k \right) \right].
\end{aligned}
$$

In view of $\mathcal{F}$-identical distribution and Lemma 2.5, $E^{\mathcal{F}} X_1 = E^{\mathcal{F}} X_2 = \cdots = E^{\mathcal{F}} X_n$ a.s. and $E^{\mathcal{F}} \exp \left(\frac{it X_1}{\sqrt{n} \sigma_{\mathcal{F}}} \right) = E^{\mathcal{F}} \exp \left(\frac{it X_2}{\sqrt{n} \sigma_{\mathcal{F}}} \right) = \cdots = E^{\mathcal{F}} \exp \left(\frac{it X_n}{\sqrt{n} \sigma_{\mathcal{F}}} \right)$ a.s., so that

$$
\begin{aligned}
& \prod_{k=1}^{n} E^{\mathcal{F}} \exp \left[\frac{it}{\sqrt{n} \sigma_{\mathcal{F}}} \left(X_k - E^{\mathcal{F}} X_k \right) \right] \\
&= \prod_{k=1}^{n} E^{\mathcal{F}} \left[\exp \left(\frac{it X_k}{\sqrt{n} \sigma_{\mathcal{F}}} \right) \cdot \exp \left(-\frac{it E^{\mathcal{F}} X_k}{\sqrt{n} \sigma_{\mathcal{F}}} \right) \right] \\
&= \prod_{k=1}^{n} \left[E^{\mathcal{F}} \exp \left(\frac{it X_k}{\sqrt{n} \sigma_{\mathcal{F}}} \right) \cdot \exp \left(-\frac{it E^{\mathcal{F}} X_k}{\sqrt{n} \sigma_{\mathcal{F}}} \right) \right]
\end{aligned}
$$

$$
\begin{aligned}
&= \left[E^{\mathcal{F}} \exp\left(\frac{itX_1}{\sqrt{n}\sigma_{\mathcal{F}}} \right) \cdot \exp\left(-\frac{itE^{\mathcal{F}}X_1}{\sqrt{n}\sigma_{\mathcal{F}}} \right) \right]^n \\
&= \left[E^{\mathcal{F}} \exp\left(\frac{it}{\sqrt{n}} \frac{X_1 - E^{\mathcal{F}}X_1}{\sigma_{\mathcal{F}}} \right) \right]^n \\
&= \left[\varphi_{\mathcal{F}}\left(\frac{t}{\sqrt{n}} \right) \right]^n,
\end{aligned}
$$

where $\varphi_{\mathcal{F}}(t)$ is the $\mathcal{F}$-characteristic function corresponding to $\left(X_1 - E^{\mathcal{F}}X_1 \right)/\sigma_{\mathcal{F}}$. However,

$$
\varphi_{\mathcal{F}}\left(\frac{t}{\sqrt{n}} \right) = 1 - \frac{t^2}{2n} - \frac{t^2}{2n}\varepsilon_2^{\mathcal{F}}\left(\frac{t}{\sqrt{n}} \right)
$$

by Lemma 2.1, where $\varepsilon_2^{\mathcal{F}}(t)$ satisfies $\varepsilon_2^{\mathcal{F}}(t/n) \to 0$ a.s. as $n \to \infty$ for every $t \in R$, and therefore

$$
E^{\mathcal{F}}\left[\exp\left(it\frac{S_n - E^{\mathcal{F}}S_n}{\sqrt{n}\sigma_{\mathcal{F}}} \right) \right] = \left[1 - \frac{t^2}{2n} - \frac{t^2}{2n}\varepsilon_2^{\mathcal{F}}\left(\frac{t}{\sqrt{n}} \right) \right]^n \to e^{-\frac{t^2}{2}} \text{ a.s.}
$$

as $n \to \infty$ for fixed $t \in \mathbb{R}$. $\qquad\square$

4. Conditional version of the general central limit theorem

This section is in fact a continuation of the previous one as we consider here a more general version of conditional central limit theorem for the case of $\mathcal{F}$-independent but not necessarily $\mathcal{F}$-identically distributed random variables.

