THE ERDÖS-RÉNYI NEW LAW OF LARGE NUMBERS FOR WEIGHTED SUMS

STEPHEN A. BOOK

Abstract. From the partial sums $S_n$ of the first $N$ random variables of a sequence of independent, identically distributed random variables, $N-K+1$ averages of the form $K^{-1}(S_{n+K}-S_n)$ can be constructed, one such average for each $n$ between 0 and $N-K$, inclusive. If we denote by $\Sigma(N, K)$ the maximum of those $N-K+1$ averages, then for a wide range of numbers $\lambda$, Erdös and Rényi (1970) proved that, as $N\to\infty$, $\Sigma(N, K)\to\lambda$ a.e. for $K=\lfloor C(\lambda)\log N \rfloor$, where $C(\lambda)$ is a constant depending only on $\lambda$, not on $N$. The objective of the present article is to extend the Erdös-Rényi theorem to the case of weighted sums. The main theorem bears a relation to the law of large numbers for weighted sums of Jamison, Orey, and Pruitt (1965) similar to the one borne by the Erdös-Rényi theorem to the ordinary strong law of large numbers.

0. Introduction. If $X_1, X_2, \cdots$ is a sequence of independent, identically distributed (i.i.d.) random variables satisfying some conditions, $S_n = \sum_{k=1}^n X_k$, and $\Sigma(N, K) = \max\{K^{-1}(S_{n+K}-S_n): 0 \leq n \leq N-K\}$, then Erdös and Rényi [3] showed that $P\lim_{N \to \infty} \Sigma(N, \lfloor C(\lambda)\log N \rfloor) = \lambda = 1$ for a wide range of positive numbers $\lambda$, including all those for which $P(X_1 > \lambda) > 0$. In what follows, $\lfloor y \rfloor$ denotes the greatest integer $\leq y$, and $S_0 = 0$. Here $C(\lambda)$ is a positive constant depending only on $\lambda$; explicitly, $C(\lambda) = -1/\log \rho(\lambda)$, where $\rho(\lambda) = \inf\{\phi(t)e^{-\lambda t}: t > 0\}$ for $\phi$ the moment-generating function (m.g.f.) of $X_1$. Erdös and Rényi call their result "a new law of large numbers" because it fills up all the middle ground between the ordinary strong law of large numbers, which asserts that $P\lim_{N \to \infty} \Sigma(N, N) = 0 = 1$ if $E(X_1) = 0$, and the fact that $P(\lim_{N \to \infty} \Sigma(N, 1) = \mu) = 1$ where $\mu \leq \infty$ is the largest possible value of $X_1$. The $\lambda$'s to which the Erdös-Rényi theorem applies range from 0 to $\mu$, exclusive of the endpoints. It is the goal of the present article to extend the Erdös-Rényi theorem to the case of weighted sums of i.i.d. random variables.

The strong law of large numbers of Jamison, Orey, and Pruitt [4] asserts that, for certain sequences of weights $\{a_k: 1 \leq k < \infty\}$, we have...
$P\{\lim_{n\to\infty} A^{-1}S_n=0\}=1$, where $S_n=\sum_{k=1}^n a_kX_k$ and $A_n=\sum_{k=1}^n a_k$. If we define $\Sigma(N, K)=\max\{(A_{n+K}-A_n)^{-1}(S_{n+K}-S_n):0\leq n\leq N-K\}$, then the strong law of large numbers becomes $P\{\lim_{N\to\infty} \Sigma(N, N)=0\}=1$, while at the other extreme we again have $P\{\lim_{N\to\infty} \Sigma(N, 1)=\mu\}=1$. The main theorem of the present paper asserts that, for a wide range of positive $\lambda$'s, there exist positive constants $C'(\lambda)$ and $C''(\lambda)$, which can be determined explicitly with $C'(\lambda)\leq C''(\lambda)$, such that

$$P\left\{\lim \inf_{N\to\infty} \Sigma(N, [C'(\lambda)\log N]) \geq \lambda\right\} = 1$$

and

$$P\left\{\lim \sup_{N\to\infty} \Sigma(N, [C''(\lambda)\log N]) \leq \lambda\right\} = 1.$$

