Seminar of I. Vekua Institute of Applied Mathematics
REPORTS, Vol. 35, 2009

ON THE RATE OF NORMAL APPROXIMATION FOR SUMS OF CONDITIONALLY INDEPENDENT RANDOM VARIABLES

Kvatadze Z., Shervashidze T.

Abstract. The main task of this note is to demonstrate the technique of reduction to the i.i.d. case of the proof of the normal approximation and its rate estimation for sums of conditionally independent random vectors with a general ergodic choice of the conditional distributions among finite number of fixed ones by the controlling sequence of random variables. The initial version with Markov chain as a controlling sequence was published in [1] with shortened proofs, next version [3] written in detail exists only as a manuscript and the present one although considering general control sequence is more transparent assigning conditional distributions as those of linear transformations of a fixed random vector having zero expectation and unit covariance matrix.

Keywords and phrases: Cramér's version of CLT for independent random vectors, conditionally independent vectors, switching conditional distributions via finite number of linear transformations by a general ergodic rule, limit theorem for sums, rate of normal approximation.

AMS subject classification (2000): 60F05; 60G50; 60G10.

We consider the problem of normal approximation of the distribution of sum of conditionally independent k-dimensional random vectors when each conditional distribution is obtained by a linear transformation of the fixed random vector.

1. Let us begin by dealing with the sequence of independent k-dimensional random vectors $X = (X_1, X_2, \dots)$, where $X_j = C_j Y_j$ and C_j are nonsingular $k \times k$ matrices, $j = 1, 2, \dots$, $Y = (Y_1, Y_2, \dots)$ are i.i.d. random vectors with $E(Y_1) = 0$ and $\operatorname{cov}(Y_1) = I$ [7]. By a result from Cramér's book [2, Ch. X] if for $n \to \infty$ the condition

$$\frac{1}{n}\sum_{j=1}^{n} C_j C_j^T \to C_0, \quad \operatorname{sp}(C_0) > 0, \tag{1}$$

and the Lindeberg condition

$$\frac{1}{n}\sum_{j=1}^{n} E\{|C_j Y_j|^2 I_{[|C_j Y_j| \ge \varepsilon\sqrt{n}]}\} \to 0 \ \ \forall \varepsilon > 0 \tag{2}$$

are fulfilled, then for $U_n = \sum_{j=1}^{n} C_j Y_j$ and Φ_B standing for normal distribution in R^k with parameters $(0, B)$ we have the weak convergence

$$P_{n^{-1/2}U_n} \xrightarrow{w} \Phi_{C_0}. \tag{3}$$

If among the matrices C_j, $j = 1, 2, \dots$, there are only a finite number of different ones $M_1, \dots, M_s$, then (2) is fulfilled automatically and (1) yields (3). If the relative frequency $\nu_n(\alpha)/n$ of M_α among $C_1, \ldots, C_n$ tends to $\pi_\alpha > 0$ as n increases to ∞, $\alpha = 1, \ldots, s$ ($\pi_1 + \cdots + \pi_s = 1$), the matrix C_0, being the limit of $\mathrm{cov}(U_n)/n$, has a form

$$C_0 = \pi_1 M_1 M_1^T + \cdots + \pi_s M_s M_s^T. \tag{4}$$

For the sum $V_n = \sum_{j=1}^n Z_j$ of i.i.d. random vectors $Z_1, \ldots, Z_n$ with $EZ_1 = 0$, $\mathrm{cov}(Z_1) = C$ the well-known estimate by Sazonov [5] refined by Senatov [6] reads as

$$\sup_{A \in \mathcal{C}^k} |P_{n^{-1/2}V_n}(A) - \Phi_C(A)| \leq c\, k \, \frac{1}{\sqrt{n}} \, E\|C^{-1/2} Z_1\|^3, \tag{5}$$

where $\mathcal{C}^k$ is the class of convex Borel subsets of R^k and c is an absolute constant.