Theorem 4.1. *Let $\{X_n, n \geq 1\}$ be a sequence of $\mathcal{F}$-independent but not necessarily $\mathcal{F}$-identically distributed random variables with $\sigma_{n,\mathcal{F}}^2 = E^{\mathcal{F}}\left(X_n - E^{\mathcal{F}}X_n \right)^2 < \infty$ a.s. for every $n \geq 1$. Define $B_{n,\mathcal{F}}^2 = E^{\mathcal{F}}(S_n - E^{\mathcal{F}}S_n)^2$. Then the following conditions:*

(i) $\lim\limits_{n\to\infty} \max\limits_{1 \leq k \leq n} \frac{\sigma_{k,\mathcal{F}}^2}{B_{n,\mathcal{F}}^2} = 0$ a.s.,

(ii) $E^{\mathcal{F}} \exp\left(it\frac{S_n - E^{\mathcal{F}}S_n}{B_{n,\mathcal{F}}} \right) \to e^{-\frac{t^2}{2}}$ as $n \to \infty$

hold if and only if for every $\varepsilon > 0$ the $\mathcal{F}$-Lindeberg condition
(4.1)
$$
\lim_{n\to\infty} \frac{1}{B_{n,\mathcal{F}}^2} \sum_{k=1}^n E^{\mathcal{F}}\left[\left(X_k - E^{\mathcal{F}}X_k \right)^2 I\left(\left| X_k - E^{\mathcal{F}}X_k \right| > \varepsilon B_{n,\mathcal{F}} \right) \right] = 0 \text{ a.s.}
$$

is satisfied.

Proof. First we will prove the sufficiency. For every $n \geq 1$ we set

$$
X_{nk} = \frac{X_k - E^{\mathcal{F}}X_k}{B_{n,\mathcal{F}}}, \ 1 \leq k \leq n,
$$

then $E^{\mathcal{F}} X_{nk} = 0$ and $\sum_{k=1}^{n} E^{\mathcal{F}} X_{nk}^2 = 1$ a.s. We see easily that (4.1) is equivalent to

$$(4.2) \qquad \lim_{n \to \infty} \sum_{k=1}^{n} E^{\mathcal{F}} \left[X_{nk}^2 I \left(|X_{nk}| > \varepsilon \right) \right] = 0 \text{ a.s.}$$

Now, for $n \geq 1$ and $1 \leq k \leq n$, we can write

$$\begin{aligned}
\frac{\sigma_{n,\mathcal{F}}^2}{B_{n,\mathcal{F}}^2} &= E^{\mathcal{F}} X_{nk}^2 \\
&= E^{\mathcal{F}} \left[X_{nk}^2 I \left(|X_{nk}| \leq \varepsilon \right) \right] + E^{\mathcal{F}} \left[X_{nk}^2 I \left(|X_{nk}| > \varepsilon \right) \right] \\
&\leq \varepsilon^2 + E^{\mathcal{F}} \left[X_{nk}^2 I \left(|X_{nk}| > \varepsilon \right) \right],
\end{aligned}$$

so that

$$\max_{1 \leq k \leq n} \frac{\sigma_{n,\mathcal{F}}^2}{B_{n,\mathcal{F}}^2} \leq \varepsilon^2 + \sum_{k=1}^{n} E^{\mathcal{F}} \left[X_{nk}^2 I \left(|X_{nk}| > \varepsilon \right) \right].$$

In view of (4.2) we see that condition (i) holds.

For the proof of (ii), let $\varphi_{nk,\mathcal{F}}(t)$ be the $\mathcal{F}$-characteristic function of X_{nk} and let $\varphi_{n,\mathcal{F}}(t)$ be the $\mathcal{F}$-characteristic function of $\sum_{k=1}^{n} X_{nk} = \left(S_n - E^{\mathcal{F}} S_n \right) / B_{n,\mathcal{F}}$. By Lemma 2.4,

$$\varphi_{n,\mathcal{F}}(t) = \prod_{k=1}^{n} \varphi_{nk,\mathcal{F}}(t), \ t \in \mathbb{R}.$$