When each $a_k=1$, it turns out that $C'(\lambda)$ and $C''(\lambda)$ are identical, so that our theorem reduces to that of Erdős and Rényi. The latter two authors base their proof on the large deviation theorem of Bahadur and Ranga Rao [1]. In place of the Bahadur-Ranga Rao theorem, our starting point is the large deviation theorem for weighted sums that appears in [2].

1. Conditions on the weights and random variables. We consider a sequence of nondegenerate i.i.d. random variables $X_1, X_2, \cdots$ such that $E(X_1)=0$ and $E(X_1^2)=1$, and a sequence $\{a_k: 1\leq k<\infty\}$ of nonnegative real numbers, not all zero. We require that the larger and smaller $a_k$'s be distributed somewhat regularly throughout the sequence. More precisely, we assume:

**Condition A.** There exist constants $\alpha$ and $\theta$, $0<\alpha\leq 1$, $0<\theta\leq 1$, such that, for every sufficiently large $K$, every set of $K$ consecutive $a_k$'s has the property that at least $\alpha K$ of them exceed or equal $\theta \cdot \max\{a_k: a_k \in \text{set of } K\}$.

Of the random variables, we require first that the m.g.f. $\phi$ of $X_1$ exist in a nondegenerate interval about the origin. In addition to this, Erdős and Rényi assumed the Bahadur-Ranga Rao condition that the function $\phi(t)e^{-\lambda t}$ takes on its minimum value at a point interior to the domain of $\phi$. If we define $Q(t)=\phi'(t)/\phi(t)$, the Bahadur-Ranga Rao condition is equivalent to the statement that $Q(t)$ assumes the value $\lambda$ at some $t$. If $P(X_1>\lambda)>0$, then that condition automatically holds if the domain of $\phi$ is large enough. We require an analogous condition of $\lambda$ and $X_1$:

**Condition B.** $\lambda$ is a positive real number for which there exists a $\lambda^*>\lambda$ such that $Q(t)$ assumes the value $\lambda^*/(\alpha\theta)^{-1}$, and $B_{\lambda^*}=B_{\alpha\theta}^{-1}(\lambda^*/\alpha\theta)$ lies in the domain of $\phi$.

The above condition holds, for example, if $\phi(t)<\infty$ for all real $t$ and $P(X_1>\lambda(\alpha\theta)^{-1})>0$. Finally, in order to make use of the large deviation theorem of [2], we must require:
Condition C. $X_1$ is absolutely continuous with density $f(x)$, and the functions $f_y(x) = e^{yf(x)}$ have uniformly bounded total variations $V_y \leq V$ for $0 \leq y \leq B_0^*$. The assumption of absolute continuity can most likely be weakened in the large deviation theorem, so there is no reason to suspect that Condition C is crucial.

Throughout the rest of the paper, we implicitly assume that Conditions A, B, and C are satisfied.

2. Application of the large deviation theorem. In this section, we apply Theorem (4.8) of [2], the large deviation theorem for weighted sums, in order to obtain estimates needed in the proof of the main theorem. The statement of the large deviation theorem is as follows:

Lemma 1. If $\{a_{nk}: 1 \leq k \leq n, 1 \leq n < \infty\}$ is an array of nonnegative real numbers such that $\sum_{k=1}^{n} a_{nk}^2 = 1$ for each $n$, and at least $a_n$ of the $a_{nk}$'s, $1 \leq k \leq n$, exceed or equal $\theta \sigma_n^*$, where $\sigma_n^* = \max\{a_{nk}: 1 \leq k \leq n\}$, $S_n^* = \sum_{k=1}^{n} a_{nk}X_k$, and $A_n^* = \sum_{k=1}^{n} a_{nk}$, then