As a special case of (5) with our initial Y_js instead of Z_js we have

$$\sup_{A \in \mathcal{C}^k} |P_{n^{-1/2}V_n}(A) - \Phi_I(A)| \leq c\, k\beta \, \frac{1}{\sqrt{n}}. \tag{6}$$

For the above situation with a finite number of weight matrices M_α with positive frequencies $\nu_n(\alpha)$, $\alpha = 1, \ldots, s$, estimate (6) and the representations

$$n^{-1/2} U_n = \sum_{\alpha=1}^{s} [\nu_n(\alpha)/n]^{1/2} M_\alpha [\nu_n(\alpha)]^{-1/2} T^{(\alpha)}_{\nu_n(\alpha)},$$

with s independent sums $T^{(\alpha)}$ of different groups of i.i.d. random vectors Y_j of cardinalities $\nu_n(\alpha)$, $\alpha = 1, \ldots, s$, and

$$\mathrm{cov}(U_n) = \sum_{\alpha=1}^{s} M_\alpha M_\alpha^T \nu_n(\alpha)$$

lead to the estimate

$$\sup_{A \in \mathcal{C}^k} |P_{n^{-1/2}U_n}(A) - \Phi_{\mathrm{cov}(U_n)/n}(A)| \leq c\, k\beta \sum_{\alpha : \nu_n(\alpha) > 0} \frac{1}{\sqrt{\nu_n(\alpha)}}. \tag{7}$$

When $\nu_n(\alpha)$ is equivalent to $n\pi_\alpha, \alpha = 1, \ldots, s$, the right-hand side of (7) is equivalent to

$$c\, k \sum_{\alpha=1}^{s} \pi_\alpha^{-1/2} \beta \, \frac{1}{\sqrt{n}}; \tag{8}$$

on the other hand, as in the latter case $\mathrm{cov}(U_n)/n$ tends to C_0 defined by (4) the normal distributions with corresponding covariances are close and due to the convolution property of the variation distance between probability distributions

$$v(\Phi_{\mathrm{cov}(U_n)/n}, \Phi_{C_0}) \leq \sum_{\alpha=1}^{s} v(\Phi_{(\nu_n(\alpha)/n)\, M_\alpha M_\alpha^T}, \Phi_{\pi_\alpha M_\alpha M_\alpha^T}) \tag{9}$$

their closeness can be estimated by the inequality

$$v(\Phi_{aC}, \Phi_{bC}) \leq \sqrt{k} \frac{|\sqrt{a} - \sqrt{b}|}{\sqrt{a+b}} \leq \sqrt{k} \frac{|a-b|}{a+b} \tag{10}$$

valid for any positive definite $k \times k$ matrix C and any two positive numbers a, b (see [4]). The latter bound together with inequalities (7) and (9) leads to the estimate

$$\sup_{A\in\mathcal{C}^k} |P_{n^{-1/2}U_n}(A) - \Phi_{C_0}(A)| \le \sum_{\alpha:\nu_n(\alpha)>0} \left\{ c\, k\beta \frac{1}{\sqrt{\nu_n(\alpha)}} + \frac{\sqrt{k}}{\pi_\alpha} \left| \frac{\nu_n(\alpha)}{n} - \pi_\alpha \right| \right\}. \quad (11)$$

Similar estimates take place for conditionally independent random vectors.

2. Let us consider a stationary two-component sequence (ξ_j, X_j), $j = 1, 2, \dots$, where ξ_j takes its values in $\{1, \dots, s\}$ and $X_j \in R^k$; denote

$$\xi = (\xi_1, \xi_2, \dots), \quad \xi_{1n} = (\xi_1, \dots, \xi_n),$$
$$X = (X_1, X_2, \dots), \quad X_{1n} = (X_1, \dots, X_n).$$

Definition. X is a sequence of conditionally independent random vectors controlled by a sequence ξ if for any natural n the conditional distribution $P_{X_{1n}|\xi_{1n}}$ of X_{1n} given ξ_{1n} is the direct product of conditional distributions of X_j given only the corresponding ξ_j, $j = 1, \dots, n$, i.e.,

$$\mathcal{P}_{X_{1n}|\xi_{1n}} = \mathcal{P}_{\xi_1} \times \dots \times \mathcal{P}_{\xi_n},$$

where $\mathcal{P}_\alpha$ is the conditional distribution of X_1 given $\{\xi_1 = \alpha\}$, $\alpha = 1, \dots, s$ (see, e.g., [1, 3, 8]). Let $\pi_\alpha = P\{\xi_1 = \alpha\}$, $\alpha = 1, \dots, s$, be a common distribution of ξ_js. For $s = 1$ X becomes a sequence of i.i.d. random variables with $\mathcal{P}_1$ as a common distribution.