We choose a $t \in \mathbb{R}$ and suppose that it is fixed throughout the rest of the proof. We first prove that

$$(4.3) \qquad \lim_{n \to \infty} \max_{1 \leq k \leq n} |\varphi_{nk,\mathcal{F}}(t) - 1| = 0 \text{ a.s.}$$

Note that

$$\varphi_{nk,\mathcal{F}}(t) - 1 = E^{\mathcal{F}} \left(e^{itX_{nk}} - 1 - itX_{nk} \right)$$

and

$$\left| e^{it} - 1 - it \right| \leq \frac{t^2}{2},$$

so that

$$(4.4) \quad |\varphi_{nk,\mathcal{F}}(t) - 1| \leq E^{\mathcal{F}} \left| e^{itX_{nk}} - 1 - itX_{nk} \right| \leq \frac{t^2}{2} E^{\mathcal{F}} X_{nk}^2 = \frac{t^2}{2} \frac{\sigma_{k,\mathcal{F}}^2}{B_{n,\mathcal{F}}^2},$$

and therefore

$$\max_{1 \leq k \leq n} |\varphi_{nk,\mathcal{F}}(t) - 1| \leq \frac{t^2}{2} \max_{1 \leq k \leq n} \frac{\sigma_{k,\mathcal{F}}^2}{B_{n,\mathcal{F}}^2}.$$

In view of (i), relation (4.3) holds.

Next we show that

$$(4.5) \qquad \lim_{n \to \infty} \left\{ \ln \varphi_{n,\mathcal{F}}(t) - \sum_{k=1}^{n} \left[\varphi_{nk,\mathcal{F}}(t) - 1 \right] \right\} = 0 \text{ a.s.}$$

From (4.3) we can choose n sufficiently large so that

$$|\varphi_{nk,\,\mathcal{F}}(t) - 1| \leq \frac{1}{2} \text{ a.s.}$$

for all $1 \leq k \leq n$. Then we have the expansion

$$\ln \varphi_{n,\,\mathcal{F}}(t) = \sum_{k=1}^{n} \ln \varphi_{nk,\,\mathcal{F}}(t) = \sum_{k=1}^{n} [\varphi_{nk,\,\mathcal{F}}(t) - 1] + R_{n,\,\mathcal{F}}(t),$$

where

$$R_{n,\,\mathcal{F}}(t) = \sum_{k=1}^{n} \sum_{j=2}^{\infty} \frac{(-1)^{j-1}}{j} [\varphi_{nk,\mathcal{F}}(t) - 1]^{j}.$$

For n sufficiently large,

$$
\begin{aligned}
|R_{n,\,\mathcal{F}}| &\leq \sum_{k=1}^{n} \sum_{j=2}^{\infty} \frac{|\varphi_{nk,\,\mathcal{F}}(t) - 1|^{j}}{j} \\
&\leq \frac{1}{2} \sum_{k=1}^{n} \frac{|\varphi_{nk,\,\mathcal{F}}(t) - 1|^{2}}{1 - |\varphi_{nk,\,\mathcal{F}}(t) - 1|} \\
&\leq \sum_{k=1}^{n} |\varphi_{nk,\,\mathcal{F}}(t) - 1|^{2} \\
&\leq \max_{1 \leq k \leq n} |\varphi_{nk,\,\mathcal{F}}(t) - 1| \cdot \sum_{k=1}^{n} |\varphi_{nk,\,\mathcal{F}}(t) - 1|.
\end{aligned}
$$

From (4.4),

$$\sum_{k=1}^{n} |\varphi_{nk,\,\mathcal{F}}(t) - 1| \leq \frac{t^2}{2} \sum_{k=1}^{n} \frac{\sigma_{k,\mathcal{F}}^2}{B_{n,\,\mathcal{F}}^2} = \frac{t^2}{2},$$

so that

$$|R_{n,\,\mathcal{F}}(t)| \leq \frac{t^2}{2} \max_{1 \leq k \leq n} |\varphi_{nk,\,\mathcal{F}}(t) - 1|,$$

and (4.5) follows easily.

