THE CENTRAL LIMIT THEOREM WITHOUT THE CONDITION OF INDEPENDENCE*

L. Szeidl (Budapest, Hungary) and V. M. Zolotarev (Moscow, Russia)

Necessary and sufficient conditions are presented for sums of asymptotically independent random variables to converge to a normal random variable in the sense of total variation distance, uniform metric for characteristic functions, and mean metric of order q.

Let us consider a sequence of random variables $X_1, X_2, \ldots$ with common distribution function $F(z)$ and characteristic function $f(t)$. We assume that $EX_1 = 0$, $EX_1^2 = 1$. Then the distribution functions $F_n(x)$ of the sums $S_n = n^{-1/2}(X_1 + \cdots + X_n)$ can be asymptotically approximated by the distribution function $\Phi(x)$ of the random variable $N$ having standard normal distribution. This limit approximation is treated in the sense of weak topology, i.e.,

$$L(S_n, N) = L(F_n, \Phi) \to 0, \quad n \to \infty,$$

where $L$ is the Lévy metric.

In fact, using a weaker condition than that of independence is a valid way to nullifying this condition completely. We demonstrate below how to prove the central limit theorem without the independence condition by making use of a stronger topology than the weak one.

Usually, weak convergence implies uniform $p$-convergence if the limit distribution is continuous, but to obtain convergence in a more strict sense, for instance in the metric of total variance

$$\sigma(X, Y) = \frac{1}{2} \int |d(F_X(z) - F_Y(z))|,$$

or

$$\chi_0(X, Y) = \sup_t |f_X(t) - f_Y(t)|,$$

where $F_X$ and $f_X$ stand for the distribution function and the characteristic function, respectively, of the random variable $X$, one should impose additional constraints. In 1952, Yu. V. Prokhorov [1] proved the following theorem.

**Theorem 1 (Prokhorov).** The condition $\sigma(F_n, \Phi) \to 0$ as $n \to \infty$ is valid if and only if the following two conditions are satisfied:

1. $L(F_n, \Phi) \to 0$ as $n \to \infty$,
2. the distribution functions $F_n$ for all sufficiently large $n$ have absolutely continuous components.

A second result of this kind, which is related to the metric $\chi_0$, was obtained by V. M. Zolotarev in 1986.

**Theorem 2 (Zolotarev).** The condition $\chi_0(F_n, \Phi) \to 0$ as $n \to \infty$ is valid if and only if the following two conditions are satisfied:

1. $L(F_n, \Phi) \to 0$ as $n \to \infty$,
2. $\lim_{n \to \infty} |f(t)| < 1$ (the so-called Cramér condition).

One more theorem of this type was obtained in [2] by V. Kruglov; he considered the metric

$$L_q(X, Y) = \left( \int |F_X(x) - F_Y(x)|^q \, dx \right)^{1/q},$$

*Supported by the Russian Foundation for Fundamental Research (grant No. 96-01-01920).
Theorem 3 (Kruglov). The condition $L_q(F_n, \Phi) \to 0$ as $n \to \infty$ is valid if and only if the following two conditions are satisfied:

1. $L(F_n, \Phi) \to 0$ as $n \to \infty$,
2. $\sup_n \int \left( \sum_{j=1}^{n} P(X_j > n^{1/2}x) \right)^q dx \to 0$ as $T \to \infty$.

It is not difficult to see that all three theorems are of the same structure. We can explain this by using the following property of probability metrics.

Let us consider the set of all distribution functions $X$ and two metrics $\mu$ and $\nu$ defined on $X$ such that $\mu < \nu$, i.e. the $\nu$-convergence of $F_n$ to $G$ implies the $\mu$-convergence of $F_n$ to $G$ but not vice versa. This means that the $\nu$-topology is stronger than the $\mu$-topology.

We say that a subset $A \subset X$ is $\nu$-compact if for any sequence $H_n \in X$, $n = 1, 2, \ldots$, we can choose a subsequence $\{H_n'\} \subset \{H_n\}$ such that for some $G \in X$ we have $\nu(H_n', G) \to 0$ as $n \to \infty$. The assertion concerning the metrics $\mu < \nu$ consists of the following.

Proposition 1. Assume that metrics $\mu$ and $\nu$ are comparable (say, $\mu \leq \nu$) on some set $A \subset X$. Then any two of the three properties given below imply the third one:

(a) the set $A$ is $\mu$-compact;
(b) the set $A$ is $\nu$-compact;
(c) the metrics $\mu$ and $\nu$ are equivalent on $A$.

Thus, from this viewpoint, condition (2) in all three theorems above is, in fact, the condition of the corresponding compactness of the set $A = \{F_n, n \geq 1\}$ with respect to the metrics $\sigma$, $\chi_0$, and $L_q$, respectively.

We would like to consider the following question: is it possible to eliminate the assumption of independence of the summands $X_j$ in the sums $S_n$ for the limit Theorems 1–3? The answer is positive, as we are going to see.

Let $J$ be the linear transformation of the $n$-dimensional vectors $X = (X_1, \ldots, X_n) \in \mathbb{R}^n$ for which $X = X_1 + \cdots + X_n$, i.e., $J = (1, \ldots, 1)$ is a row-matrix. It is clear that $J$ is a measurable mapping of $\mathbb{R}^n$ onto $\mathbb{R}$, and $S = JX$ is a random variable. Let us introduce the metric

$$
\lambda(X, Y) = \min_{T > 0} \max_{t \in T} \left( \frac{1}{2} \max_{|t| \leq T} |f_X(t) - f_Y(t)|, 1/T \right),
$$

where $f_X(t)$ and $f_Y(t)$ are the characteristic functions of the $n$-dimensional random vectors $X$ and $Y$, respectively.

The metric $\lambda$ induces the weak topology and hence is equivalent to the Lévy metric $L$ and the Lévy–Prokhorov metric $\pi$.

Lemma 1. For any random vectors $X'$ and $X''$, the following relations are true:

$$
\sigma(S', S'') \leq \sigma(X', X''), \quad \chi_0(S', S'') \leq \chi_0(X', X''), \quad \lambda(S', S'') \leq \lambda(X', X''),
$$

where $S' = JX'$, $S'' = JX''$.

Proof. First, let us deal with the metric $\sigma$. Since $\sigma$ is the minimal metric with respect to the indicator metric $i(X, Y) = E[I(X \neq Y)] = P(X \neq Y)$, we can write

$$
\{X = X''\} \subseteq \{JX' = JX''\} \implies \{X \neq X''\} \supseteq \{JX' \neq JX''\},
$$

$$
i(X', X'') \geq i(JX', JX'') \implies \sigma(X', X'') \geq \sigma(JX', JX'') = \sigma(S', S'').
$$

Now, let us consider the metric $\chi_0$. By definition,

$$
\chi_0(X', X'') = \sup_{t} |f_{X'}(t) - f_{X''}(t)|, \quad t \in \mathbb{R}^n.
$$

$$
\chi_0(S', S'') = \sup_{\tau} |f_S(\tau) - f_{S''}(\tau)|, \quad \tau \in \mathbb{R}^n.
$$

It is clear that for $t = (t_1, \ldots, t_n)$, $t_j \in \mathbb{R}$, we obtain

$$
f_S(\tau) - f_{S''}(\tau) = f_{X'}(t) - f_{X''}(t).
$$

3003