ON THE RATE OF CONVERGENCE
IN A RANDOM CENTRAL LIMIT THEOREM

BY

K. S. KUBACKI AND D. SZYNAL (LUBLIN)

Abstract. We extend the random central limit theorem of Rényi [8] and theorems on the convergence rate for random summation of [3] and [1] to the case where a larger class of random indices is considered.

1. Let \( \{X_k, k \geq 1\} \) be a sequence of independent random variables (r.v.'s) with \( EX_k = 0, EX_k^2 = \sigma_k^2 < \infty, k \geq 1 \). Suppose that there exists a probability measure \( \mu \) such that
\[
Y_n := \frac{S_n}{s_n} \Rightarrow \mu, \quad n \to \infty \quad \text{(converges weakly)},
\]
where
\[
S_n = \sum_{k=1}^{n} X_k, \quad s_n^2 = \sum_{k=1}^{n} \sigma_k^2 < \infty
\]
for all \( n \), and \( s_n^2 \to \infty, n \to \infty \).

We are going to prove the following results:

THEOREM 1. Let \( \{X_k, k \geq 1\} \) be a sequence of independent r.v.'s with \( EX_k = 0, EX_k^2 = \sigma_k^2, k \geq 1 \), satisfying (1.1), and let \( \{N_n, n \geq 1\} \) be a sequence of positive integer-valued r.v.'s such that
\[
s_n^2/s_{\{N_n\}}^2 \overset{p}{\to} 1, \quad n \to \infty \quad \text{(converges in probability)},
\]
where \( \lambda \) is a positive r.v. having a discrete distribution, and \( \{v_n, n \geq 1\} \) is a sequence of positive integer-valued r.v.'s independent of \( \{\lambda, X_k, k \geq 1\} \) with \( v_n \overset{p}{\to} \infty, n \to \infty \). Then
\[
Y_n = \frac{S_n}{s_n} \Rightarrow \mu, \quad n \to \infty.
\]

THEOREM 2. Let \( \{X_k, k \geq 1\} \) be a sequence of independent r.v.'s with \( EX_k = 0, EX_k^2 = \sigma_k^2, k \geq 1 \), satisfying the following conditions:
\[
E|X_k|^{2+\delta} = \beta_k^{2+\delta} < \infty, \quad k \geq 1, \quad \text{for some } 0 < \delta \leq 1;
\]
there exist positive numbers $b_1$ and $b_2$ such that, for every positive integers $n > k \geq 1$,

\begin{equation}
\sum_{k=1}^{n} \beta_k^{2+\delta} = O(s_n^2), \quad \text{where } B_n^{2+\delta} = \sum_{k=1}^{n} \beta_k^{2+\delta};
\end{equation}

Suppose that $\{N_n, n \geq 1\}$ is a sequence of positive integer-valued r.v.'s such that, for a constant $C_1$,

\begin{equation}
P[s_n^2/s_{v_n}^2 - 1 > C_1 \varepsilon_n] = O(\sqrt{\varepsilon_n}),
\end{equation}

where $\{v_n, n \geq 1\}$ is a sequence of positive integer-valued r.v.'s independent of $\{X_k, k \geq 1\}$ with

\begin{equation}
P[s_n^2 < C_2 \varepsilon_n^{1/\delta}] = O(\sqrt{\varepsilon_n}) \text{ for a constant } C_2,
\end{equation}

and $\{\varepsilon_n, n \geq 1\}$ is a sequence of positive numbers with $\varepsilon_n \to 0$, $n \to \infty$. Then

\begin{equation}
\sup_{x} |P[S_{N_n} < xs_n] - \Phi(x)| = O(\sqrt{\varepsilon_n})
\end{equation}

and

\begin{equation}
\sup_{x} |P[S_{N_n} < xs_n] - \Phi(x)| = O(\sqrt{\varepsilon_n}),
\end{equation}

where $\Phi$ denotes the standard normal distribution function.

From Theorem 1 we get a generalization of Rényi's result [8].