We now return to the proof of condition (ii) and write

$$\sum_{k=1}^{n} [\varphi_{nk,\,\mathcal{F}}(t) - 1] = -\frac{t^2}{2} + \rho_{n}^{\mathcal{F}}(t),$$

where

$$\rho_{n}^{\mathcal{F}}(t) = \frac{t^2}{2} + \sum_{k=1}^{n} E^{\mathcal{F}} \left(e^{itX_{nk}} - 1 - itX_{nk} \right).$$

Let $\varepsilon > 0$. Since

$$\sum_{k=1}^{n} E^{\mathcal{F}} X_{nk}^2 = 1 \text{ a.s.,}$$

we can rewrite $\rho_n^{\mathcal{F}}(t)$ as

$$\rho_n^{\mathcal{F}}(t) = \sum_{k=1}^{n} E^{\mathcal{F}} \left\{ \left[e^{itX_{nk}} - 1 - itX_{nk} - \frac{1}{2}(itX_{nk})^2 \right] I\left(|X_{nk}| \leq \varepsilon\right) \right\}$$
$$+ \sum_{k=1}^{n} E^{\mathcal{F}} \left\{ \left[e^{itX_{nk}} - 1 - itX_{nk} - \frac{1}{2}(itX_{nk})^2 \right] I\left(|X_{nk}| > \varepsilon\right) \right\},$$

so that

$$\left| \rho_n^{\mathcal{F}}(t) \right|$$
$$= \frac{|t|^3}{6} \sum_{k=1}^{n} E^{\mathcal{F}} \left[|X_{nk}|^3 I\left(|X_{nk}| \leq \varepsilon\right) \right] + t^2 \sum_{k=1}^{n} E^{\mathcal{F}} \left[|X_{nk}|^2 I\left(|X_{nk}| > \varepsilon\right) \right]$$
$$\leq \frac{|t|^3 \varepsilon}{6} \sum_{k=1}^{n} E^{\mathcal{F}} \left[X_{nk}^2 I\left(|X_{nk}| \leq \varepsilon\right) \right] + t^2 \sum_{k=1}^{n} E^{\mathcal{F}} \left[X_{nk}^2 I\left(|X_{nk}| > \varepsilon\right) \right]$$
$$= \frac{|t|^3 \varepsilon}{6} + t^2 \left(1 - \frac{|t| \varepsilon}{6} \right) \sum_{k=1}^{n} E^{\mathcal{F}} \left[|X_{nk}|^2 I\left(|X_{nk}| > \varepsilon\right) \right].$$

Here, we have used a inequality $\left| e^{it} - 1 - it - (it)^2 \big/ 2! \right| \leq |t|^3 \big/ 3!$, which can be derived from Lemma 5.14 in Kallenberg [3]. Using (4.2), we conclude that $\rho_n^{\mathcal{F}}(t) \to 0$ a.s. as $n \to \infty$. Finally, we see that $\varphi_n^{\mathcal{F}}(t) \to e^{-t^2/2}$ a.s. as $n \to \infty$ by using (4.5) and complete the proof of (ii).

We now turn to the proof of the necessity. (4.3) holds from condition (i), and consequently (4.5) holds. On the other hand, we conclude that $\ln \varphi_{n,\mathcal{F}}(t) \to -t^2/2$ a.s. as $n \to \infty$ from condition (ii). Combining the two results, we obtain

$$\sum_{k=1}^{n} [\varphi_{nk,\mathcal{F}}(t) - 1] \to -\frac{t^2}{2} + o(1) \text{ a.s. as } n \to \infty.$$

Taking the real part on both sides, we obtain

$$\frac{t^2}{2} - \sum_{k=1}^{n} E^{\mathcal{F}} (1 - \cos tX_{nk}) = o(1) \text{ a.s. as } n \to \infty.$$

Let $\varepsilon > 0$. Then we can rewrite the last relation as

(4.6)
$$\frac{t^2}{2} - \sum_{k=1}^{n} E^{\mathcal{F}} (1 - \cos tX_{nk}) I\left(|X_{nk}| \leq \varepsilon\right)$$
$$= \sum_{k=1}^{n} E^{\mathcal{F}} (1 - \cos tX_{nk}) I\left(|X_{nk}| > \varepsilon\right) + o(1) \text{ a.s. as } n \to \infty.$$