$$P(S_n^* > \lambda A_n^*) = (2\pi)^{-1/2} (\sigma_n^* h_n^*)^{-1} \exp(-h_n^* \lambda A_n^*) \left( \prod_{k=1}^{n} \phi(h_n^* a_{nk}) \right) (1 + O(n^{-1/2})), $$

where $0 < b_\lambda = Q^{-1}(\lambda/\sigma_\lambda^2) \leq h_n^* \sigma_n^* \leq Q^{-1}(\lambda/\sigma_\lambda^2) = B_\lambda < \infty$, and $0 < d_\lambda^2 \leq \sigma_\lambda^2 \leq D_\lambda^2 < \infty$ for $d_\lambda^2 = \min\{Q'(t): 0 \leq t \leq B_\lambda\}$ and $D_\lambda^2 = \max\{Q'(t): 0 \leq t \leq B_\lambda\}$.

What we will need are asymptotic representations for $P_n(K, \lambda) = P\{S_{n+K} - S_n > \lambda(A_{n+K} - A_n)\}$ where $0 \leq n \leq N - K$. We fix $n$ and define $a_{Kk} = a_k(\sum_{k=n+1}^{n+K} a_k^2)^{-1/2}$ for $n+1 \leq k \leq n+K$, so that

$$P_n(K, \lambda) = P\left\{ \sum_{k=n+1}^{n+K} a_{Kk}X_k > \lambda \sum_{k=n+1}^{n+K} a_{Kk} \right\},$$

where $\sum_{k=n+1}^{n+K} a_{Kk}^2 = 1$, and at least $aK$ of the $a_{Kk}$'s, $n+1 \leq k \leq n+K$, exceed or equal $\theta \sigma_k^*$, because the $a_k$'s satisfy Condition A. As in [2], we define $\sigma_k^* = \max\{a_{Kk}: n+1 \leq k \leq n+K\}$, $h_k^*$, and $\sigma_k^*$, and by Condition A, we still have that $b_\lambda \leq h_k^* \sigma_k^* \leq B_\lambda$ and $d_\lambda \leq \sigma_k^* \leq D_\lambda$. It follows from Lemma 1 that

Lemma 2. $P_n(K, \lambda) = (2\pi)^{-1/2} (\sigma_k^* h_k^*)^{-1} \rho_{n,K}(\lambda)(1 + O(K^{-1/2}))$, where

$$\rho_{n,K}(\lambda) = \exp\left( -\lambda h_k^* \sum_{k=n+1}^{n+K} a_{Kk} \phi(h_k^* a_{Kk}) \right).$$
From Lemma 2, we obtain the large deviation result in the form needed for the proof of the main theorem:

**Lemma 3.** If \( C_{n,K}(\lambda) = (-K^{-1} \log \rho_{n,K}(\lambda))^{-1} \), then

\[
\frac{1 + O(K^{-1/2})}{(2\pi K)^{1/2} D_{\lambda} \beta_{\lambda}} e^{-K/C_{n,K}(\lambda)} \leq P_n(K, \lambda)
\]

\[
\leq \frac{1 + O(K^{-1/2})}{(2\pi K)^{1/2} (\alpha \theta^2)^{1/2} d_{\lambda} b_{\lambda}} e^{-K/C_{n,K}(\lambda)}.
\]

**Proof.** Because \( (\alpha \theta^2)^{1/2} K^{1/2} \leq (\sigma_{K}^*)^{-1} \leq K^{1/2} \), it follows that \( \sigma_{K}^* h_{K} \leq D_{\lambda} \beta_{\lambda} (\sigma_{K}^*)^{-1} \leq K^{1/2} D_{\lambda} \beta_{\lambda} \) and that \( \sigma_{K}^* h_{K} \leq d_{\lambda} b_{\lambda} (\sigma_{K}^*)^{-1} \leq K^{1/2} (\alpha \theta^2)^{1/2} d_{\lambda} b_{\lambda} \).