**Corollary 1.** Let $\{X_k, k \geq 1\}$ be a sequence of independent and identically distributed r.v.'s with $EX_1 = 0$, $EX_1^2 = \sigma^2 > 0$, and let $\{N_n, n \geq 1\}$ be a sequence of positive integer-valued r.v.'s such that

\begin{equation}
N_n/\sqrt{v_n} \xrightarrow{P} \lambda, \quad n \to \infty,
\end{equation}

where $\lambda$ is a positive r.v. having a discrete distribution, and $\{v_n, n \geq 1\}$ is a sequence of positive integer-valued r.v.'s independent of $\{\lambda, X_k, k \geq 1\}$ with $v_n \to \infty$, $n \to \infty$. Then

\begin{equation}
S_{N_n}/\sigma \sqrt{N_n} \Rightarrow \mathcal{N}_{0,1}, \quad n \to \infty.
\end{equation}

Theorem 2 gives us the following generalization of the Callaert and Janssen's result [1]:

**Corollary 2.** Let $\{X_k, k \geq 1\}$ be a sequence of independent and identically distributed r.v.'s with $EX_1 = 0$, $EX_1^2 = \sigma^2 > 0$, $E|X_1|^{2+\delta} < \infty$ for some $0 < \delta \leq 1$ and, let $\{N_n, n \geq 1\}$ be a sequence of positive integer-valued r.v.'s such that, for a constant $C_1$,
(1.13) \[ P \left( \left| \frac{N_n}{v_n} - 1 \right| > C_1 \varepsilon_n \right) = O(\sqrt{\varepsilon_n}), \]

where \( \{v_n, n \geq 1\} \) is a sequence of positive integer-valued r.v.'s independent of \( \{X_k, k \geq 1\} \) with.

(1.14) \[ P [v_n < C_2 \varepsilon_n^{-1/\delta}] = O(\sqrt{\varepsilon_n}) \text{ for a constant } C_2, \]

and \( \{\varepsilon_n, n \geq 1\} \) is a sequence of positive numbers with \( \varepsilon_n \to 0, n \to \infty \). Then

(1.15) \[ \sup_x P [S_{N_n} < x \sigma \sqrt{v_n}] - \Phi(x) = O(\sqrt{\varepsilon_n}) \]

and

(1.16) \[ \sup_x P [S_{N_n} < x \sigma \sqrt{N_n}] - \Phi(x) = O(\sqrt{\varepsilon_n}). \]

2. In order to prove Theorems 1 and 2 we need the following auxiliary results.

**Lemma 1.** Let \( \{X_k, k \geq 1\} \) be a sequence of independent r.v.'s with \( EX_k = 0, EX_k^2 = \sigma_k^2 < \infty, k \geq 1 \), satisfying (1.1), and let \( \lambda \) be a positive r.v. having a discrete distribution. If \( \{v_n, n \geq 1\} \) is a sequence of positive integer-valued r.v.'s independent of \( \{\lambda, X_k, k \geq 1\} \) with \( v_n \overset{P}{\to} \infty, n \to \infty \), then

(2.1) \[ Y_{[\lambda v_n]} \Rightarrow \mu, \quad n \to \infty. \]

Moreover, the sequence \( \{Y_n, n \geq 1\} \) satisfies the Anscombe random condition ((A***) of [2]) with norming sequence \( \{s_n, n \geq 1\} \) and filtering sequence \( \{[\lambda v_n], n \geq 1\} \), i.e., for every \( \varepsilon > 0 \) there exists a \( \delta > 0 \) such that

(2.2) \[ \limsup_{n \to \infty} P \left( \max_{i} |Y_i - Y_{[\lambda v_n]}| > \varepsilon \right) \leq \varepsilon, \]

where \( I = \{i: |s_i^2 - s_{[\lambda v_n]}^2| < \delta s_{[\lambda v_n]}^2\} \).