Clearly

$$\sum_{k=1}^{n} E^{\mathcal{F}} (1 - \cos tX_{nk}) I\left(|X_{nk}| \leq \varepsilon\right) \leq \frac{t^2}{2} \sum_{k=1}^{n} E^{\mathcal{F}} X_{nk}^2 I\left(|X_{nk}| \leq \varepsilon\right)$$

$$= \frac{t^2}{2}\left[1 - \sum_{k=1}^{n} E^{\mathcal{F}} X_{nk}^2 I\left(|X_{nk}| > \varepsilon\right)\right],$$

so that

(4.7) $\quad \dfrac{t^2}{2} - \displaystyle\sum_{k=1}^{n} E^{\mathcal{F}}\left(1 - \cos tX_{nk}\right) I\left(|X_{nk}| \le \varepsilon\right) \ge \dfrac{t^2}{2}\sum_{k=1}^{n} E^{\mathcal{F}} X_{nk}^2 I\left(|X_{nk}| > \varepsilon\right).$

On the other hand,

(4.8) $\quad \displaystyle\sum_{k=1}^{n} E^{\mathcal{F}}\left(1 - \cos tX_{nk}\right) I\left(|X_{nk}| > \varepsilon\right) \le 2\sum_{k=1}^{n} E^{\mathcal{F}} I\left(|X_{nk}| > \varepsilon\right)$

$$\le \frac{2}{\varepsilon^2}\sum_{k=1}^{n} E^{\mathcal{F}} X_{nk}^2 I\left(|X_{nk}| > \varepsilon\right)$$

$$\le \frac{2}{\varepsilon^2}.$$

Combining (4.6), (4.7) and (4.8), we get

$$0 \le \sum_{k=1}^{n} E^{\mathcal{F}} X_{nk}^2 I\left(|X_{nk}| > \varepsilon\right) \le \frac{2}{t^2}\left[\frac{2}{\varepsilon^2} + o\left(1\right)\right] \text{ a.s. as } n \to \infty.$$

Taking the limits on both sides, first as $n \to \infty$ and then as $|t| \to \infty$, we see that the $\mathcal{F}$-Lindeberg condition (4.1) is satisfied. $\qquad\square$

Remark 4.2. In Theorem 4.1, condition (ii) implies

$$\frac{S_n - E^{\mathcal{F}} S_n}{B_{n,\,\mathcal{F}}} \to N\left(0, 1\right) \text{ in distribution as } n \to \infty.$$

Proof. This is because of the continuity theorem and the observation

$$\left|E \exp\left(it\frac{S_n - E^{\mathcal{F}} S_n}{B_{n,\,\mathcal{F}}}\right) - e^{-\frac{t^2}{2}}\right|$$

$$= \left|E\left[E^{\mathcal{F}} \exp\left(it\frac{S_n - E^{\mathcal{F}} S_n}{B_{n,\,\mathcal{F}}}\right) - e^{-\frac{t^2}{2}}\right]\right|$$

$$\le E\left\{\left|E^{\mathcal{F}} \exp\left(it\frac{S_n - E^{\mathcal{F}} S_n}{B_{n,\,\mathcal{F}}}\right) - e^{-\frac{t^2}{2}}\right|\right\}. \qquad\square$$

Remark 4.3. Theorem 3.1 is a special case to Theorem 4.1.

Proof. Suppose that $\{X, X_n, n \ge 1\}$ is a sequence of $\mathcal{F}$-identical distributed random variables with $\sigma_{\mathcal{F}}^2 = E^{\mathcal{F}}\left(X - E^{\mathcal{F}} X\right)^2 < \infty$ a.s. Since

$$\sum_{k=1}^{\infty} P^{\mathcal{F}}\left(\left|\frac{X - E^{\mathcal{F}} X}{\sigma_{\mathcal{F}}}\right| > k\right) \le E^{\mathcal{F}}\left|\frac{X - E^{\mathcal{F}} X}{\sigma_{\mathcal{F}}}\right| \le 1 \text{ a.s.}$$

implies

$$P^{\mathcal{F}}\left(\left|\frac{X - E^{\mathcal{F}}X}{\sigma_{\mathcal{F}}}\right| > k\right) \to 0 \text{ a.s. as } k \to \infty,$$

and consequently

$$P\left(\left|\frac{X - E^{\mathcal{F}}X}{\sigma_{\mathcal{F}}}\right| > k\right) \to 0 \text{ as } k \to \infty.$$