If we define \( L(t) = t Q(t) - \log \phi(t) \), Corollary (5.1) of [2] yields:

**Lemma 4.** \( \alpha \theta K \beta_t^*(\lambda) \leq -\log \rho_{n,K}(\lambda) \leq K \beta_t^*(\lambda) \) for all fixed \( n \) and sufficiently large \( K \), where \( \beta_t^*(\lambda) = \alpha L(\theta b_{\lambda}) > 0 \) and \( \beta_t^*(\lambda) = (\alpha \theta^2)^{-1} L(B_{\lambda}) < \infty \).

**Proof.** Examining the proof of Corollary (5.1) of [2], we see that

\[
\beta_t^*(\lambda) (\sigma_{K}^*)^{-1} \sum_{k=n+1}^{n+K} a_{Kk} \leq -\log \rho_{n,K}(\lambda) \leq \beta_t^*(\lambda) (\sigma_{K}^*)^{-1} \sum_{k=n+1}^{n+K} a_{Kk}.
\]

But \( \alpha \theta \sigma_{K}^* \leq \sum_{k=n+1}^{n+K} a_{Kk} \leq K \sigma_{K}^* \) by Condition A and the definition of \( \sigma_{K}^* \).

It immediately follows from Lemma 4 that:

**Lemma 5.** \( C'(\lambda) \leq C_{n,K}(\lambda) \leq C''(\lambda) \), where \( C'(\lambda) = \alpha \theta^2 / L(B_{\lambda}) \) and \( C''(\lambda) = (\alpha \theta^2 \theta L(\theta b_{\lambda}))^{-1} \).

The following lemmas depend on the fact that \( Q(t) = \phi'(t)/\phi(t) \) is monotonically increasing for \( 0 \leq t \leq B_{\theta}^* \), which is true because \( Q'(t) \) is the (strictly positive) variance of the "associated" distribution introduced in \( \S 1 \) of [2]. From Lemma (1.3) of [2], we know also that \( Q(t) \) is a one-to-one mapping from \( [0, B_{\theta}^*] \) to \( [0, Q(B_{\theta}^*]) \).

**Lemma 6.** \( \alpha \) is a monotonically decreasing function of \( \lambda \).

**Proof.** \( \beta_{\lambda} \) is monotonically increasing in \( \lambda \).

**Lemma 7.** \( C'(\lambda) \) and \( C''(\lambda) \) are monotonically decreasing in \( \lambda \).

**Proof.** \( L(B_{\lambda}) \) increases monotonically in \( \lambda \), as does \( L(\theta b_{\lambda}) \).

3. **Proof of the main theorem.** We are now ready to state and prove the extension of the Erdös-Rényi theorem to weighted sums. We have a sequence \( \{X_k: 1 \leq k < \infty\} \) of i.i.d. nondegenerate random variables with common m.g.f. \( \phi \), \( E(X_1) = 0 \), \( E(X_1^2) = 1 \), and a sequence \( \{a_k: 1 \leq k < \infty\} \) of
nonnegative real numbers, at least one nonzero. We have defined \( Q(t) = \phi(t)/\phi(t) \), \( L(t) = tQ(t) - \log \phi(t) \), \( B_\lambda = \theta^{-1}Q^{-1}(\lambda/\sqrt{\theta}) \), \( b_\lambda = Q^{-1}(\lambda\sqrt{\theta}) \), and

\[
\Sigma(N, K) = \max\{(A_{n+K} - A_n)^{-1}(S_{n+K} - S_n) : 0 \leq n \leq N - K\}.
\]

**THEOREM.** Under Conditions A, B, and C, there exist positive constants \( C'(\lambda) = \alpha\theta^2/L(B_\lambda) \) and \( C''(\lambda) = \{\alpha^2\theta L(\theta b_\lambda)\}^{-1} \) such that

\[
P\left( \limsup_{N \to \infty} \max_{[C''(\lambda)\log N]} \Sigma(N, [C'(\lambda)\log N]) \leq \lambda \leq \liminf_{N \to \infty} \Sigma(N, [C'(\lambda)\log N]) \right) = 1.
\]