**Lemma 2.** Let \( \{X_k, k \geq 1\} \) be a sequence of independent r.v.'s with \( EX_k = 0, EX_k^2 = \sigma_k^2, E|X_k|^{2+\delta} = \beta_k^{2+\delta} < \infty \) for some \( 0 < \delta \leq 1, k \geq 1 \). Then there exists a constant \( C \) such that, for every positive integers \( n \) and \( k \) and for every \( x \),

(2.3) \[ P [S_n < x; S_{n+k} \geq x] \leq C \left\{ B_n^{2+\delta}/s_n^{2+\delta} + \sqrt{(s_n^{2+\delta} - s_k^{2+\delta})/s_n^{2+\delta}} \right\} \]

and

(2.4) \[ P [S_n \geq x; S_{n+k} < x] \leq C \left\{ B_n^{2+\delta}/s_n^{2+\delta} + \sqrt{(s_n^{2+\delta} - s_k^{2+\delta})/s_n^{2+\delta}} \right\}. \]

**Proof of Lemma 1.** Since for every \( n \) the r.v. \( v_n \) is independent of \( \{\lambda, X_k, k \geq 1\} \), we have

\[ P[Y_{[\lambda v_n]} < x] = \sum_{k=1}^{\infty} P[v_n = k] P[Y_{[\lambda k]} < x]. \]
But, by Lemma 6 of [2],
\[
\lim_{k \to \infty} P[Y_{\lambda k} < x] = F(x)
\]
for every continuity point \(x\) of \(F\), where \(F(\cdot) = \mu((-\infty, \cdot))\). Furthermore, since \(v_n \to \infty, n \to \infty\), we have
\[
\lim_{n \to \infty} P[v_n = k] = 0 \quad \text{for every } k \geq 1.
\]
Thus, by Toeplitz lemma (cf. [5], p. 238),
\[
\lim_{n \to \infty} P[Y_{\lambda v_n} < x] = F(x)
\]
for every continuity point \(x\) of \(F\), which proves (2.1).

For the proof of (2.2) let us note that, for arbitrarily fixed \(M > 0\),
\[
P[\max_{I_1} |Y_i - Y_{\lambda v_n}| \geq \varepsilon] \leq P[v_n \leq M] + \sum_{k > M} P[v_n = k] P[\max_{I_2} |Y_i - Y_{\lambda k}| \geq \varepsilon].
\]
where \(I_1 = |s_i^2 - s_{I_1 v_n}^2| \leq \delta s_{I_1 v_n}^2, I_2 = |s_i^2 - s_{I_2 k}^2| \leq \delta s_{I_2 k}^2\), and \(\lim_{n \to \infty} P[v_n \leq M] = 0\). Moreover, by Lemma 7 of [2], we can choose \(M\) so large that
\[
P[\max_{I_2} |Y_i - Y_{\lambda k}| \geq \varepsilon] \leq \varepsilon \quad \text{for all } k > M.
\]
Hence we get the desired result (2.2).

Proof of Lemma 2. Since (2.4) follows from (2.3) by replacing \(X_k\) by \(-X_k\), we prove (2.3) only.

We put \(D(n; x) = P[S_n \leq x; S_{n+k} \geq x]\). Then, by the theorem of Fubini, we obtain
\[
D(n; x) = \int_{E_1} \cdots dF_{X_1}(x_1) \cdots dF_{X_{n+k}}(x_{n+k})
\]
\[
= \int_{E_1} \cdots P[S_n \leq x; S_n + \sum_{i=1}^{n+k} x_i \geq x] dF_{X_{n+1}}(x_{n+1}) \cdots dF_{X_{n+k}}(x_{n+k})
\]
\[
= \int_{E_1} \cdots P[x - \sum_{i=1}^{n+k} x_i \leq S_n \leq x] dF_{X_{n+1}}(x_{n+1}) \cdots dF_{X_{n+k}}(x_{n-k}),
\]
where \(E_1 = [\sum_{i=1}^{n+k} x_i \leq x; \sum_{i=1}^{n+k} x_i \geq x]\). Hence
\[
D(n; x) = \int_{E_2} \{P\left[\frac{S_n}{s_n} \leq \frac{x}{s_n}\right] - \Phi\left(\frac{x}{s_n}\right) +
\]
\[
+ \Phi\left(\frac{x}{s_n} - \frac{\sum_{i=n+1}^{n+k} x_i}{s_n}\right) - P\left[\frac{S_n}{s_n} < \frac{x}{s_n} - \frac{\sum_{i=n+1}^{n+k} x_i}{s_n}\right]\}
\]
Rate of convergence

\[
\Phi \left( \frac{x}{s_n} \right) - \Phi \left( \frac{x}{s_n} - \sum_{i=n+1}^{n+k} \frac{x_i}{s_n} \right) \leq \frac{1}{\sqrt{2\pi}} |q| E_2 \leq 1.
\]

where \( E_2 = \sum_{i=n+1}^{n+k} x_i \geq 0 \).