To make further argument more transparent let us restrict ourselves by the case when conditional distributions given ξ_j are generated linearly from Y_1 via matrices $M_1, \dots, M_s$:

$$\mathcal{P}_\alpha = P_{M_\alpha Y_1}, \alpha = 1, \dots, s,$$

which can be expressed easier in the following way

$$X_j = M_{\xi_j} Y_j, \quad j = 1, 2, \dots. \quad (12)$$

The control sequence ξ is to be chosen ergodic, i.e., such that a.s. the relative frequency $\frac{\nu_n(\alpha)}{n} \to \pi_\alpha$, $\alpha = 1, \dots, s$, where for any α $\nu_n(\alpha) = \sum_{j=1}^n I_{(\xi_j=\alpha)}$ and

$$E\left| \frac{\nu_n(\alpha)}{n} - \pi_\alpha \right| \le \frac{d}{\sqrt{n}}, \quad d = const. \quad (13)$$

Theorem 1. *If in the stationary two-component sequence each member of the ergodic control sequence $\xi = (\xi_1, \xi_2, \dots)$ takes values in $\{1, \dots, s\}$ having the common distribution $\pi = (\pi_1, \dots, \pi_s)$ and the sequence of conditionally independent random vectors X given by the sequence $Y = (Y_j, j = 1, 2, \dots)$ of i.i.d. random vectors in R^k such that $EY_1 = 0, \operatorname{cov}(Y_1) = I$ and by the nonsingular matrices $M_1, \dots, M_s$ according to* (12)*, then for the normalized sum $S_n = \frac{1}{\sqrt{n}} \sum_{j=1}^n X_j$ we have*

$$1.\ P_{S_n|\xi_{1n}} \xrightarrow{w} \Phi_{C_0} \quad a.s., \quad 2.\ P_{S_n} \xrightarrow{w} \Phi_{C_0}$$

with C_0 given in the form (4).

The proof is based on the following decomposition formula for a given trajectory ξ_{1n}

$$S_n = \sum_{\alpha=1}^{s} \sqrt{\frac{\nu_n(\alpha)}{n}} S_n^{\alpha}, \tag{14}$$

where for each $\alpha = 1, \ldots, s$

$$S_n^{\alpha} = I_{(\nu_n(\alpha)>0)} \sqrt{\frac{1}{\nu_n(\alpha)}} M_{\alpha} \sum_{j=1}^{s} I_{(\xi_j=\alpha)} Y_j. \tag{15}$$

Theorem 2. *If under the conditions of Theorem* 1 $\beta = E|Y_1|^3 < \infty$ *and for relative frequencies the moment condition* (13) *is fulfilled, then*

$$\sup_{A\in\mathcal{C}^k} |P_{S_n}(A) - \Phi_{C_0}(A)| \le \frac{c(k,\pi,d,\beta)}{\sqrt{n}}\,,$$

where $c(k,\pi,d,\beta) = \sum_{\alpha=1}^{s}[c\,k(2/\pi_\alpha)^{1/2}\beta + (\sqrt{k}+2)d/\pi_\alpha]$.

For the proof we use the inequality

$$\begin{aligned}
&\sup_{A\in\mathcal{C}^k} |P_{S_n}(A) - \Phi_{C_0}(A)| \\
&\quad\le \sum_{\alpha=1}^{s} E\Big\{ I_{(\nu_n(\alpha)>0)} \sup_{A\in\mathcal{C}^k} |P_{[\frac{\nu_n(\alpha)}{n}]^{1/2} S_n^{\alpha}}(A) - \Phi_{\frac{\nu_n(\alpha)}{n} M_\alpha M_\alpha^T}(A)|\xi_{1,n}\Big\} \\
&\quad+ \sum_{\alpha=1}^{s} E\Big\{ I_{(\nu_n(\alpha)>0)} \sup_{A\in\mathcal{C}^k} |\Phi_{\frac{\nu_n(\alpha)}{n} M_\alpha M_\alpha^T}(A) - \Phi_{\pi_\alpha M_\alpha M_\alpha^T}(A)|\xi_{1,n}\Big\} =: \Sigma_1 + \Sigma_2.
\end{aligned}$$

Each summand of Σ_1 does not exceed the following sum

$$\begin{aligned}
&E\Big\{ I_{(\frac{\nu_n(\alpha)}{n} \ge \frac{\pi_\alpha}{2})} \sup_{A\in\mathcal{C}^k} |P_{[\frac{\nu_n(\alpha)}{n}]^{1/2} S_n^{\alpha}}(A) - \Phi_{\frac{\nu_n(\alpha)}{n} M_\alpha M_\alpha^T}(A)|\xi_{1,n}\Big\} \\
&\qquad + E\Big\{ I_{(0<\frac{\nu_n(\alpha)}{n} < \frac{\pi_\alpha}{2})} \sup_{A\in\mathcal{C}^k} |P_{[\frac{\nu_n(\alpha)}{n}]^{1/2} S_n^{\alpha}}(A) - \Phi_{\frac{\nu_n(\alpha)}{n} M_\alpha M_\alpha^T}(A)|\xi_{1,n}\Big\}.
\end{aligned}$$