**PROOF.** Setting \( \Sigma''(N, \lambda) = \Sigma(N, [C''(\lambda)\log N]) \) for notational ease, we show first that

\[
P(\limsup_{N \to \infty} \Sigma''(N, \lambda) \leq \lambda) = 1.
\]

Set \( \lambda'' = \lambda + \epsilon < \lambda^* \), where \( \epsilon > 0 \) is arbitrarily small. Then, with \( S''_{N, \lambda}(n) = (A_{n+K} - A_n)^{-1}(S_{n+K} - S_n) \) for \( K = [C''(\lambda)\log N] \) and \( 0 \leq n \leq N - K \),

\[
P(\Sigma''(N, \lambda) \geq \lambda'') = \frac{\max_{0 \leq n \leq N - K} S''_{N, \lambda}(n)}{\sum_{n=0}^{N-K} P(S''_{N, \lambda}(n) \geq \lambda'')} \leq \frac{N-K}{\sum_{n=0}^{N-K} B''K^{-1/2}e^{-K/C_{n,K}(\lambda'')}} \leq B''K^{-1/2}e^{-K/C_{n,K}(\lambda'')},
\]

by Lemma 3 for \( N \) sufficiently large, where \( B'' \) is a constant depending only on \( \lambda \) (and \( \lambda^* \) which is fixed). Continuing from Lemma 5,

\[
P(\Sigma''(N, \lambda) \geq \lambda'') \leq B''K^{-1/2}(N - K + 1)e^{-K/C''(\lambda'')}
\]

where, by Lemma 7, \( C''(\lambda)/C''(\lambda') = 1 + \delta_1 \) for \( \delta_1 > 0 \) depending only on \( \lambda \) and \( \epsilon \), not on \( N \), and \( B'' \) depends only on \( \lambda \) (and \( \lambda^* \) which is fixed). For all values of \( N \) such that \([C''(\lambda)\log N] = j \), we have \( N \geq \exp(j/C''(\lambda)) \). We define \( N_j \) to be the largest integer such that \([C''(\lambda)\log N_j] = j \). It follows that

\[
\sum_{j=1}^{\infty} P(\Sigma''(N_j, \lambda) \geq \lambda'') \leq B''N_j^{-\delta_1} \leq B''\sum_{j=1}^{\infty} (e^{-\delta_1/C''(\lambda)})^j < \infty.
\]

The Borel-Cantelli lemma then implies that the probability is one that \( \Sigma''(N_j, \lambda) < \lambda'' \) for all but a finite number of values of \( j \). But for each value of \( N \) between \( \exp(j/C''(\lambda)) \) and \( \exp((j+1)/C''(\lambda)) \), we have \( \Sigma''(N, \lambda) \leq \Sigma''(N_j, \lambda) \). Therefore the probability is one that \( \Sigma''(N, \lambda) < \lambda'' \) for all but finitely many \( N \). Equivalently, \( P(\limsup_{N \to \infty} \Sigma''(N, \lambda) < \lambda''\) = 1, and since \( \lambda'' = \lambda + \epsilon, \epsilon > 0 \) arbitrary, we obtain the first half of the theorem.

We next set \( \Sigma'(N, \lambda) = \Sigma(N, [C'(\lambda)\log N]) \), and show that

\[
P\left( \liminf_{N \to \infty} \Sigma'(N, \lambda) \geq \lambda \right) = 1.
\]
Set \( \lambda' = \lambda - \varepsilon \) for \( 0 < \varepsilon < \lambda \), \( \varepsilon \) arbitrarily small. Now with \( S'_{N, \lambda}(n) = (A_{n+K} - A_n)^{-1}(S_{n+K} - S_n) \) for \( K = \lfloor C'(\lambda) \log N \rfloor \) and \( 0 \leq n \leq N - K \), we have