Since, for every \( q \) (cf. [7], inequality (3.4) on p. 143),

\[
\sup_y |\Phi(y+q) - \Phi(y)| \leq |q|/\sqrt{2\pi}
\]

and, by Berry-Esseen inequality (cf. [7], Theorem 6 on p. 144),

\[
(2.5) \quad \sup_y \left| P \left[ \frac{S_n}{s_n} < y \right] - \Phi(y) \right| = O \left( B_n^{2+\delta}/s_n^{2+\delta} \right),
\]

there exists a constant \( C \) such that

\[
D(n; x) \leq C \left( B_n^{2+\delta}/s_n^{2+\delta} + \sqrt{(s_n^{2+\delta} - s_n^2)/s_n^2} \times \right.
\]

\[
\left. \sum_{i=n+1}^{n+k} x_i \int \ldots \int \frac{dF_{x_{n+1}}(x_{n+1}) \ldots dF_{x_{n+k}}(x_{n+k})}{\sqrt{s_n^2 - s_{n+k}^2}} \right).
\]

But

\[
\sum_{i=n+1}^{n+k} x_i \int \ldots \int \frac{dF_{x_{n+1}}(x_{n+1}) \ldots dF_{x_{n+k}}(x_{n+k})}{\sqrt{s_n^2 - s_{n+k}^2}}
\]

\[
= E \left| \frac{S_{n+k} - S_n}{\sqrt{s_n^2 - s_{n+k}^2}} \right| \leq E^{1/2} \left( \frac{S_{n+k} - S_n}{\sqrt{s_n^2 - s_{n+k}^2}} \right)^2 = 1.
\]

Therefore, we get the desired result (2.3). The proof of Lemma 2 is complete.

Proof of Theorem 1. The proof is easily based on Lemma 1 and Corollary 2 of [2] and is not detailed here.

Proof of Theorem 2. The proof contains some ideas of [4] and bases on Lemma 2 and Lemma 6.1 of [9].

First we observe that

\[
(2.6) \quad \sup_{x} \left| P \left[ S_{x_n} < x s_n \right] - \Phi(x) \right| = O \left( \sqrt{\epsilon_n} \right).
\]

Indeed, by (1.5), (2.5) and the fact that \( v_n \) is independent of \( \{ X_k, k \geq 1 \} \)
we obtain

\[(2.7) \quad \sup_x |P[S_{v_n} < x s_{v_n}] - \Phi(x)| \leq \sum_{k=1}^{\infty} P[v_n = k] \sup_x |P[S_k < x s_k] - \Phi(x)| \leq C \sum_{k=1}^{\infty} P[v_n = k] \{B_k^{2+\delta}/s_k^{2+\delta}\} \leq \bar{C} E \{s_{v_n}^{-\delta}\},\]

where \(C\) and \(\bar{C}\) are some constants independent of \(n\) and \(k\). But, by (1.8), we have (with the assumption \(\sigma_1^2 > 0\))

\[(2.8) \quad E \{s_{v_n}^{-\delta}\} = E \{s_{v_n}^{-\delta} I[s_{v_n}^2 < C_2 e_n^{-1/\delta}] + s_{v_n}^{-\delta} I[s_{v_n}^2 \geq C_2 e_n^{-1/\delta}]\} \leq \sigma_1^{-\delta} P[s_{v_n}^2 < C_2 e_n^{-1/\delta}] + (C_2 e_n^{-1/\delta})^{-\delta/2} = O(\sqrt{e_n}).\]

Combining (2.7) and (2.8) we get the desired result (2.6).