To estimate the first part of the latter sum we use estimates (5), (6) and obtain $c\,k(2/\pi_\alpha)^{1/2}\beta n^{-1/2}$ for each α (note that an influence of M_α in this part is eliminated and in others, too). In the second part the difference of the probability measures does not exceed 1 and taking expectation of the indicator left using the condition (13) and we reach the sufficient order in n: $(2d/\pi_\alpha)n^{-1/2}$. What Σ_2 concerns its estimate is obtained similar to (9) and can be treated further using (11): $\Sigma_2 \le \sqrt{k}\sum_{\alpha=1}^{s} E\Big\{I_{(\nu_n(\alpha)>0)}\pi_\alpha^{-1}|\frac{\nu(n)}{n} - \pi_\alpha|\Big\} \le \sqrt{k}d\sum_{\alpha=1}^{s}\pi_\alpha^{-1}\frac{1}{\sqrt{n}}$. Finally, we achieve at the estimate

$$\sup_{A\in\mathcal{C}^k} |P_{S_n}(A) - \Phi_{C_0}(A)| \le n^{-1/2} \sum_{\alpha=1}^{s}[c\,k(2/\pi_\alpha)^{1/2}\beta + (\sqrt{k}+2)d/\pi_\alpha] = n^{-1/2}c(k,\pi,d,\beta).$$

Note that the case $E(Y_1) \neq 0$ was considered in [1, 3] when ξ is a regular stationary Markov chain and the same rate of convergence is obtained.

When ξ and Y are the independent sequences unless observable control sequence, probably the case where one may have an interest to the conditional theorems (see [1, 3]), the study is reduced to the i.i.d. case with the matrix C_0.

Acknowledgement. The second author was partially supported by the Georgian National Science Foundation (grant GNSF/ST07/3-172).

REFERENCES

1. Bokuchava I.V., Kvatadze Z.A., Shervashidze T.L. On limit theorems for random vectors controlled by a Markov chain. *Probability Theory and Mathematical Statistics,* (Vilnius, 1985), vol. I, 231–250, *VNU Sci. Press, Utrecht,* 1987.

2. Cramér H. Random Variables and Probability Distributions. 2nd ed. *Cambridge Tracts in Mathematics and Mathematical Physics. No.*36. *Cambridge: At the University Press*, 1962, 119 p.

3. Kvatadze Z. Limit Theorems for Some Classes of Dependent Random Variables (Georgian), The thesis for awarding the Scientific Degree of a Candidate of Physical and Mathematical Sciences, Tbilisi. N. Muskhelishvili Institute of Computational Mathematics of the Georgian Academy of Scienses, Tbilisi, 1996, 92 pp.

4. Kvatadze Z., Shervashidze T. On the accuracy of Kraft upper bound for the L_1-distance between Gaussian densities in R^k. *Georgian Math. J.* **12**, 4 (2005), 679-682.

5. Sazonov V.V. Normal Approximation - Some Recent Advances, Lecture Notes in Mathematics, 879. Berlin-Heidelberg-New York: Springer-Verlag, **VII**, (1981), 105 p.

6. Senatov V.V. Several uniform estimates of the rate of convergence in the multidimensional central limit theorem. (Russian) Teor. Veroyatn. Primen. **25** (1980), 757-770 .

7. Shervashidze T. Limit theorems for weighted sums of independent identically distributed random vectors, *Proc. A. Razmadze Math. Inst.* **136** (2004), 127-134.

8. Shervashidze T. On statistical analysis of a class of conditionally independent observations, *The International Scientific Conference "Problems of Cybernetics and Informatics", October* 24-26, 2006, *Baku, Azerbaijan,* **1**, 17-19. *Institute of Information Technologies of NASA, Baku,* 2006.www.pci2006.science.az/2/05.pdf

Received 7.07.2009; revised 8.09.2009; accepted 7.10.2009.

Authors' addresses:

Z. Kvatadze
Georgian Technical University
77, M. Kostava St., Tbilisi 0175
Georgia

T. Shervashidze
A. Razmadze Mathematical Institute
1, M. Aleksidze St., Tbilisi 0193
Georgia
E-mail: sher@rmi.acnet.ge