\[
P(\Sigma'(N, \lambda) \leq \lambda') = P\left( \max_{0 \leq n \leq N-K} S'_{N, \lambda}(n) \leq \lambda' \right)
\]
\[
\leq P(S'_{N, \lambda}[r] \leq \lambda' \text{ for } 0 \leq r \leq K^{-1}N - 1)
\]
\[
\leq \prod_{r=0}^{K^{-1}N-1} P(S'_{N, \lambda}[r] \leq \lambda'),
\]
where \( S'_{N, \lambda}[r] = (A_{(r+1)K} - A_{rK})^{-1}(S_{(r+1)K} - S_{rK}) \), by independence of the increments. By Lemma 3, for \( N \) sufficiently large,

\[
P(S'_{N, \lambda}[r] \leq \lambda') \leq 1 - B'K^{-1/2}e^{-K/C\log N},
\]
where \( B' \) depends only on \( \lambda \) and \( \varepsilon \). By Lemmas 5 and 7,

\[
P(S'_{N, \lambda}[r] \leq \lambda') \leq 1 - B'_1(\log N)^{-1/2}e^{-C'(\lambda)\log N/C(\lambda')}
\]
\[
\leq 1 - B'_1(\log N)^{-1/2}N^{-(1-3\delta_2)}
\]
\[
\leq 1 - B'_1N^{-(1-3\delta_2)} \leq \exp(-B'_1N^{-(1-3\delta_2)}),
\]
where \( B'_1 \) and \( \delta_2 > 0 \) depend only on \( \lambda \) and \( \varepsilon \), not on \( N \). Therefore, for \( N \) sufficiently large,

\[
P(\Sigma'(N, \lambda) \leq \lambda') \leq \exp(-B'_2N^{-(1-2\delta_2)}K^{-1}N) \leq \exp(-B'_2N^{\delta_2}),
\]
where \( B'_2 \) and \( \delta_2 > 0 \) depend only on \( \lambda \) and \( \varepsilon \). The infinite series

\[
\sum_{N=1}^{\infty} \exp(-B'_2N^{\delta_2})
\]
converges by the integral test, so that the Borel-Cantelli lemma implies that, with probability one, only finitely many of the events \( \Sigma'(N, \lambda) \leq \lambda' \) occur. Therefore \( P(\liminf_{N \to \infty} \Sigma'(N, \lambda) > \lambda') = 1 \). Since \( \lambda' = \lambda - \varepsilon \), where \( \varepsilon > 0 \) is arbitrarily small, the second half of the theorem follows.

4. **Reduction to the Erdös-Rényi theorem.** The theorem of the previous section reduces to the assertion of the Erdös-Rényi theorem, for absolutely continuous random variables, if we take all \( a_k = 1 \). Then Condition A holds with \( a = \theta = 1 \), so that \( b_k = Q^{-1}(\lambda) = B_\lambda \) and \( C'(\lambda) = 1/L(Q^{-1}(\lambda)) = C''(\lambda) \). If we define \( C(\lambda) \) to be the common value of \( C'(\lambda) \) and \( C''(\lambda) \), the assertion of the theorem becomes \( P(\lim_{N \to \infty} \Sigma(N, [C(\lambda)\log N]) = \lambda) = 1 \). If we set \( \tau = Q^{-1}(\lambda) \), then \( \phi'(\tau)/\phi(\tau) = \lambda \), and so \( C(\lambda) = (\log \rho(\lambda))^{-1} \), where \( \rho(\lambda) = e^{-\tau\phi(\tau)} = \inf\{e^{-\lambda t}\phi(t) : t > 0\} \).

5. **Acknowledgments.** The author would like to express his gratitude to Professor C. M. Deo, whose courses in probability theory at the University
of Oregon inspired this work. He would also like to thank Professor D. R. Truax of the University of Oregon, for conjecturing the worth of the function $L(t)$ in connection with the large deviation theorem.

**References**


**Department of Mathematics, California State College, Dominguez Hills, California 90246**