Now let us put

\[I_n = \{k: (1 - C_1 e_n)s_{v_n}^2 \leq s_{k}^2 \leq (1 + C_1 e_n)s_{v_n}^2\}\]

and

\[I_{n,r} = \{k: (1 - C_1 e_n)s_{v_n}^2 \leq s_{k}^2 \leq (1 + C_1 e_n)s_{r}^2\}, \quad r \geq 1.\]

By (1.7) we have

\[P[\max_{k \in I_n} S_k < x s_{v_n}] - O(\sqrt{e_n}) \leq P[S_{v_n} < x s_{v_n}] \leq P[\min_{k \in I_n} S_k < x s_{v_n}] + O(\sqrt{e_n}).\]

Furthermore,

\[P[\max_{k \in I_{n,r}} S_k < x s_{v_n}] \leq P[S_{v_n} < x s_{v_n}] \leq P[\min_{k \in I_{n,r}} S_k < x s_{v_n}].\]

Hence, and by (2.6), we obtain

\[(2.9) \quad \sup_x |P[S_{v_n} < x s_{v_n}] - \Phi(x)| \leq O(\sqrt{e_n}) + \sup_x |P[\min_{k \in I_n} S_k < x s_{v_n}] - P[\max_{k \in I_n} S_k < x s_{v_n}]|.

Let

\[p_n = \min \{k: s_k^2 \geq (1 - C_1 e_n)s_{v_n}^2\}\]

and

\[q_n = \max \{k: s_k^2 \leq (1 + C_1 e_n)s_{v_n}^2\}.

Let

\[k_{1,n} = \{k: s_k^2 < (1 - C_1 e_n)s_{v_n}^2\} \quad \text{and} \quad k_{2,n} = \{k: s_k^2 > (1 + C_1 e_n)s_{v_n}^2\}.

Therefore

\[P[\min_{k \in I_n} S_k < x s_{v_n}] \leq P[\max_{k \in I_n} S_k < x s_{v_n}] + O(\sqrt{e_n}).\]
Then

\[(2.10) \quad P[\min_{k \in I_n,v_n} S_k < x_{s_{y_n}}] - P[\max_{k \in I_n,v_n} S_k < x_{s_{y_n}}] = \sum_{r=1}^{\infty} P[v_n = r] K_{n,r},\]

where

\[K_{n,r} := P[\min_{p_n \leq k \leq q_n} S_k < x_{s_{y_n}}] - P[\max_{p_n \leq k \leq q_n} S_k < x_{s_{y_n}}].\]

According to Lemma 6.1 of [9] there exists a constant C, independent of \(n, r\) and \(x\), such that

\[(2.11) \quad K_{n,r} \leq C \{P[S_{p_n} < x_{s_{y_n}}; S_{q_n} > x_{s_{y_n}}] + P[S_{p_n} > x_{s_{y_n}}; S_{q_n} < x_{s_{y_n}}]\}.

For the completeness of the proof we repeat here the proof of inequality (2.11).

First we note that \(K_{n,r} = P[S_j < x_{s_{y_n}}; S_k > x_{s_{y_n}}]\) for some \(j\) and \(k\), \(p_n < j < k < q_n\) = \(P(A)\), say. Furthermore,

\[P(A) = P(A \cap [S_{p_n} < x_{s_{y_n}}; S_{q_n} < x_{s_{y_n}}]) + P(A \cap [S_{p_n} < x_{s_{y_n}}; S_{q_n} > x_{s_{y_n}}]) + P(A \cap [S_{p_n} > x_{s_{y_n}}; S_{q_n} > x_{s_{y_n}}]) + P(A \cap [S_{p_n} > x_{s_{y_n}}; S_{q_n} < x_{s_{y_n}}]) + P(A \cap [S_{p_n} > x_{s_{y_n}}; S_{q_n} < x_{s_{y_n}}]) + P(S_{p_n} > x_{s_{y_n}}; S_{q_n} < x_{s_{y_n}})

\]

Hence it is sufficient to prove that there exists a constant \(C\) (independent of \(n, r\) and \(x\)) such that

\[(2.12) \quad P(A \cap [S_{p_n} < x_{s_{y_n}}; S_{q_n} < x_{s_{y_n}}]) \leq C \cdot P[S_{p_n} < x_{s_{y_n}}; S_{q_n} > x_{s_{y_n}}]

and

\[(2.13) \quad P(A \cap [S_{p_n} > x_{s_{y_n}}; S_{q_n} > x_{s_{y_n}}]) \leq C \cdot P[S_{p_n} > x_{s_{y_n}}; S_{q_n} < x_{s_{y_n}}].