We conclude that from Lemma 2.5 and the conditional monotone convergence theorem,

$$\lim_{n \to \infty} \frac{1}{B_{n,\mathcal{F}}^2} \sum_{k=1}^{n} E^{\mathcal{F}}\left[\left(X_k - E^{\mathcal{F}}X_k\right)^2 I\left(\left|X_k - E^{\mathcal{F}}X_k\right| > \varepsilon B_{n,\mathcal{F}}\right)\right]$$

$$= \lim_{n \to \infty} E^{\mathcal{F}}\left[\left(\frac{X - E^{\mathcal{F}}X}{\sigma_{\mathcal{F}}}\right)^2 I\left(\left|\frac{X - E^{\mathcal{F}}X}{\sigma_{\mathcal{F}}}\right| > \sqrt{n}\varepsilon\right)\right] = 0.$$

That is, the $\mathcal{F}$-Lindeberg condition (4.1) is satisfied. $\square$

References

[1] V. Basawa and B. L. S. Prakasa Rao, *Statistical Inference for Stochastic Processes*, London, Academic press, 1980.

[2] W. Grzenda and W.Zieba, *Conditional central limit theorems*, Int. Math. Forum **3** (2008), no. 29-32, 1521–1528.

[3] O. Kallenberg, *Foundations of Modern Probability*, 2nd Edition, Now York, Springer-Verlag, 2002.

[4] M. Loève, *Probability Theory II*, 4th Edition, Now York, Springer-Verlag, 1978.

[5] D. Majerek, W. Nowak, and W. Zieba, *Conditional strong law of large number*, Int. J. Pure Appl. Math. **20** (2005), no. 2, 143–157.

[6] M. Ordóñez Cabrera, A. Rosalsky, and A. Volodin, *Some theorems on conditional mean convergence and conditional almost sure convergence for randomly weighted sums of dependent random variables*, TEST **21** (2012), no. 2, 369–385.

[7] B. L. S. Prakasa Rao, *Conditional independence, conditional mixing and conditional association*, Ann. Inst. Statist. Math. **61** (2009), no. 2, 441–460.

[8] G. G. Roussas, *On conditional independence, mixing, and association*, Stoch. Anal. Appl. **26** (2008), no. 6, 1274–1309.

[9] A. N. Shiryaev, *Probability*, 2nd Edition, Now York, Springer-Verlag, 1995.

[10] D. M. Yuan, J. An, and X. S. Wu, *Conditional limit theorems for conditionally negatively associated random variables*, Monatsh. Math. **161** (2010), no. 4, 449–473.

[11] D. M. Yuan and L. Lei, *Some conditional results for conditionally strong mixing sequences of random variables*, Sci. China Math. **56** (2013), no. 4, 845–859.

[12] D. M. Yuan and Y. K. Yang, *Conditional versions of limit theorems for conditionally associated random variables*, J. Math. Anal. Appl. **376** (2011), no. 1, 282–293.

[13] D. M. Yuan and Y. Xie, *Conditional limit theorems for conditionally linearly negative quadrant dependent random variables*, Monatsh. Math. **166** (2012), no. 2, 281–299.

De-Mei Yuan
School of Mathematics and Statistics
Chongqing Technology and Business University
Chongqing 400067, P. R. China
E-mail address: yuandemei@163.com

Li-Ran Wei
College of Mathematics and Computer Science
Yangtze Normal University
Fuling 408100, P. R. China
E-mail address: 517109353@qq.com

Lan Lei
School of Mathematics and Statistics
Chongqing Technology and Business University
Chongqing 400067, P. R. China
E-mail address: leilan@ctbu.edu.cn