Define, for \(p_n + 1 \leq k \leq q_n\), \(A_k = [S_k < x_{s_{y_n}}\) for \(p_n < j < k-1; S_k > x_{s_{y_n}}\]. Then

\[P(A \cap [S_{p_n} < x_{s_{y_n}}; S_{q_n} < x_{s_{y_n}}]) = \sum_{k=p_n+1}^{q_n} P(A_k \cap [S_{p_n} < x_{s_{y_n}}]) \leq \sum_{k=p_n+1}^{q_n-1} P(A_k) P[S_{p_n} - S_k \leq 0] (by \(2.6)) \leq b_2 \sum_{k=p_n+1}^{q_n-1} P(A_k) P[S_{p_n} - S_k \leq 0] (by \(2.6)) \leq b_2 \sum_{k=p_n+1}^{q_n-1} P(A_k \cap [S_{p_n} > S_K]) \leq b_2 \sum_{k=p_n+1}^{q_n-1} P(A_k) \leq b_2 P[S_{p_n} > x_{s_{y_n}}; S_{q_n} > x_{s_{y_n}}],

which proves (2.12).

Inequality (2.13) follows similarly. Thus we get (2.11).
Using (2.11) and Lemma 2 we have, for a constant $C$,
\[ K_{n,r} \leq C \left\{ \frac{B_{\delta_{n,r}}}{s_{\delta_{n,r}}} + \sqrt{\frac{s_{\overline{r}_n}^2 - s_{\overline{r}_n}^2}{s_{\overline{r}_n}^2}} \right\}, \]
where, by (1.5),
\[ B_{\delta_{n,r}} \leq \tilde{C} s_{\delta_{n,r}}^{-\delta} + \sqrt{(1 + C_1 \varepsilon_n) s_{\overline{r}_n}^2 - (1 - C_1 \varepsilon_n) s_{\overline{r}_n}^2}/(1 - C_1 \varepsilon_n) s_{\overline{r}_n}^2 \]
\[ \leq \tilde{C} (1 - C_1 \varepsilon_n)^{-\delta/2} s_{\overline{r}_n}^{-\delta} + \sqrt{2C_1 \varepsilon_n/(1 - C_1 \varepsilon_n)} \leq \tilde{C}_1 s_{\overline{r}_n}^{-\delta} + O(\sqrt{\varepsilon_n}) \]
for some constants $\tilde{C}$ and $\tilde{C}_1$ independent of $n$ and $r$. Hence and by (2.8) we obtain, for a constant $C$,
\[ \sum_{r=1}^{\infty} P[y_n = r] K_{n,r} \leq CE \{ s_{\delta_{n,r}}^{-\delta} \} + O(\sqrt{\varepsilon_n}) = O(\sqrt{\varepsilon_n}), \]
which, combined with (2.9) and (2.10), yields (1.9).

Further on, by (1.7), (1.9) and Lemma 1 of [6], stating that if \{X_n, n \geq 1\} and \{Y_n, n \geq 1\} are sequences of r.v.'s such that
\[ \sup_x |P[X_n < x] - \Phi(x)| = O(a_n) \quad \text{and} \quad P[|Y_n - 1| > a_n] = O(a_n), \]
then
\[ \sup_x |P[X_n < xY_n] - \Phi(x)| = O(a_n), \]
we obtain (1.10). The proof of Theorem 2 is complete.

Remark. One can see that each sequence \{X_k, k \geq 1\} of independent and identically distributed r.v.'s, with $EX_1 = 0$ and $\sigma^2 X_1 = \sigma^2 < \infty$, satisfies (1.6) ([9], p. 95). This observation and Theorem 2 imply Corollary 2.

REFERENCES
Rate of convergence


Received on 26. 3. 1986;
new version on 27. 10. 1986